

LMAP Working Group
Internet-Draft
Intended status: Informational
Expires: January 17, 2014

A. Akhter
P. Aitken
Cisco Systems
July 16, 2013

A Framework and Inventory for a Large Scale Measurement System
draft-akhter-lmap-framework-00.txt

Abstract

This LMAP framework document reviews the LMAP Working Group charter, considers the necessary building blocks, and looks at what we already have in the IETF and what's missing, so that LMAP Working Group attention can be focused on where the gaps are.

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <http://datatracker.ietf.org/drafts/current/>.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on January 17, 2014.

Copyright Notice

Copyright (c) 2013 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1. Introduction	3
1.1. Working Group Scope	4
1.2. Out of Scope	5
1.3. LMAP Working Group Goals	5
2. Terminology	6
3. Architecture	7
3.1. Measurement Agent	8
3.1.1. Measurement Agent Embedded in Site Gateway	9
3.1.2. Measurement Agent behind Site NAT/Firewall	9
3.1.3. Measurement Agent in-line with Site Gateway	9
3.1.4. Measurement Agent in Multi Homed Site	10
3.2. Remote Measurement Test Target	11
3.3. Controller	11
3.4. Collector	12
3.5. Information Model	12
3.6. Transport Protocols	13
3.7. Scaling	14
3.8. Device Discovery	14
4. Active Measurements	15
4.1. What building blocks exist today?	15
4.1.1. Single Sided Client Tests	15
4.1.2. OWAMP - One Way Active Measurement Protocol	15
4.1.3. TWAMP - Two Way Active Measurement Protocol	17
4.1.4. Cisco Service-Level Assurance Protocol	18
4.1.5. IPPM Performance Metrics	18
4.2. Missing building blocks	18
4.2.1. Time Synchronization	18
4.2.2. Shared Secret Distribution	19
4.2.3. NAT/Firewall Traversal for Control and Test Protocols	19
4.2.4. IPPM Metrics Registry	19
4.2.5. OWAMP/TWAMP configuration	19
5. Passive Measurements	20
5.1. What building blocks exist today?	20
5.1.1. Measuring Packets	20
5.1.2. Measuring Flows	20
5.1.3. Defining new Information Elements	22
5.1.4. Exporting Process	22
5.1.5. Mediation	23
5.1.6. Configuration	23
5.2. Missing building blocks	23
5.2.1. Performance metrics definition in the IPFIX registry	23
5.2.2. Mediation Configuration	24
6. LMAP: Standards Re-usability	24
6.1. Existing Building Blocks	24
6.2. Missing Building Blocks	24
6.2.1. Task Definitions	24

6.2.2. Instructions Setup	25
6.2.3. Task Scheduling	26
6.2.4. Combining Active and Passive Measurements	26
7. Security considerations	27
8. IANA Considerations	28
9. Acknowledgements	28
10. References	28
10.1. Normative References	28
10.2. Informative References	28
Authors' Addresses	32

1. Introduction

There is a desire to be able to coordinate the execution of broadband measurements and the collection of measurement results across a large scale set of diverse devices. These devices could be software based agents on PCs, embedded agents in consumer devices (e.g. blu-ray players), service provider controlled devices such as set-top players and home gateways, or simply dedicated probes. It is expected that such a system could easily comprise 100k devices. Such a scale presents unique problems in coordination, execution and measurement result collection. Broad users of such a system include governmental regulators looking for service compliance; network and service operators (including over the top content providers) for diagnostics, compliance and planning; and end users for diagnostics and service compliance. The various detailed uses of such a large scale measurement system are covered in [I-D.linsner-lmap-use-cases].

Over the years various efforts inside and outside the IETF have worked on independent components of such a system. There are also existing systems that are deployed today. However, these are either proprietary, closed, and/or not standardized. The IETF Large-Scale Measurement of Broadband Performance (LMAP) Working Group is chartered to specify the information model, associated data models, and select/extend one or more protocols for secure measurement control and measurement result collection.

With standardization, LMAP compliant Measurement Agents will be more pervasive in gateways and end systems and offer a base common service across vendor implementations.

A set of Measurement Agents is to be controlled by a single organization. The Measurement Agents do not coordinate with Measurement Agents under the control of other organisations. While some of the capabilities are meant for end users, an end user is not meant to directly control Measurement Agents, except for his own Measurement Agent. The end user may interact with a service provider portal to schedule and execute a measurement task using the Measurement Agent on their premises.

The measurements themselves may be on IPv4, IPv6, and on various services (DNS, HTTP, XMPP, FTP, VoIP, etc.). The Measurement Agents may have multiple interfaces (WiFi, Ethernet, DSL, fiber, etc.) and the measurements may specify any one of these. The measurement tasks may generate synthetic traffic to perform the measurement (active measurement), only observe existing traffic (passive measurement), or may do a combination of both active and passive measurement.

Given the usage of passive measurements (and even in the case of active measurement) there are valid concerns regarding privacy of the measurement results and any user identifiable information.

1.1. Working Group Scope

The Large-Scale Measurement of Broadband Performance (LMAP) Working Group is chartered to standardize the LMAP measurement system for performance measurements of broadband access devices.

The Working Group is chartered to specify an information model, the associated data models, and select/extend one or more protocols for secure communication:

- o A Control Protocol, from a Controller to instruct Measurement Agents what performance metrics to measure, when to measure them, and when and how to report the measurement results to a Collector.
- o A Report Protocol, for a Measurement Agent to report the results to the Collector.

The data models should be extensible for new and additional measurements. LMAP will consider re-use of existing data models languages.

The LMAP architecture will allow for measurements that utilize either IPv4 or IPv6, or possibly both. Devices containing Measurement Agents may have several interfaces using different link technologies. Multiple address families and interfaces must be considered in the Control and Report protocols.

Both active and passive measurements are in scope, although there may be differences in their applicability to specific use cases, or in the security measures needed according to the threats specific to each measurement category. LMAP will not standardize performance metrics.

The LMAP Working Group will consider privacy as a core requirement and will ensure that by default measurement and collection mechanisms and protocols operate in a privacy-sensitive manner, ie that privacy features are at least well-defined.

1.2. Out of Scope

There are a number of items that are currently explicitly out of scope for the LMAP Working Group:

- o Inter-organization coordination and sharing of results is out of scope
- o Discovery of service parameters on Measurement Agents is out of scope
- o Sharing the service parameters between Measurement Agents is out of the scope.
- o Decision on the set of measurements to run is out of scope
- o Protection against intentional / malicious gaming is out of scope
- o Standardizing control of end users Measurement Agents is out of scope.
- o The management protocol to bootstrap the Measurement Agents in measurement devices is out of scope.

1.3. LMAP Working Group Goals

The LMAP Working Group will produce the following work items:

1. The LMAP Framework - provides common terminology, basic architecture elements, and justifies the simplifying constraints
2. The LMAP Use Cases - provides the motivating use cases as a basis for the work

3. Information Model, the abstract definition of the information carried from the Controller to the Measurement Agent and the information carried from the Measurement Agent to the Collector. It includes:
 - * The metric(s) that can be measured and values for its parameters such as the Peer Measurement Agent participating in the measurement and the desired environmental conditions (for example, only conduct the measurement when there is no user traffic observed)
 - * The schedule: when the measurement should be run and how the results should be reported (when and to which Collector)
 - * The report: the metric(s) measured and when, the actual result, and supporting metadata such as location. Result reports may be organized in batches or may be reported immediately, such as for an on-demand measurement.
4. The Control protocol and the associated data model: The definition of how instructions are delivered from a Controller to a Measurement Agent; this includes a Data Model consistent with the Information Model plus a transport protocol. This may be a simple instruction - response protocol, and LMAP will specify how it operates over an existing protocol (to be selected, perhaps REST-style HTTP(s) or NETCONF).
5. The Report protocol and the associated data model: The definition of how the Report is delivered from a Measurement Agent to a Collector; this includes a Data Model consistent with the Information Model plus a transport protocol (to be selected, perhaps REST-style HTTP(s) or IPFIX).

2. Terminology

Terms used in this document and are to be interpreted as defined in the following documents:

- o LMAP terms used in this document are defined in [I-D.eardley-lmap-terminology].
- o IPFIX terms are defined in the Terminology section of the IPFIX Architecture [RFC5470]
- o PSAMP terms are defined in the Terminology section of the PSAMP Protocol [RFC5476]
- o TODO: where are IPPM terms defined?

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].

3. Architecture

A large scale measurement system is composed of several basic parts:

- o Measurement Agents
- o Remote Measurement Test Target(s)
- o Controller(s)
- o Collector(s)

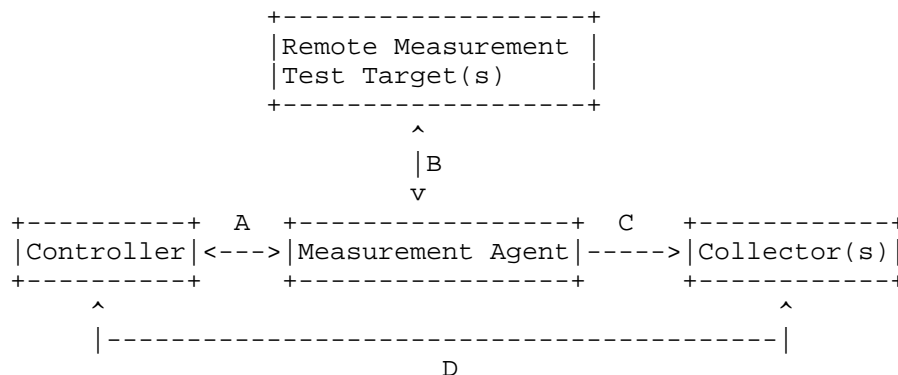


Figure 1: LMAP Basic Parts

A full system would include other components such as a data parsing module, report generation module, and subscriber database module. However, these are not covered by the high level Figure 1.

There is also the concept of the measurement task and the measurement result.

A measurement task generates a measurement result, which is composed of one or many metrics as well as supplemental information from the process of the task. A measurement task might time the download of a specific web page. The measurement results might include the DNS query time, query result used (IP address), the time to download the web page and individual times for associated objects, the maximum bit rate, and average bit rate. Note that some of the results are derived metrics (based on other measurements, like maximum bit rate),

while others may not be not metrics at all (IP address used). Also note that it is possible that the size of the result is not known at the time of measurement task scheduling - the number of objects on the web page might change, or if the measurement task includes a traceroute the path will be different from many of the Measurement Agents.

The measurement task is scheduled by the Controller via an instruction.

3.1. Measurement Agent

The Measurement Agent is the component that is responsible for executing a test. The Measurement Agent could take a number of forms: a dedicated probe, software on a PC, embedded into an appliance, or even embedded into a gateway. A single site (home, branch office etc.) that is participating in a test could make use of one or multiple Measurement Agents in a single measurement test. e.g., if there are multiple output interfaces, there might be a Measurement Agent per interface.

The Measurement Agent's configuration (specifically which Controller to initially connect to), is out of scope within LMAP. However, depending on the type of probe, it could be manually configured by the user, pre-configured before shipment to the end user, or configured by the application (in the case of some PC based Measurement Agents). For example, a Measurement Agent that is included in the app for a content provider might be configured automatically by the content provider to use the content provider's LMAP Controller. That said, there should be an element of local premises configuration that allows the Measurement Agent (especially in the case of active measurements) to mimic performance of user applications at the same site. For example, making use of the same DNS server as the remainder of the site.

The Measurement Agent could be deployed in a variety of locations. Not all deployment locations are available to every kind of Measurement Agent operator. There are also a variety of limitations and trade-offs depending on the final placement. The next sections outline some of the locations a Measurement Agent may be deployed. This is not an exhaustive list and combinations of the below may also apply.

3.1.1. Measurement Agent Embedded in Site Gateway

A Measurement Agent embedded with the site gateway (e.g. in the case of a branch office in a managed service environment) is one of better places the Measurement Agent could be deployed.

All site to ISP traffic would traverse through the gateway and passive measurements could easily be performed. Similarly, due to this user traffic visibility, an active measurement task could be rescheduled so as not to compete with user traffic.

Generally NAT and firewall services are built into the gateway, allowing the Measurement Agent the option to offer its Controller facing management interface outside of the NAT/firewall. This placement of the management interface allows the Controller to unilaterally contact the Measurement Agent for instructions.

However, if the site gateway is owned and operated by the service provider, the Measurement Agent will generally not be available for over the top providers, the regulator, end users or enterprises.

3.1.2. Measurement Agent behind Site NAT/Firewall

The Measurement Agent could also be embedded behind a NAT, a firewall, or both. In this case the Controller may not be able to unilaterally contact the Measurement Agent unless either static port forwarding configuration or firewall pin holing is configured. This would require user intervention, and ultimately might not be an option available to the user (perhaps due to permissions).

The Measurement Agent may originate a session towards the Controller and maintain the session for bidirectional communications. This would alleviate the need to have user intervention on the gateway, but would reduce the overall scalability of the Controller as it would have to maintain a higher number of active sessions.

That said, sending keepalives to prop open the firewall could serve a dual purpose in testing network reachability for the Measurement Agent.

An alternative would be to use a protocol such as UPnP or PCP [RFC6887] to control the NAT/firewall if the gateway supports this kind of control.

3.1.3. Measurement Agent in-line with Site Gateway

As mentioned earlier, there are benefits in the Measurement Agent's ability to observe the site's user traffic. In the case of active

measurement it allows the Measurement Agent to back-off on a potentially disruptive measurement task to avoid impacting the user. For the case of passive measurement, access to the user traffic allows the Measurement Agent to gather data without a traffic footprint (of interest to both the site user and network operator) as well as potentially provide a greater number of samples for a measurement task.

A Measurement Agent behind the gateway would generally not be privy to observation of the user traffic unless the Measurement Agent was placed in-line with the site gateway or the site gateway traffic was replicated to the Measurement Agent (a capability generally not found in home broadband gateways).

3.1.4. Measurement Agent in Multi Homed Site

A broadband site may be multi-homed. For example, the site may be connected to multiple broadband ISPs (perhaps for redundancy or load-sharing), or have a broadband as well as mobile/WiFi connectivity. It may also be helpful to think of dual stack IPv4 and IPv6 broadband sites as multi-homed.

In these cases, there needs to be clarity on which network connectivity option is being measured. Sometimes this is easily resolved by the location of measurement agent itself. For example, if the measurement agent is built into the gateway (and the gateway only has a single WAN side interface), there is little confusion or choice. However, for multi-homed gateways or devices behind the gateway(s) of multi-homed sites it would be preferable to explicitly select the network to measure (e.g. [RFC5533]) but the network measured should be included in the Measurement Result.

Section 3.2 of [I-D.ietf-homenet-arch] describes dual-stack and multi-homing topologies that might be encountered in a home network (which is generally a broadband connected site). The Multiple Interfaces (mif) working group covers cases where hosts are either directly attached to multiple networks (physical or virtual) or indirectly (multiple default routers, etc.). [xref target="RFC6419"/>](#) provides the current practices of multi-interfaces hosts today. As some of the end goals of a LMAP Measurement Agent is to replicate the network experience as an end user would, it is important to understand the current practices.

3.2. Remote Measurement Test Target

A remote measurement test target is the other side of the measurement test - the test target of the Measurement Agent. The remote measurement test target could also take many different forms: a web site, a service (VoIP), a DNS server, an application specific server (e.g., webex), a well known web site (e.g., youtube, Google Search), another Measurement Agent in another home, a powerful Measurement Agent that is well network connected (Anchor Measurement Agent), or even a collection of home based Measurement Agents.

An Anchor Measurement Agent is a remote measurement test target that is well placed bandwidth-wise and is meant to handle test traffic in a highly scaled (1000s of test sessions) environment. Similar to the measurement agent sitting at a broadband site, it is under the direction of an LMAP Controller, but might support multi-tenancy. It is generally expected to respond to broadband site Measurement Agents rather than initiate tests.

As illustrated in Figure 2, a measurement task may not only involve a similar LMAP Measurement Agent, but multiple such Measurement Agents. An example where this arrangement would be useful is when an Anchor Measurement Agent in a path capacity measurement is unable to saturate a path, while horizontal scaling properties of multiple Measurement Agents can. This arrangement also alleviates any one remote Measurement Agent from saturating its own access link as the load is distributed.

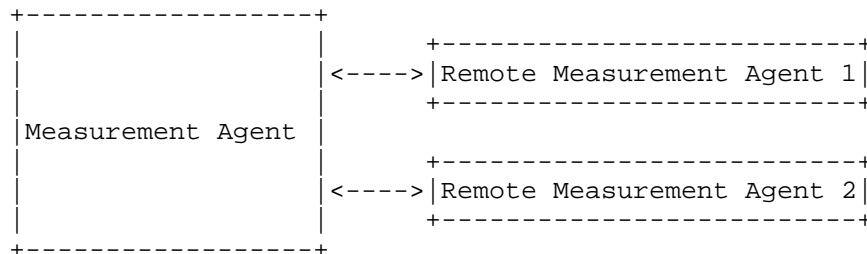


Figure 2: Measurement Task Involving Multiple Remote Measurement Agents

3.3. Controller

A Controller is responsible for providing the Measurement Agent with instructions which include the test schedule, test parameters, etc. It is basically the entity controlling the Measurement Agents in a LMAP domain.

For scaling purposes there may be several Collectors as well as several Controllers, perhaps regionally located. A large scale test making use of multiple Controllers would need a master Controller that is the ultimate source of direction.

3.4. Collector

A Collector is responsible for receiving the test results from the Measurement Agent at the end of a test. It may have additional features such as aggregating the results across multiple Measurement Agents, remove outliers, create additional statistics, (depending on usage of data) anonymization of results for privacy reasons (if not done already in the Measurement Agents) etc. The work of anonymization of user identifiable data has been addressed for IPFIX via RFC6235 [RFC6235].

For scaling purposes there may be several Collectors as well as several Controllers, perhaps regionally located. A large scale test making use of multiple Collectors would need to aggregate/consolidate their results for the complete picture.

3.5. Information Model

For definitions and examples of Information Models and Data Models, refer to [RFC3444]

The information shared between LMAP devices would be organized into an LMAP information model [I-D.burbridge-lmap-information-model] covering:

- o Controlling the Measurement Agent (from the Controller)
- o The Measurement Agent submitting the results (to the Collector)

In some cases, the Collector and Controller could be co-resident on the same device but the information models would continue to be separate.

The IETF IPFIX working group has defined an extensible information model in [I-D.ietf-ipfix-information-model-rfc5102bis] which could be used to organize the result metrics back to the LMAP Collector.

3.6. Transport Protocols

The information shared between the components that is in the information model needs a transport protocol. Similar to the information model the transport protocols would map to the following main functions:

- o Control of the Measurement Agent by the Controller
- o Submission of measurement test results to the Collector
- o Controller to Collector Test Configuration synchronization

Note that each of these could use different transport protocols. However, for implementation simplification and keeping a small memory footprint having the option of a single transport protocol can be helpful.

The Controller to Controller Test Configuration Synchronization provides a direct way for the Controller to communicate test configuration information to the Collector. An alternate would have been for the Measurement Agent to echo back the configuration to the Controller. However, given several hundred thousand reports this would be much duplication of data. It would be optimal to transfer the test identification and configuration information directly to the Collector(s) a single time. In this case, the Controller to Collector Test Configuration Synchronization would be no different in communications that with a Measurement Agent, except the Collector would not execute the test-- it would simply understand it. Explicit Collector directives (if they exist) by the Controller should not be sent via the Measurement Agent.

The collated results from the Collector would have a pre-configured path to publication or data storage.

To reduce the complexity and memory footprint needs of the Measurement Agent it is possible that the control and report protocol are the same (but still using the independent information models).

There are a number of transport candidate transport protocols. Depending on the placement and use of the Measurement Agent certain transport protocols may be preferred over others. For example if the Measurement Agent is behind a NAT or firewall, it would be difficult to make use of SNMP, or for the Controller to connect to the Measurement Agent if REST were used. Regardless of transport protocol used, the information model MUST be consistent.

3.7. Scaling

Scalability is a key issue, since LMAP is expected to scale to 10,000's of Measurement Agents. Therefore the architecture shown in Figure 1 may include a hierarchy of Controllers and a hierarchy of Collectors, e.g. with sub-controllers and sub-collectors distributed topographically or geographically. Note that sub-collectors are effectively IPFIX mediators [RFC6183].

Separating the Control Protocol and Report Protocol allows these hierarchies to scale independently, which would not be possible with a single command-response protocol which requires a co-hosted Controller/Collector implementation.

A scalability optimisation is discussed in Section 6.2.2.

3.8. Device Discovery

In a large-scale system, an LMAP controller must somehow discover which Measurement Agents are available in the LMAP domain, and which is the most appropriate collector for these agents to report test results to. ie, for each LMAP Measurement Agent, which LMAP domain is it in?

Possibilities include:

- o The call home mechanism from netconf [I-D.ietf-netconf-reverse-ssh]
- o DNS SRV
- o DNS anycast, selecting the nearest
- o ALTO
- o DHCP

Also note one corner case: a Collector may have to ignore test results from a Measurement Agent which is mis-configured, or which has been moved from one LMAP domain to another. ie, where the test results are being reported out of scope.

4. Active Measurements

4.1. What building blocks exist today?

4.1.1. Single Sided Client Tests

A good number of active measurement tasks simply require that the Measurement Agent perform client side duties and interact with a Remote Measurement Test Target as a general user application would have. Examples of single sided client tests include HTTP GETs from private or public servers, DNS queries, FTP transfers etc. The metrics are generally based on response time, bit rate of transfer etc.

There are no requirements against the server and in fact it is likely that the server might even be under the operational control of an entirely different entity. Generally the servers in this will be providing a user side function (a new website, DNS services, etc.) so care must be taken in running too many synthetic tests as Denial of Service (DoS) may be achieved or (more likely) automated DoS protection mechanisms may come into play ultimately rejecting traffic from that broadband site.

4.1.2. OWAMP - One Way Active Measurement Protocol

The One Way Active Measurement Protocol [RFC4656] allows for the generation of a unidirectional test stream (by the Session-Sender) and the measurement of that test stream by a remote entity (Session-Receiver). The test stream is known as OWAMP-Test. The OWAMP-Control protocol is used to (at the time of the test) negotiate OWAMP-Test port numbers (for example UDP or TCP port numbers) between the Session-Sender and Session-Receiver. As the test stream in OWAMP is unidirectional the Session-Receiver has to compute the metrics. The OWAMP-Control protocol is then used to retrieve the metrics. Depending on the metrics being computed, it may be necessary for the Session-Sender and Session-Receiver to be time synchronized. The one-way-latency measurement is an example of this case. The OWAMP-Control protocol is also used to convey from the Session-Sender to the Session-Receiver any packets that were unable to be sent so the Session-Receiver does not mistakenly count these non-existent packets as loss.

Figure 3 describes the individual roles and relationships in OWAMP. Any unlabeled links are unspecified by OWAMP and may be proprietary protocols.

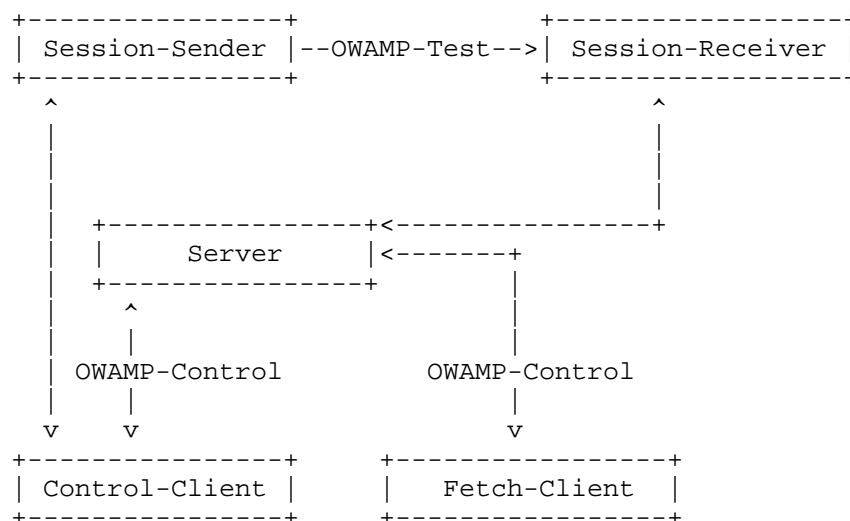


Figure 3: OWAMP Individual Roles and Relationships

Figure 4 describes simplified individual roles and relationships in OWAMP such that only two hosts are used.

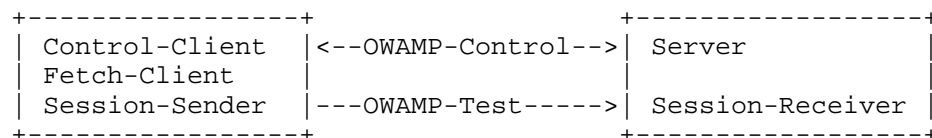


Figure 4: OWAMP Two Host Implementation

For the cases of many IPPM defined metrics the OWAMP is a natural fit and the OWAMP-Test stream could certainly be utilized between two Measurement Agents. The Session-Sender would be the local broadband site Measurement Agent, while the Session-Receiver would be a Remote Measurement Test Target, perhaps an anchor Measurement Agent or a Measurement Agent at a broadband site.

In the case that Session-Receiver/Server is behind a firewall, it would be challenging for the OWAMP-Control protocol to reach the OWAMP Server. The usage of PCP by the Session-Receiver/Server might be utilized. Another solution would be for the LMAP Controller to take on the role of the OWAMP server-- with the understanding that the LMAP server has open lines of communication to all Measurement Agents. The case of the OWAMP-Test stream would also be challenged in crossing such a firewall. The OWAMP-Test transport ports are dynamically negotiated to prevent special handling by the underlying

network. The LMAP Controller line of communication may be of little help in this case and other techniques would need to be used.

The OWAMP-Control protocol provides an authenticated control channel that prevents unauthorized usage (and thereby conserving test resources and bandwidth) as well as tampering with the results as they are fetched from the Session-Receiver. Additionally, encryption is also offered to prevent a third-party from improving the results that reality. If authentication and encryption is to be used in an LMAP scenario, the shared-secret would need to be deployed to both the Session-Sender and the Session-Receiver.

4.1.3. TWAMP - Two Way Active Measurement Protocol

The Two-Way Active Measurement Protocol [RFC5357] builds on top of the OWAMP work to allow two-way active measurement. In TWAMP, the Session-Receiver becomes a Session-Reflector that does not keep state for the test metric and simply reflects back the received test stream (with some additions and modifications to account for the reverse direction trip. The metric computation is performed at the Session-Sender.

Figure 5 describes the individual roles and relationships in TWAMP. Note that due to the Session-Reflector, the diagram is simplified compared to OWAMP. Any unlabeled links are unspecified by OWAMP and may be proprietary protocols.

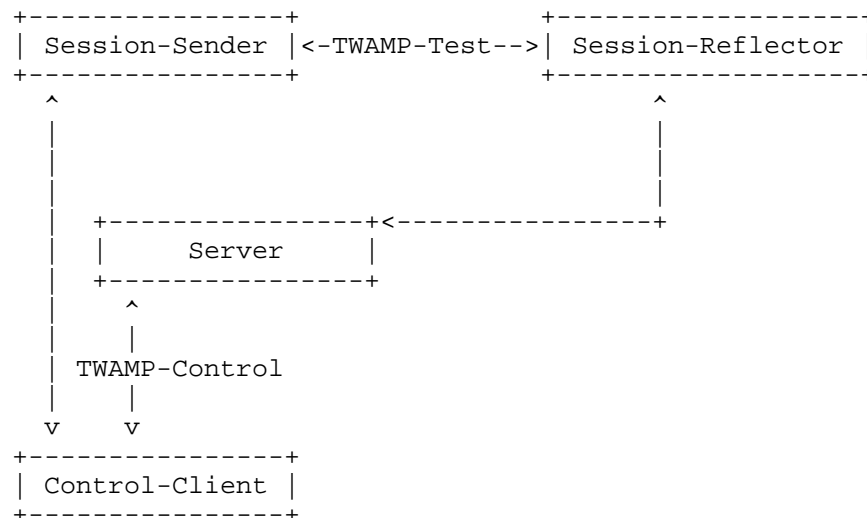


Figure 5: TWAMP Individual Roles and Relationships

Figure 6 describes simplified individual roles and relationships in TWAMP such that only two hosts are used.

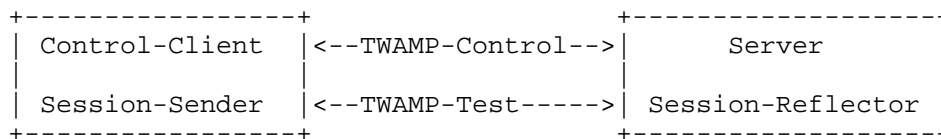


Figure 6: TWAMP Two Host Implementation

For the cases of many IPPM defined metrics the TWAMP is a natural fit and the TWAMP-Test stream could certainly be utilized between two Measurement Agents. The Session-Sender would be the local broadband site Measurement Agent, while the Session-Reflector would be a Remote Measurement Test Target, perhaps an anchor Measurement Agent or a Measurement Agent at a broadband site.

In the case that Session-Reflector/Server is behind a firewall, the same challenges for TWAMP-Control and TWAMP-Test as described for OWAMP earlier would apply to TWAMP.

4.1.4. Cisco Service-Level Assurance Protocol

The Cisco Service Level Assurance Protocol [RFC6812] is similar to OWAMP and TWAMP in that there is a control phase that negotiates transport port usage and a measurement phase. The issues described previously to remote test targets situated behind a firewall would continue to apply to CSLA.

4.1.5. IPPM Performance Metrics

A good number of performance methodologies and metrics exist today and have been defined via various works by the IETF IPPM working group. Recently, guidelines have been published for new performance metrics in [RFC6390]. These guidelines are applied by the Performance Metrics Directorate [pm-dir] when reviewing new metrics.

4.2. Missing building blocks

4.2.1. Time Synchronization

A variety of the metrics require the time synchronization to a common clock. This time synchronization is not guaranteed, especially at broadcast sites. Given that the Measurement Agents are communicating to a common set of Controller(s) this should present an opportunity to provide a fall-back common clock. In this particular case it may not be in the best interest of the test to use broadband site local

NTP server configuration as the disparate Measurement Agents might be on different NTP hierarchies.

4.2.2. Shared Secret Distribution

A secured control protocol and test stream between the Measurement Agents requires the distribution of a shared key. Such a shared key might be distributed by the Controller or by the Measurement Agent provisioning system.

4.2.3. NAT/Firewall Traversal for Control and Test Protocols

A NAT or firewall in between the Measurement Agents can become problematic. In this case the traditional method of control communications between the Measurement Agents would be impaired. Whereas the control protocols are on well known ports, the test streams are on negotiated port values. In this case, the test traffic may need to also be well known port values. There may be alternative mechanisms to reach the Measurement Agent behind the firewall such as via the Controller line of communication or the use of PCP.

4.2.4. IPPM Metrics Registry

The IPPM WG defined an IPPM Metrics Registry [RFC4148] . However this was obsoleted by [RFC6248] as the registry was found to be insufficiently detailed to uniquely identify IPPM metrics. Calls to the community regarding the registry were unanswered in 2010.

Such a registry (and the unique identification issue resolved) will be needed to by the LMAP system as the controller needs to designate which test (or to be more precise, which metric within a test) is to be run by the measurement agent. There's currently no IPPM metrics registry since [RFC4148] was obsoleted by [RFC6248]. Proposals for such a registry can be found at [I-D.claise-ippm-perf-metric-registry-00], [I-D.bagnulo-ippm-new-registry], and [I-D.bagnulo-ippm-new-registry-independent]. The first provides a simplified model, taking into account PMOL [RFC6390] and the IPFIX Information Model [I-D.ietf-ipfix-information-model-rfc5102bis]. The second has a single registry with sub-registries while the third proposes a more distributed registry for the components involved. As this new registry is created it would be extremely helpful that the metrics added to the registry confirm to the performance metrics guidelines outlined in [RFC6390].

4.2.5. OWAMP/TWAMP configuration

Currently there are no standardised way to configure OWAMP and TWAMP. An information model is required. A YANG module (data model) would be a plus, if NETCONF is chosen as the LMAP Configuration Protocol.

5. Passive Measurements

5.1. What building blocks exist today?

5.1.1. Measuring Packets

The PSAMP Framework [RFC5474] specifies a framework for Packet Sampling ("PSAMP").

The PSAMP protocol [RFC5476] selects packets from a stream according to a set of standardized selectors, forms a stream of reports on the selected packets, and exports the reports to a Collector. The PSAMP Framework [RFC5474] defines packet selection processes, with various types of filtering and sampling. It defines the exporting process and packet reports.

The architecture shown in section 3.1 of the PSAMP Framework [RFC5474] corresponds well to the LMAP architecture discussed in Section 3 above. The PSAMP Metering Process corresponds to LMAP Measurement Agents; the PSAMP protocol corresponds to the LMAP Report Protocol; the PSAMP Collector corresponds to the LMAP Collector.

[RFC5477] defines an Information Model for Packet Sampling Exports which is used by the PSAMP protocol for encoding sampled packet data and information related to the sampling process. This includes confidence intervals, measurement error, and observation timestamps.

In PSAMP, a Selector ID identifies a Primitive Selector, and a Selection Sequence ID identifies a combination of Selectors. LMAP should follow a similar model, using a global ID to identify a complex test built up from a set of test primitives.

5.1.2. Measuring Flows

The IPFIX Metering Process defined in [RFC5470] is designed to meter flows, which are defined as:

A Flow is defined as a set of IP packets passing an Observation Point in the network during a certain time interval. All packets belonging to a particular Flow have a set of common properties.

Inserting IPFIX terminology into Figure 1 above gives the architecture shown in Figure 7:

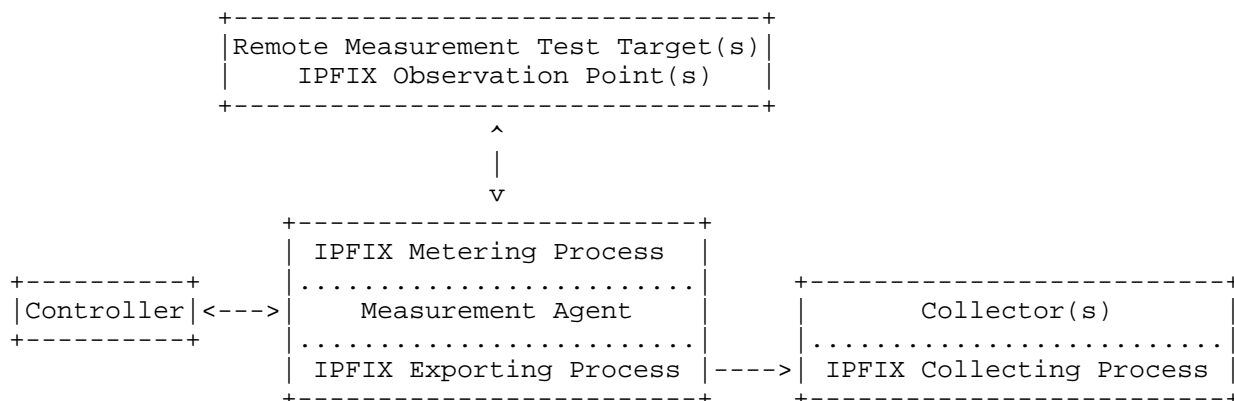


Figure 7: LMAP IPFIX Architecture

The only component of the LMAP architecture which doesn't have a parallel in IPFIX is the LMAP Controller. Therefore the IPFIX Architecture is clearly a key component for passive measurements in an LMAP Measurement Agent.

Inputs to the IPFIX Metering Process are packet headers, packet characteristics, and packet treatment. The Metering Process consists of a set of functions that includes packet header capture, timestamping, sampling, classifying, and maintaining Flow Records.

A packet belongs to a Flow if it completely satisfies all the defined properties of the Flow. This definition covers the range from a Flow containing all packets observed at a network interface to a Flow consisting of just a single packet between two applications. It includes packets selected by a sampling mechanism.

[I-D.ietf-ipfix-protocol-rfc5101bis] specifies the IPFIX Protocol which transmits information from the IPFIX Metering Process over the network, from an Exporting Process to a Collecting Process. The Protocol defines a common data representation and a standard means of communicating information over a number of transport protocols from an Exporting Process to a Collecting Process.

The IPFIX protocol is a candidate for the LMAP Report Protocol between Measurement Agents and Collectors.

[I-D.ietf-ipfix-information-model-rfc5102bis] defines the Information Model for IPFIX export, for which IANA's IPFIX registry [iana-ipfix-assignments]) is now the normative reference. The information model encodes measured traffic information and information related to the traffic Observation Point, the traffic

Metering Process, and the Exporting Process - ie, both details of the measured traffic and metadata about the Measurement Agent.

Although developed for the IPFIX Protocol, the information model is defined in an open way that readily allows it to be used in other protocols and applications. The information model is maintained as an IANA registry [iana-ipfix-assignments]).

[I-D.ietf-ipfix-ie-doctors] provides guidelines for how define new IPFIX Information Elements. It provides instructions on using the proper conventions for Information Elements to be registered in the IANA IPFIX Information Element registry, and provides guidelines for expert reviewers to evaluate new registrations.

[RFC6759] specifies an extension to the IPFIX information model to export application information, including application ID, name, description, and classification which would be useful if the test needs to be run, or test results must be reported, per application.

5.1.3. Defining new Information Elements

The IPFIX Information Model defined in [I-D.ietf-ipfix-information-model-rfc5102bis] is extensible: new elements may be defined by following the process defined in [I-D.ietf-ipfix-ie-doctors]. New Information Elements may be registered in IANA's IPFIX Information Element registry, or may be enterprise specific.

The IPFIX protocol supports export of both standard Information Elements (as defined in IANA's IPFIX registry [iana-ipfix-assignments]), and enterprise-specific Information Elements which allows non-standard (ie, proprietary) information to be carried in the protocol.

This extensibility allows new information to be carried in the IPFIX protocol without any modification to the underlying protocol.

5.1.4. Exporting Process

An IPFIX Exporting Process [RFC5470] transmits information generated by one or more IPFIX [RFC5470] or PSAMP [RFC5474] Metering Processes to one or more Collecting Processes. IPFIX export is specified over SCTP, UDP, and TCP, with authentication and security.

In LMAP terms, the Exporting Process uses the Reporting Protocol to transmit test information from Measurement Agents to the Collector.

5.1.5. Mediation

The sharing of information for monitoring applications having different requirements raises issues in terms of measurement system scalability, measurement flexibility, and export reliability which are described in [RFC5982]. Mediation fills the gap between restricted metering capabilities and the requirements of measurement applications by introducing an intermediate device called the Mediator.

[RFC6183] describes a framework for IPFIX Mediation. It introduces a generalized concept for intermediate entities, describes the high-level Mediation architecture, key architectural components, and mediation characteristics.

Mediation could be anonymization [RFC6235], aggregation [I-D.ietf-ipfix-a9n], or flow selection [I-D.ietf-ipfix-flow-selection-tech]. Removing user identifiable information eg by aggregation is especially important for passive measurements.

Aggregation is needed in the ISP use case, when the ISP needs to report the information to the regulator.

Note that IPFIX is required to report the output of any mediation function, possibly with stricter rules to support LMAP.

5.1.6. Configuration

[RFC6728] specifies the Configuration Data Model for IPFIX and PSAMP exporting and metering process configuration, and for Collecting Processes. The model is specified using YANG [RFC6020]. The configuration data is encoded in Extensible Markup Language (XML).

YANG is a data modeling language used to model configuration and state data manipulated by the Network Configuration Protocol (NETCONF), NETCONF remote procedure calls, and NETCONF notifications.

5.2. Missing building blocks

5.2.1. Performance metrics definition in the IPFIX registry

IANA's IPFIX Information Elements registry [iana-ipfix-assignments]) defines around 400 elements, ranging from layer 2, 3, and 4 packet fields to layer 7 application details, and including timestamps, pre/post NAT fields, sampling and filtering details. However, the registry includes very few performance metrics.

The IPPM WG defined an IPPM Metrics Registry [RFC4148]. However this was obsoleted by [RFC6248].

LMAP requires a standardised performance metrics registry, i.e. a PMOL IANA registry based on section 5.4.4 of [RFC6390]. See [I-D.claise-ippm-perf-metric-registry-00],

5.2.2. Mediation Configuration

In the Collector infrastructure, mediation changes traffic granularity, provides time and/or spatial data composition, data anonymization, and data retention.

[RFC5982] indicates that increasing numbers of data exporters, traffic, and the variety of treatments expected to be performed on the data make it more and more difficult to implement all measurement applications within a single Collector. To increase the collecting bandwidth capacity and processing capacity, distributed Collectors need to be deployed close to Exporters. In this case, those Collectors become mediators, re-exporting data on demand to centralized applications.

Although the IPFIX WG has published a Mediation Problem Statement [RFC5982] and a Mediation Framework [RFC6183], and is currently working on a mediation protocol [I-D.ietf-ipfix-mediation-protocol], there's currently no configuration model for mediation.

6. LMAP: Standards Re-usability

6.1. Existing Building Blocks

The LAMP charter has been defined [lmap-wg-charter]. The Working Group is in the process of defining the framework.

6.2. Missing Building Blocks

6.2.1. Task Definitions

The central part of LMAP is the Measurement Task itself which performs the measurement and generates the Measurement Result that is shared with the Collector.

An information model is needed to organize the Measurement Task configuration, scheduling, and result posting of measurement tasks. A proposal of such an information model can be found at [I-D.burbridge-lmap-information-model].

The Measurement Task is an instance of the Measurement Method at specific time (schedule) and place (Measurement Agent). The Measurement Method is the methodology used to generate the metrics. Therefore for comparable metrics the Measurement Method needs to be well understood and agreed upon. Additionally, the manner to reference the Measurement Method in the Instruction setup should be from a well-known registry. From experience, there are a number of existing methods to generate similarly named metrics. However, the results of these methods is not comparable as the algorithm used is not the same. The well-known registry should not simply list the measurement methods but also clearly define scope and usability of such metrics to avoid result comparison confusion.

A Measurement Task would include not only the Measurement Method but also configuration parameters such as (in the case of passive monitoring) what traffic to monitor or (in the case of active monitoring) that Remote Measurement Test Target(s) and test parameters etc. Some of these configuration parameters may not be explicit but implicit based on local state on the Measurement Agent. For example, the Controller may give Instruction to provide reachability (e.g. ping) information from the 1st and 2nd hop device towards a destination IP address. In this case, each 1st and 2nd hop device would not be known to the Controller and would be different at each Measurement Agent. Another example is the selection of the specific Controllers to which the Measurement Results should be posted to. The Controller may use ALTO [I-D.ietf-alto-protocol] to discover which Collector is the best one to use for each specific Measurement Agent, or the Controller may delegate the Controller selection to the Measurement Agent (ALTO, DNS SRV, etc.).

A Measurement Method could include a multi-part set of tests which chain information together to replicate a user workflow. For example the method might start with a DNS query to a specific website, a measurement on the DNS response time, and the DNS query result used in a HTTP GET (while using the VHOST of the website) and the download bitrate measured.

The Measurement Task could be spread across multiple Measurement Agents each generating and submitting their Measurement Results to the Collector(s). A Measurement Task ID would need to be allocated by the Controller to identify the Task to the Collector which would further aggregating the results from Measurement Agents. This Task ID would need to be unique across Controller reboots to prevent collision of different Measurement Tasks on to the Collector.

6.2.2. Instructions Setup

The Controller uses the Control Protocol to communicate with the Measurement Agents, to schedule tests. (As a scalability optimisation, the Controller may also use the Control Protocol to inform the Collector of the requested test(s). Else, every Measurement Agent would have to repeat the Test details to the Collector, along with the Test results.)

Which protocol should be used as the Control Protocol? Several possibilities exist, including NetConf [RFC6241], and YANG [RFC6020], Apache thrift, REST-style HTTP(s), TR-069, ALTO, ...

The Control Protocol should be transport independent, and available over a variety of transports. e.g., SCTP, TCP, and UDP, in both IPv4 and IPv6 networks, since Measurement Agents will be located in different kinds of networks. e.g., Home router versus branch office.

6.2.3. Task Scheduling

In one use case, tests are run immediately upon receipt of a command and reported immediately to the Collector. In a different use case, tests are configured ahead of time, perhaps across multiple Measurement Agents with the intention that all the Agents run the test at about the same time. In yet another case, a test may be run repeatedly or may otherwise make observations at several discrete times.

Therefore the Control Protocol must be able to clearly indicate to the Measurement Agent(s) when the test is scheduled, and the Reporting Protocol must be able to clearly indicate when the test was run.

These time indications may be either absolute ("at 10:23") or relative ("in 300 seconds"). Absolute timestamps require good clock synchronisation between the Controller, Measurement Agents, and Collector. Relative timestamps don't require any clock synchronisation. However, they're susceptible to delays.

The IPFIX WG has standardised many timestamps [iana-ipfix-assignments]). Each time stamp is available in multiple resolutions: seconds, milliseconds, microseconds, nanoseconds, being a trade-off between range and resolution.

6.2.4. Combining Active and Passive Measurements

The balanced use of both active and passive measurements would be needed in a large scale measurement system. While it is certainly possible to run active measurements to variety of test targets this can be disruptive to user traffic (and to the test if the active

measurement backs off) but also the remote measurement test targets that have user facing services. Additionally, active measurement would be taking away bandwidth certainly from the broadband site but potentially also from the ISP if the remote measurement test target is outside of the ISP.

Many questions can be answered by simple observation rather than explicit active measurement. For example, response times for DNS queries can be gleaned by observation of user traffic rather than explicit probing. In fact, it is possible to gather more samples of measurement that would have been acceptable under active measurement. Similarly, observation of user traffic of a Video on Demand stream to well known content provider can reveal information about the network conditions along the path to the content provider's server.

One proposal for making use of both active and passive measurement is to allow the Measurement Agent to make local decisions on which technique to use to deliver a particular metric-- as long as the specific method is included in the report. For example, DNS response time could be answered by passive monitoring as well as active monitoring. The Measurement Task could provide guidelines along how long to delay an active measurement in case passive measurement is unable to provide the result. If passive measurement is unable to provide a result, active measurement would be engaged.

Similarly, rather than completely backing off on an last mile path capacity active measurement in the presence of user traffic the Measurement Agent might keep a historical record of the high watermark of user traffic utilization and attempt to actively probe the delta current utilization and the high-water mark or the configured service profile (that the broadband site is 20mbps connected).

In all cases of the combined usage of active and passive measurement the results need to clearly indicate which method was used to what extent.

7. Security considerations

The privacy aspects of the end user measurements are important. The potentially large number of Measurement Agents capable of driving network traffic can be an attractive target for taking control of utilized for Denial of Service (DoS) attacks. The sizable resources associated also with the anchor Measurement Agents needs to be protected from unauthorized usage. Finally, as the Measurement Results could have potentially damaging commercial and regulatory effects they need to be protected as well.

The security considerations related to LMAP will be completed in the future.

8. IANA Considerations

There are no IANA considerations in this memo.

9. Acknowledgements

Thanks to all the authors of all the referenced works, and to the experts at Cisco who helped to make this draft possible.

Thanks to our families for their patience and understanding while we wrote this draft.

10. References

10.1. Normative References

- [I-D.burbridge-lmap-information-model]
Burbridge, T., Eardley, P., Bagnulo, M., and J. Schoenwaelder, "Information Model for Large-Scale Measurement Platforms (LMAP)", draft-burbridge-lmap-information-model-00 (work in progress), July 2013.
- [I-D.eardley-lmap-terminology]
Eardley, P., Morton, A., Bagnulo, M., and T. Burbridge, "Terminology for Large Measurement Platforms (LMAP)", draft-eardley-lmap-terminology-02 (work in progress), July 2013.
- [I-D.linsner-lmap-use-cases]
Linsner, M., Eardley, P., and T. Burbridge, "Large-Scale Broadband Measurement Use Cases", draft-linsner-lmap-use-cases-03 (work in progress), July 2013.
- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

10.2. Informative References

- [I-D.bagnulo-ippm-new-registry-independent]
Bagnulo, M., Burbridge, T., Crawford, S., Eardley, P., and A. Morton, "A registry for commonly used metrics. Independent registries", draft-bagnulo-ippm-new-registry-independent-01 (work in progress), July 2013.
- [I-D.bagnulo-ippm-new-registry]

Bagnulo, M., Burbridge, T., Crawford, S., Eardley, P., and A. Morton, "A registry for commonly used metrics", draft-bagnulo-ippm-new-registry-01 (work in progress), July 2013.

[I-D.ietf-alto-protocol]

Alimi, R., Penno, R., and Y. Yang, "ALTO Protocol", draft-ietf-alto-protocol-17 (work in progress), July 2013.

[I-D.ietf-homenet-arch]

Chown, T., Arkko, J., Brandt, A., Troan, O., and J. Weil, "Home Networking Architecture for IPv6", draft-ietf-homenet-arch-09 (work in progress), July 2013.

[I-D.ietf-ipfix-a9n]

Trammell, B., Wagner, A., and B. Claise, "Flow Aggregation for the IP Flow Information Export (IPFIX) Protocol", draft-ietf-ipfix-a9n-08 (work in progress), November 2012.

[I-D.ietf-ipfix-flow-selection-tech]

D'Antonio, S., Zseby, T., Henke, C., and L. Peluso, "Flow Selection Techniques", draft-ietf-ipfix-flow-selection-tech-18 (work in progress), May 2013.

[I-D.ietf-ipfix-ie-doctors]

Trammell, B. and B. Claise, "Guidelines for Authors and Reviewers of IPFIX Information Elements", draft-ietf-ipfix-ie-doctors-07 (work in progress), October 2012.

[I-D.ietf-ipfix-information-model-rfc5102bis]

Claise, B. and B. Trammell, "Information Model for IP Flow Information eXport (IPFIX)", draft-ietf-ipfix-information-model-rfc5102bis-10 (work in progress), February 2013.

[I-D.ietf-ipfix-mediation-protocol]

Claise, B., Kobayashi, A., and B. Trammell, "Operation of the IP Flow Information Export (IPFIX) Protocol on IPFIX Mediators", draft-ietf-ipfix-mediation-protocol-05 (work in progress), June 2013.

[I-D.ietf-ipfix-protocol-rfc5101bis]

Claise, B. and B. Trammell, "Specification of the IP Flow Information eXport (IPFIX) Protocol for the Exchange of Flow Information", draft-ietf-ipfix-protocol-rfc5101bis-10 (work in progress), July 2013.

[I-D.ietf-netconf-reverse-ssh]

- Watson, K., "Reverse Secure Shell (Reverse SSH)", draft-ietf-netconf-reverse-ssh-01 (work in progress), June 2013.
- [RFC3444] Pras, A. and J. Schoenwaelder, "On the Difference between Information Models and Data Models", RFC 3444, January 2003.
- [RFC4148] Stephan, E., "IP Performance Metrics (IPPM) Metrics Registry", BCP 108, RFC 4148, August 2005.
- [RFC4656] Shalunov, S., Teitelbaum, B., Karp, A., Boote, J., and M. Zekauskas, "A One-way Active Measurement Protocol (OWAMP)", RFC 4656, September 2006.
- [RFC5357] Hedayat, K., Krzanowski, R., Morton, A., Yum, K., and J. Babiarz, "A Two-Way Active Measurement Protocol (TWAMP)", RFC 5357, October 2008.
- [RFC5470] Sadasivan, G., Brownlee, N., Claise, B., and J. Quittek, "Architecture for IP Flow Information Export", RFC 5470, March 2009.
- [RFC5470] Sadasivan, G., Brownlee, N., Claise, B., and J. Quittek, "Architecture for IP Flow Information Export", RFC 5470, March 2009.
- [RFC5474] Duffield, N., Chiou, D., Claise, B., Greenberg, A., Grossglauser, M., and J. Rexford, "A Framework for Packet Selection and Reporting", RFC 5474, March 2009.
- [RFC5474] Duffield, N., Chiou, D., Claise, B., Greenberg, A., Grossglauser, M., and J. Rexford, "A Framework for Packet Selection and Reporting", RFC 5474, March 2009.
- [RFC5476] Claise, B., Johnson, A., and J. Quittek, "Packet Sampling (PSAMP) Protocol Specifications", RFC 5476, March 2009.
- [RFC5477] Dietz, T., Claise, B., Aitken, P., Dressler, F., and G. Carle, "Information Model for Packet Sampling Exports", RFC 5477, March 2009.
- [RFC5533] Nordmark, E. and M. Bagnulo, "Shim6: Level 3 Multihoming Shim Protocol for IPv6", RFC 5533, June 2009.
- [RFC5982] Kobayashi, A. and B. Claise, "IP Flow Information Export (IPFIX) Mediation: Problem Statement", RFC 5982, August 2010.

- [RFC6020] Bjorklund, M., "YANG - A Data Modeling Language for the Network Configuration Protocol (NETCONF)", RFC 6020, October 2010.
- [RFC6183] Kobayashi, A., Claise, B., Muenz, G., and K. Ishibashi, "IP Flow Information Export (IPFIX) Mediation: Framework", RFC 6183, April 2011.
- [RFC6235] Boschi, E. and B. Trammell, "IP Flow Anonymization Support", RFC 6235, May 2011.
- [RFC6241] Enns, R., Bjorklund, M., Schoenwaelder, J., and A. Bierman, "Network Configuration Protocol (NETCONF)", RFC 6241, June 2011.
- [RFC6248] Morton, A., "RFC 4148 and the IP Performance Metrics (IPPM) Registry of Metrics Are Obsolete", RFC 6248, April 2011.
- [RFC6390] Clark, A. and B. Claise, "Guidelines for Considering New Performance Metric Development", BCP 170, RFC 6390, October 2011.
- [RFC6728] Muenz, G., Claise, B., and P. Aitken, "Configuration Data Model for the IP Flow Information Export (IPFIX) and Packet Sampling (PSAMP) Protocols", RFC 6728, October 2012.
- [RFC6759] Claise, B., Aitken, P., and N. Ben-Dvora, "Cisco Systems Export of Application Information in IP Flow Information Export (IPFIX)", RFC 6759, November 2012.
- [RFC6812] Chiba, M., Clemm, A., Medley, S., Salowey, J., Thombare, S., and E. Yedavalli, "Cisco Service-Level Assurance Protocol", RFC 6812, January 2013.
- [RFC6887] Wing, D., Cheshire, S., Boucadair, M., Penno, R., and P. Selkirk, "Port Control Protocol (PCP)", RFC 6887, April 2013.
- [iana-ipfix-assignments]
Internet Assigned Numbers Authority, ., "IP Flow Information Export Information Elements (<http://www.iana.org/assignments/ipfix/ipfix.xml>)", .
- [lmap-wg-charter]
., "LMAP Working Group Charter (<http://tools.ietf.org/wg/lmap/charters>)", .

[pm-dir] , "Performance Metrics Directorate (<http://www.ietf.org/iesg/directorate/performance-metrics.html>)", .

Authors' Addresses

Aamer Akhter
Cisco Systems, Inc.
7025 Kit Creek Road
RTP, NC 27709
USA

Email: aakhter@cisco.com

Paul Aitken
Cisco Systems, Inc.
96 Commercial Street
Edinburgh, Scotland EH6 6LX
UK

Email: paitken@cisco.com

Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: March 14, 2015

M. Bagnulo
UC3M
T. Burbridge
BT
S. Crawford
SamKnows
J. Schoenwaelder
V. Bajpai
Jacobs University Bremen
September 10, 2014

Large MeAsurement Platform Protocol
draft-bagnulo-lmap-http-03

Abstract

This documents specifies the LMAP protocol based on HTTP for the Control and Report in Large Scale Measurement Platforms.

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <http://datatracker.ietf.org/drafts/current/>.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on March 14, 2015.

Copyright Notice

Copyright (c) 2014 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must

include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1. Introduction	2
2. Overview	4
3. Naming Considerations	4
4. Information model	6
5. Transport protocol	7
5.1. Pre-configured information	7
5.2. Control Protocol	7
5.2.1. Retrieving Instructions	8
5.2.2. Handling communication failures	10
5.2.3. Pushing Information from the Controller to the MA	10
5.3. Report protocol	11
5.3.1. Handling communication failures	12
6. LMAP Data Model	13
6.1. Timing Information	13
6.2. Channels	15
6.3. Configuration	15
6.4. Instruction	16
6.5. Measurement Supression	16
6.6. Measurement Task Configurations	16
6.7. Measurement Schedules	17
6.8. Logging	18
6.9. Capability and Status	18
6.10. Reporting	19
7. Security considerations	22
8. IANA Considerations	24
9. Acknowledgments	24
10. References	24
10.1. Normative References	24
10.2. Informative References	25
Authors' Addresses	25

1. Introduction

A Large MeAsurement Platform (LMAP) is an infrastructure deployed in the Internet that enables performing measurements from a very large number of vantage points.

The main components of a LMAP are the following:

- o The Measurement Agents (MAs): these are the processes that perform the measurements. The measurements can be both active or passive measurements.

- o The Controller: this is the element that controls the MAs. In particular it provides configuration information and it instructs the MA to perform a set of measurements.
- o The Collector: this is the repository where the MAs send the results of the measurements that they have performed.

These and other terms used in this document are defined in [I-D.ietf-lmap-framework]. We only include the definition of the main elements in this document so it is self-contained and can be read without the need to consult other documents. The reader is referred to the terminology draft for further details.

In order for a LMAP to work, the following protocols are required:

- o Measurement protocols: These are the protocols used between the MA and the Measurement Peer in active measurements. These are the actual packets being used for the measurement operations.
- o Control Protocol. This is the protocol between the Controller and the MAs. This protocol is used to convey measurement Instruction(s) from the Controller to the MA as well as logging, failure and capabilities information from the MA to the Controller.
- o Report Protocol. This is the protocol between the MAs and the Collector. This protocol conveys information about the results of the measurements performed by the MA to the Collector.

Both the Control protocol and the Report protocol have essentially two parts: a transport and a data model. The data model represents the information about measurement instructions and logging/failure/capabilities (in the Control protocol) and the information about measurement results (in the Report protocol) that is being exchanged between the parties. The transport is the underlying protocol used to exchange that information. This document specifies the use of HTTP 1.1 [RFC7230] [RFC7231] [RFC7232] [RFC7233] [RFC7234] [RFC7235] as a transport for the Control and the Report protocol. This document also defines the data model for the Control and Report protocols. The data model described in this document follows the information model described in [I-D.ietf-lmap-information-model]. The Measurement protocols are out of the scope for this document.

At this stage, the goal of this document is to explore different options that can be envisioned to use the HTTP protocol to exchange LMAP information and to foster discussion about which one to use (if any). Because of that, the document contains several discussion paragraphs that explore different alternative approaches to perform the same function.

2. Overview

This section provides an overview of the architecture envisioned for a LMAP using HTTP as transport protocol. As we described in the previous section, a LMAP is formed by a large number of MAs, one or more Controllers and one or more Collectors. We assume that before the MAs are deployed, it is possible to pre-configure some information in them. Typically this includes information about the MA itself (like its identifier), security information (like some certificates) and information about the Controller(s) available in the measurement platform. Once that the MA is deployed it will retrieve additional configuration information from the pre configured Controller. After obtaining the configuration information, the MA is ready to receive Instructions from the Controller and initiate measurement tasks. The MA will perform the following operations:

- o It will obtain Instructions from one of the configured Controllers. These Instructions include information about the set of measurement tasks to be performed, a schedule for the execution of the measurements as well as a set of report channels. This information is downloaded by the MA from the Controller. The MA will periodically check whether there are new Instructions available from the Controller. This document specifies how the MA uses the HTTP protocol to retrieve information from the Controller.
- o The MA will execute measurement tasks either by passively listening to traffic or by actively sending and receiving measurement packets. How this is done is out of the scope of this document.
- o After one or more measurements have been performed, the MA reports the results to the Collector. The timing of these uploads is specified in the measurement Instruction i.e. each measurement specified in a measurement Instruction contains a report information, defining when the MA should report the results back to the Collector. This document specifies how the MA uses the HTTP protocol to upload the measurement results to the Collector.
- o In addition, the MA will periodically report back to the Controller information about its capabilities (like the number of interfaces it has, the corresponding IP addresses, the set of measurement methods it supports, etc) and also logging information (whether some of the requested measurement tasks failed and related information).

3. Naming Considerations

In this section we define how the different elements of the LMAP architecture are identified and named.

The Controller and the Collectors can be assumed to have both an IP address and a Fully Qualified Domain Name (FQDN). It is natural to use these as identifiers for these elements. In this document we will use FQDNs, but IP addresses can be used as well.

The MAs on the other hand, are likely to be executed in devices located in the end user premises and are likely to be located behind a NAT box. It is reasonable to assume they have neither a public IP address nor a FQDN. We propose then that the MAs are identified using an Universally Unique Identifier URN as defined in RFC 4122 [RFC4122]. In particular each MA has a version 4 UUID, which is randomly or pseudo randomly generated.

DISCUSSION:

MA ID Configuration: Some open issues related to this are: a) whether the MA ID is configured before or after the MA is deployed, b) if configured after deployment whether the MA ID is generated locally and posted or fetched from the Controller and c) whether this is within the scope of this (or other) specification if any. These issues seem also to be related to the nature of the MA platform (whether the MA is a software downloaded into a general purpose device or it is a special purpose hardware box). Consider the case that the MA is located in a special purpose hardware box, then having the MA ID pre configured before deployment requires a per device customization that is expensive. It would be more costly efficient to reuse an existent (hopefully) unique identifier available in the hardware (such as a MAC address) to serve as a one-time pre configured identifier to be used to fetch (or post a self generated) the MA ID from the Controller once the MA is deployed. The requirement for such one-time identifier is that they must be unique (which is not always true for the MACs). About the local generation of the MA ID (as opposed to fetch it from the Controller), the generation process performed in the MA MUST be idempotent, i.e. if the MA was factory-reset then the server would still see it with the same MA ID when it came back up. This is probably easier to achieve if it is generated in the Controller and then fetched by the MA. Finally, it is not clear at this stage if this needs to be specified in this document or in the information model document or left open to the implementers. Group identifiers. In some cases, like the case of measurements in mobile devices, it may be important because of privacy considerations for the MA not to have a unique identifier. It is possible then to assign "Group identifiers" to a set of devices that share relevant characteristics from the measurement perspective (e.g. devices from the same operator, with the same type of contract or other relevant feature). In this case, the MAs within the same group would retrieve common measurement

Instructions from the controller by presenting the same Group ID and would report results including the Group ID in the report. This would imply that it would not be possible for the platform to correlate specific measurement data with any given MA. The downside of this is that some MAs may be over-represented while other under-represented in the measurement data and it would not be possible to detect this case (for instance a given MA may have reported 20 results while another one only one). In order to deal with this issue, the MA behaviour must be programmed accordingly (e.g. the MA should not perform more than one measurement every given period of time). In addition, it should be noted that privacy is only achieved in a holistic way. This means that really anonymity of the MA is incompatible with strong authentication. In particular, if a measurement platform's goal is to keep MAs anonymous, it cannot require any form of strong authentication (other than weak group authentication e.g. a password shared by a group), which has security implications. In particular, the threat for report forgery (i.e. enabling an attacker to submit forged reports as discussed in the security considerations) increases.

There are additional naming considerations related to:

- o The measurements. In order to enable a Controller to properly convey a measurement schedule, it must be possible for the Controller to specify a measurement to be performed while providing the needed input parameters. While this is critical, it is out of the scope of this document. There is a proposed registry for metrics/measurements in [I-D.bagnulo-ippm-new-registry-independent])
- o The resources being exchanged, namely, the configuration information, the measurement Instructions and the reports. These are being discussed in the upcoming sections.

4. Information model

The information model for LMAP is described [I-D.ietf-lmap-information-model]. It contains basically two models one for the control information (i.e. the Instructions from the Controller to the MA) and a model for the Report information. We briefly describe their overall structure here.

The control information (or Instruction) has the following five elements:

- o The Set of Measurement Task Configurations: This element defines the measurements/test that the MA will perform without defining the schedule when they will be performed.

- o The Set of Report Channels: This element defines the set of collectors as well as the reporting schedules for the reports.
- o The Set of Measurement Schedules for Repeated Tasks: defines the schedules for the repeated measurements, by referencing the measurement tasks defined in the second element.
- o Suppression information

Summary of Report information model here.

Summary of Capability and Status information model here.

Summary of Logging information model here.

5. Transport protocol

5.1. Pre-configured information

As we mentioned earlier, the MAs contain pre-configured information before being deployed. The pre-configured information is the following:

- o The UUID for the MA. This should be pre-configured so that the Controller is aware of the MA and can feed configuration information and measurement Instructions to it.
- o Information about one or more Controllers. The MA MUST have enough information to create the URL for the Instruction resources. This includes the the FQDN of each of the Controller or the IP addresses of the Controller, as well as the well-known path prefix and its identifier.
- o The certificate for the Certification authority that is used in the platform to generate the certificates for the Controller and the Collector. See the Security considerations section below.
- o The security related information for the MA (it can be a certificate for the MA and the corresponding private key, or simply a key/password depending on the security method used, see the security considerations section below).

5.2. Control Protocol

The Control protocol is used by the MA to retrieve Instruction information from the Controller. In this section we describe how to use HTTP to transport Instructions. The Instruction information is structured as defined in the LMAP Information model [I-D.ietf-lmap-information-model] as described in the previous section. The MA uses the Control protocol to retrieve all the resources described above, namely, the Agent information, the Set of Measurement Task Configurations, the Set of Report Channels, the Set of Measurement Schedules for Repeated Tasks and the Set of

Measurement Schedules for Isolated Tasks. The main difference from the HTTP perspective is that the MA MUST have the URL for the Agent Information resource pre-configured as described in the previous section, while the URLs for all the other resources are contained in the Agent Information resource itself.

5.2.1. Retrieving Instructions

In order to retrieve the Instruction resources from the Controller the MA can use either the GET or the POST method using the corresponding URL.

5.2.1.1. Using the GET method

One way of using the GET method to retrieve configuration information is to explicitly name the configuration information resources and then apply the GET method. The MA retrieves its Instruction when it is first connected to the network and periodically after that. The frequency for the periodical retrieval is contained in the Agent Information (???).

The URL for the Agent Information resource is formed as the FQDN of the Controller, a well-known path prefix and the MA UUID. The well-known path prefix is /.well-known/lmap/ma-info. The URL for the remaining resources that compose the Instruction are contained in the Agent Information.

Agent Information retrieval: In order to retrieve the Agent information the MA uses the HTTP GET method follows:

```
GET /.well-known/lmap/ma-info/ < ma-iid> HTTP/1.1
Host: FQDN or IP of the Controller
Accept: application/json (as per [RFC7159])
```

The Agent Information should contain the Configuration Retrieval Schedule (i.e. how often the MA should retrieve configuration information) and also the Measurement Instruction Retrieval Schedule (i.e. how often the MA should retrieve the Measurement Instruction from the Controller). COMMENT: this is missing from the Data Model

The retrieval of the remaining resources of the Instruction using the GET method is analogous, only that the URL is extracted from the Agent Information file rather than constructed with pre-configured information.

The format for the response should be described here

Periodical Instruction retrieval: After having downloaded the initial Instruction information, the MA will periodically look for updated Instruction information. The frequency with which the MA polls for the new Instructions from the Controller is contained in the last Agent Information downloaded. In order to retrieve the Agent Information, the MA uses the GET method as follows:

```
GET /.well-known/lmap/ma-info/ma-iid/ HTTP/1.1
Host: FQDN or IP of the Controller
Accept: application/json (as per [RFC7159])
If-None-Match: the eTag of the last retrieved Agent Information
(an alternative option here is to use If-Modified-Since, not sure
which one is best)
```

For the other Instruction resources, the GET method is applied in the same way just that the URL used are the ones retrieved in the last Agent Information.

The format for the response should be described here

Alternatively, instead of explicitly naming the Instruction resources for each MA, it is possible to perform a query using the GET method as well. In this case, the MA could perform a GET for the following URI `http://controller.example.org/?ma=maid & q=ma-info` (similar queries can be constructed for the other Instruction resources). (I am not sure how to express in this case the condition that the MA wishes to retrieve the configuration if it is newer than the last one it downloaded.)

5.2.1.2. Using the POST method

An alternative to retrieve Instruction resources is to use the POST method to perform a query (similar to the query using GET). In this case there is no explicit naming of the Instruction information of each MA, but a general Instruction resource and the POST method is used to convey a query for the Instruction information of a particular MA. For the case of the Agent Information resource, this would look like as follows:

```
POST /.well-known/lmap/ma-info/ma-iid/ HTTP/1.1
Host: controller.example.com
Content-Type: application/lmap-maid+json
Accept: application/lmap-config+json
{
  "ma-id" : "550e8400-e29b-11d4-a716-446655440000",
}
```

The reply for this query would contain the actual configuration information as follows:

```
HTTP/1.1 200 OK
Content-Length: xxx
Content-Type: application/lmap-config+json
{
// whatever config goes here
}
```

In this case, the URLs contained in the Agent information can be generic and not MA specific, since the MA will use the POST method including its own identifier when retrieving the Instruction resources.

The argument for this approach is that this is much more extensible since the POST can carry complex information and there is no need to "press" arguments into the strict hierarchy of URIs.

We need to describe how to use this to retrieve newer information in the periodic case.

5.2.2. Handling communication failures

The cases that the MA is unable to retrieve the Instructions are handled as follows:

- o The MA will use a timeout for the communication of TIMEOUT seconds. The value of TIMEOUT MUST be configurable via the aforementioned Configuration Information retrieval protocol. The default value for the TIMEOUT is 3 seconds. If after the timeout, the communication with the Controller has not been established, the MA will retry doing an exponential backoff and doing a round robin between the different Controllers it has available.
- o If a HTTP error message (5xx) is received from the Controller as a response to the GET request, the MA will retry doing an exponential back-off and doing a round robin between the different Controllers it has available. The 5xx error codes indicate that this Controller is currently incapable of performing the requested operation.

5.2.3. Pushing Information from the Controller to the MA

The previous sections described how the MA periodically polls the Controller to retrieve Instruction information. The frequency of the downloads is configurable. The question is whether this is enough or a mechanism for pushing Instruction information is needed. Such method would enable to contact the MA in any moment and take actions

like triggering a measurement right away or for instance to stop an ongoing measurement (e.g. because it is disturbing the network). The need for such a mechanism is likely to depend on the use case of the platform. Probably the ISP use case is more likely to require this feature than the regulator/benchmarking use case. It is probably useful then to provide this as an optional feature.

The main challenge in order to provide this feature is that the MAs are likely to be placed behind NATs, so it is not possible for the Controller to initiate a communication with the MA unless there is a binding in the NAT to forward the packets to the MA. There are several options that can be considered to enable this communication:

- o The MA can use one of the NAT control protocols, such as PCP or UPNP. If this approach is used, the MA will create a binding in the NAT opening a hole. After that, the MA should inform the Controller about which is the IP address and port available for communication. It would be possible to re-use existing protocols to forward this information. The problem with this is that the NAT may not support these protocols or they may not be activated. In any case, a solution should try to use them in the case they are available.
- o If it is not possible to use a NAT control protocol, then the MA can open a hole in the NAT by establishing a connection to the Controller and keeping it open. This allows the Controller to push information to the MA through that connection. One concern with this approach is that the MA is playing the role of the client and the Controller is playing the role of the server (the MA is initiating the TCP connection), but it would be the Controller who would use the PUT method towards the MA reversing the roles. An alternative approach is that the MA has a long running GET pending which is answered by the server if the measurement Instruction changes (or the server times out, in which case the MA restarts the long running GET. More discussion is needed about whether one of these options is acceptable or not. In addition, this would imply that the Controller should maintain as many open sessions as MAs it is managing, which imposes additional burden in the Controller. There are security considerations as well, but these are covered in the Security Considerations section below.

5.3. Report protocol

The MA after performing the measurements reports the results to a collector. There can be more than one collector within a LMAP framework. Each collector is identified by its FQDN or IP address which is retrieved as part of the Agent information from a pre-configured controller as previously discussed. The number of

Collectors that the MA uploads the results to as well as the schedule when it does so is defined in the measurement Instruction previously downloaded from the Controller. The MA themselves are identified by a UUID.

There are two options that can be considered for the MA to upload reports to the Collector either to use the PUT method or to use the POST method.

If the PUT method option is used, then the MA need to perform the PUT method using an explicit name for the report resource it is transferring to the Collector. The name of the resource is contained in the Agent Information previously retrieved by the MA

The other option is for the MA to use the POST method to upload the measurement reports to one or more Collectors. In this case,, the POST message body can contain the identifier of the MA and additional information describing the report in addition to the report itself.

One argument to consider is that PUT is idempotent. This means that if the network is bad at some point and the MA is not sure whether its request made it through, it can send it a second (or nth) time, and it is guaranteed that the request will have exactly the same effect as sending it for the first time. POST does not by itself guarantee this. This can be achieved by verifying the report data itself, and contrast it with data already stored in the Collector database.

5.3.1. Handling communication failures

The MA will use a timeout for the communication with the Collector of TIMEOUT seconds. The value of TIMEOUT MUST be configurable via the aforementioned Configuration Information retrieval protocol. The default value for the TIMEOUT is 3 seconds.

If the MA is uploading the report to several Collectors and it manages to establish the communication before TIMEOUT seconds with at least one of them, but not with one or more of the other Collectors, then the MA gives up after TIMEOUT seconds and it MAY issue an alarm. The definition of how to do that operation is out of the scope of this document.

If the MA is uploading the report to only one Collector, and it does not manages to establish a communication before TIMEOUT seconds, then it retry doing an exponential backoff and doing a round robin between the different Collectors it has available.

Similarly, if an HTTP error message (5xx) is received from the Collector as a response to the PUT request, the MA will retry doing an exponential backoff and doing a round robin between the different Collectors it has available. The 5xx error codes indicate that this Collector is currently incapable of performing the requested operation.

In order to support this, the information model must express the difference between a report sent to multiple collectors and multiple collectors used for fallback.

6. LMAP Data Model

This section will contain the data model in json.

6.1. Timing Information

An example immediate timing object with no defined randomness is shown below:

```
{
  "timings": [
    {
      "id": 1,
      "ma_periodic_end": 1410017611,
      "ma_periodic_interval": 3600000,
      "ma_periodic_start": 1410017613,
      "ma_randomness_spred": 0,
      "ma_timing_name": "hourly"
    },
    {
      "id": 3,
      "ma_periodic_end": 1410017611,
      "ma_periodic_interval": 86400,
      "ma_periodic_start": 1410017613,
      "ma_randomness_spred": 0,
      "ma_timing_name": "daily"
    },
    {
      "id": 2,
      "ma_periodic_end": 1410017611,
      "ma_periodic_interval": 3600000,
      "ma_periodic_start": 1410017613,
      "ma_randomness_spred": 0,
      "ma_timing_name": "hourly"
    }
  ]
}
```

```
"id": 4,
"ma_calendar_days_of_month": "",
"ma_calendar_days_of_week": "tuesday, thursday, sunday",
"ma_calendar_end": 1410017613,
"ma_calendar_hours": "18",
"ma_calendar_minutes": "04",
"ma_calendar_months": "",
"ma_calendar_seconds": "42",
"ma_calendar_start": 1410017612,
"ma_calendar_timezone_offset": 2,
"ma_randomness_spred": 0,
"ma_timing_name": "tuesday-thursday-sunday"
},
{
  "id": 5,
  "ma_calendar_days_of_month": "",
  "ma_calendar_days_of_week": "",
  "ma_calendar_end": 1410017619,
  "ma_calendar_hours": "0, 6 12 18",
  "ma_calendar_minutes": "0",
  "ma_calendar_months": "",
  "ma_calendar_seconds": "0",
  "ma_calendar_start": 1410017612,
  "ma_calendar_timezone_offset": 2,
  "ma_randomness_spred": 21600000,
  "ma_timing_name": "once-every-six-hours"
},
{
  "id": 6,
  "ma_one_off_time": 1410017613,
  "ma_randomness_spred": 0,
  "ma_timing_name": "immediate"
},
{
  "id": 7,
  "ma_one_off_time": 1410017613,
  "ma_randomness_spred": 0,
  "ma_timing_name": "immediate"
},
{
  "id": 8,
  "ma_randomness_spred": 12345,
  "ma_timing_name": "startup"
}
]
}
```

6.2. Channels

An example channel object using the aforementioned timing object is shown below:

```
{
  "channels": [
    {
      "id": 1,
      "ma_channel_credentials": "MIIFeZCCAvsCAQEwDQYJ...",
      "ma_channel_interface_name": "eth0",
      "ma_channel_name": "default-collector-channel",
      "ma_channel_target": "collector.example.org"
    },
    {
      "id": 2,
      "ma_channel_credentials": "MIIFeZCCAvsCAQEwDQYJ...",
      "ma_channel_interface_name": "eth0",
      "ma_channel_name": "default-controller-channel",
      "ma_channel_target": "controller.example.org"
    }
  ]
}
```

6.3. Configuration

An example config object using the aforementioned channel objects is shown below:

```
{
  "config": [
    {
      "id": 1,
      "ma_agent_id": "550e8400-e29b-41d4-a716-446655440000",
      "ma_channel_name": "default-controller-channel",
      "ma_control_channel_fail_tresh": "10",
      "ma_credentials": "MIIFeZCCAvsCAQEwDQYJ...",
      "ma_device_id": "01:23:45:67:89:ab",
      "ma_group_id": "550e8400-e29b-41d4-a716-446655440123",
      "ma_report_ma_id_flag": "1"
    }
  ]
}
```

6.4. Instruction

The instruction object is essentially a wrapper around suppression, schedule, task, channel objects.

6.5. Measurement Supression

An example supression object used by the aforementioned instruction object is shown below:

```
{
  "supression": [
    {
      "id": 1,
      "ma_supression_enabled": 0,
      "ma_supression_end": 0,
      "ma_supression_schedule_names": "icmp-latency-immediate",
      "ma_supression_start": 1410037509,
      "ma_supression_stop_ongoing_task": 0,
      "ma_supression_task_names": "iperf-server"
    }
  ]
}
```

6.6. Measurement Task Configurations

An example task object used by the aforementioned instruction object is shown below:

```
{
  "tasks": [
    {
      "id": 1,
      "ma_role": "",
      "ma_task_cycle_id": "1",
      "ma_task_name": "udp-latency-test",
      "ma_task_options": "",
      "ma_task_registry_entry": "urn:...",
      "ma_task_supress_default": "true"
    },
    {
      "id": 5,
      "ma_role": "",
      "ma_task_cycle_id": "1",
      "ma_task_name": "reporting-daily",
      "ma_task_options": "",
      "ma_task_registry_entry": "urn:...",
      "ma_task_supress_default": "true"
    }
  ]
}
```



```

    },
    {
      "id": 2,
      "ma_role": "",
      "ma_task_cycle_id": "1",
      "ma_task_name": "icmp-latency-test",
      "ma_task_options": "",
      "ma_task_registry_entry": "urn:...",
      "ma_task_suppress_default": "true"
    },
    {
      "id": 3,
      "ma_role": "",
      "ma_task_cycle_id": "1",
      "ma_task_name": "iperf-server",
      "ma_task_options": "{\\"name\\":\\"role\\",
        \\"value\\":\\"server\\"}",
      "ma_task_registry_entry": "server",
      "ma_task_suppress_default": "false"
    },
    {
      "id": 4,
      "ma_role": "",
      "ma_task_cycle_id": "1",
      "ma_task_name": "lmap-reporting-task",
      "ma_task_options": "",
      "ma_task_registry_entry": "lmap-reportd",
      "ma_task_suppress_default": "true"
    }
  ]
}

```

6.7. Measurement Schedules

An example schedule object used by the aforementioned instruction object is shown below:

```
{
  "schedules": [
    {
      "id": 1,
      "ma_sched_channel_interface_select": "0",
      "ma_sched_channel_names": "default-collector-channel",
      "ma_sched_task_downstream_config_names": "reporting-daily",
      "ma_sched_task_output_selection": "1",
      "ma_schedule_name": "reporting-immediate",
      "ma_schedule_task_name": "icmp-latency-test",
      "ma_timing_name": "immediate"
    }
  ]
}
```

6.8. Logging

An example log object is shown below:

```
{
  "logging": [
    {
      "id": 1,
      "ma_log_agent_id": "0e49b32b01falle4bcaf10ddb1bd23b5",
      "ma_log_code": "200",
      "ma_log_description": "OK",
      "ma_log_event_time": 1404313752
    }
  ]
}
```

6.9. Capability and Status

An example status object is shown below:

```
{
  "status": [
    {
      "id": 1,
      "ma_agent_id": "c54c284a01ee11e48dd310ddb1bd23b5",
      "ma_condition_code": "8081",
      "ma_condition_text": "Cond_Text",
      "ma_device_id": "urn:dev:mac:0024beffffe804ff1",
      "ma_firmware": "4560",
      "ma_hardware": "TL-MR3020",
      "ma_interface_dns_server": "8.8.8.8",
      "ma_interface_gateway": "192.168.1.1",
      "ma_interface_ip_address": "192.168.1.10",
      "ma_interface_name": "eth0",
      "ma_interface_speed": "100Mbps",
      "ma_interface_type": "100baseTX",
      "ma_last_config": "140423245",
      "ma_last_instruction": "140431312",
      "ma_last_measurement": "1404315031",
      "ma_last_report": "1404315053",
      "ma_link_layer_addr": "01:23:45:67:89:ab",
      "ma_task_name": "Report",
      "ma_task_registry": "urn:ietf:lmmap:report:http_report",
      "ma_task_role": "Role",
      "ma_version": "Busybox"
    }
  ]
}
```

6.10. Reporting

An example report object is shown below:

```
{
  "reporting": [
    {
      "id": 1,
      "ma_report_agent_id": "550e8400-e29b-41d4-a716-446655440000",
      "ma_report_date": 1404315031,
      "ma_report_group_id": "550e8400-e29b-41d4-a716-446655440123",
      "ma_report_result_conflict_task": "0",
      "ma_report_result_cross_traffic": 20,
      "ma_report_result_end_time": 1404315031,
      "ma_report_result_start_time": 1404315031,
      "ma_report_result_values": "result_values",
      "ma_report_task_column_labels": "\"start-time\"",
      "\"conflicting-tasks\"", "\"cross-traffic\"", "\"mean\"",
    }
  ]
}
```

```

        \ "min\", \ "max\"",
        "ma_role": "",
        "ma_task_cycle_id": "1",
        "ma_task_name": "udp-latency-test",
        "ma_task_options": "",
        "ma_task_registry_entry": "urn:...",
        "ma_task_supress_default": "true"
    },
    {
        "id": 2,
        "ma_report_agent_id": "550e8400-e29b-41d4-a716-446655440000",
        "ma_report_date": 1404315031,
        "ma_report_group_id": "550e8400-e29b-41d4-a716-446655440123",
        "ma_report_result_conflict_task": "0",
        "ma_report_result_cross_traffic": 20,
        "ma_report_result_end_time": 1404315031,
        "ma_report_result_start_time": 1404315031,
        "ma_report_result_values": "result_values",
        "ma_report_task_column_labels": "\"start-time\",
        \"conflicting-tasks\", \"cross-traffic\",
        \"mean\", \"min\", \"max\"",
        "ma_role": "",
        "ma_task_cycle_id": "1",
        "ma_task_name": "icmp-latency-test",
        "ma_task_options": "",
        "ma_task_registry_entry": "urn:...",
        "ma_task_supress_default": "true"
    },
    {
        "id": 3,
        "ma_report_agent_id": "550e8400-e29b-41d4-a716-446655440000",
        "ma_report_date": 1404315031,
        "ma_report_group_id": "550e8400-e29b-41d4-a716-446655440123",
        "ma_report_result_conflict_task": "0",
        "ma_report_result_cross_traffic": 20,
        "ma_report_result_end_time": 1404315031,
        "ma_report_result_start_time": 1404315031,
        "ma_report_result_values": "result_values",
        "ma_report_task_column_labels": "\"start-time\",
        \"conflicting-tasks\", \"cross-traffic\",
        \"mean\", \"min\", \"max\"",
        "ma_role": "",
        "ma_task_cycle_id": "1",
        "ma_task_name": "iperf-server",
        "ma_task_options": "{\\"name\\":\\"role\\",
        \\"value\\":\\"server\\"}",
        "ma_task_registry_entry": "server",
        "ma_task_supress_default": "false"
    }

```

```
    },
    {
      "id": 4,
      "ma_report_agent_id": "550e8400-e29b-41d4-a716-446655440000",
      "ma_report_date": 1404315031,
      "ma_report_group_id": "550e8400-e29b-41d4-a716-446655440123",
      "ma_report_result_conflict_task": "0",
      "ma_report_result_cross_traffic": 20,
      "ma_report_result_end_time": 1404315031,
      "ma_report_result_start_time": 1404315031,
      "ma_report_result_values": "result_values",
      "ma_report_task_column_labels": "\"start-time\",
      \"conflicting-tasks\", \"cross-traffic\",
      \"mean\", \"min\", \"max\"",
      "ma_role": "",
      "ma_task_cycle_id": "1",
      "ma_task_name": "lmap-reporting-task",
      "ma_task_options": "",
      "ma_task_registry_entry": "lmap-reportd",
      "ma_task_suppress_default": "true"
    },
    {
      "id": 5,
      "ma_report_agent_id": "550e8400-e29b-41d4-a716-446655440000",
      "ma_report_date": 1404315031,
      "ma_report_group_id": "550e8400-e29b-41d4-a716-446655440123",
      "ma_report_result_conflict_task": "0",
      "ma_report_result_cross_traffic": 20,
      "ma_report_result_end_time": 1404315031,
      "ma_report_result_start_time": 1404315031,
      "ma_report_result_values": "result_values",
      "ma_report_task_column_labels": "\"start-time\",
      \"conflicting-tasks\", \"cross-traffic\",
      \"mean\", \"min\", \"max\"",
      "ma_role": "",
      "ma_task_cycle_id": "1",
      "ma_task_name": "reporting-daily",
      "ma_task_options": "",
      "ma_task_registry_entry": "urn:...",
      "ma_task_suppress_default": "true"
    }
  ]
}
```

7. Security considerations

Large Measurement Platforms may result in a security hazard if they are not properly secured. This is so because they encompass a large number of MAs that can be managed and coordinated easily to generate traffic and they can potentially be used for generating DDoS attacks or other forms of security threats.

From the perspective of the protocols described in this documents, we can identify the following threats:

- o Hijacking: Probably the worst threat is that an attacker takes over the control of one or more MAs. In this case the attacker would be able to instruct the MAs to generate traffic or to eavesdrop traffic in their location. It is then critical that the MA is able to strongly authenticate the Controller. An alternative way to achieve this attack is to alter the communication between the Controller and the MAs. In order to prevent this form of attack, integrity protection of the communication between the Controller and the MAs is required.
- o Polluting: Another type of attack is that an attacker is able to pollute the Collectors database by providing false results. In this case, the attacker would attempt to impersonate one or more MAs and upload fake results in the Collector. In order to prevent this, the authentication of the MAs with the Collector is needed. An alternative way to achieve this is for an attacker to alter the communication between the MA and the Collector. In order to prevent this form of attack, integrity protection of the communication between the MA and the Collector is needed.
- o Disclosure: Another threat is that an attacker may gather information about the MAs and their configuration and the Measurement schedules. In order to do that, it would connect to the Controller and download the information about one or more MAs. This can be prevented by using MA authentication with the Controller. An alternative mean to achieve this would be for the attacker to eavesdrop the communication between the MA and the Controller. In order to prevent this, confidentiality in the communication between the MA and the Controller is required. Similarly, an attacker may wish to obtain measurement result information by eavesdropping the communication between the MA and the Collector. In order to prevent this, confidentiality in the communication between the MA and the Collector is needed.

In order to address all the identified threats, the HTTPS protocol must be used for LMAP (i.e. using HTTP over TLS). HTTPS provides confidentiality, integrity protection and authentication, satisfying all the aforementioned needs. Ideally, mutual authentication should be used. In any case, server side authentication MUST be used. In

order to achieve that, both the Controller and the Collector MUST have certificates. The certificate of the CA used to issue the certificates for the Controller and the Collector MUST be pre configured in the MAs, so they can properly authenticate them. As mentioned earlier, ideally, mutual authentication should be used. However, this implies that certificates for the MAs are needed. Certificate management for a large number of MAs may be expensive and cumbersome. Moreover, the major threats identified are the ones related to hijacking of the MAs, which are prevented by authenticating the Controller. MAs authentication is needed to prevent Polluting and Disclosure threats, which are less severe. So, in this case, alternative (cheaper) methods for authenticating MAs can be considered. The simplest method would be to simply use the MA UUID as a token to retrieve information. Since the MA UUID is 128 bit long, it is hard to guess. It would be also possible to use a password and use the HTTP method for authentication. It is not obvious that managing passwords for a large number of MAs is easier than managing certificates though.

An additional security consideration is posed by the mechanism to push information from the Controller to the MAs. If this method is used, it would be possible its abuse by an attacker to control the MAs. This threat is prevented by the use of HTTPS. If HTTPS is used in the established connection between the MA and the Controller, the only effect that a packet generated by an external attacker to the MA or the Controller would be to reset the HTTPS connection, requiring the connection to be re-established.

It is required in this document that both the Controller and that the Collector are authenticated using digital certificates. The current specification allows for the MA to have information about the certificate of the Certification authority used for generating the Controller and Collector certificates while the actual certificates are exchanged in band using TLS. Another (more secure) option is to perform certificate pinning i.e. to configure in the MAs the actual certificates rather than the certification authority certificate. Another measure to increase the security would be to limit the domains that the FQDNs of the Controller and/or the Collector (e.g. only names in the exmample.org domain).

Large scale measurements can have privacy implications, especially in some scenarios like mobile devices performing measurements. In this memo we have considered using Group IDs to the MA in order to avoid the possibility for the platform to track each individual MA that is feeding results.

8. IANA Considerations

Registration of the well-known URL

9. Acknowledgments

We would like to thank Vlad Victor Ungureanu (Jacobs University Bremen) for providing us external support.

Marcelo Bagnulo, Trevor Burbridge, Sam Crawford, Juergen Schoenwaelder and Vaibhav Bajpai work in part on the Leone research project, which receives funding from the European Union Seventh Framework Programme [FP7/2007-2013] under grant agreement number 317647.

10. References

10.1. Normative References

- [RFC4122] Leach, P., Mealling, M., and R. Salz, "A Universally Unique IDentifier (UUID) URN Namespace", RFC 4122, July 2005.
- [RFC7159] Bray, T., "The JavaScript Object Notation (JSON) Data Interchange Format", RFC 7159, March 2014.
- [RFC7230] Fielding, R. and J. Reschke, "Hypertext Transfer Protocol (HTTP/1.1): Message Syntax and Routing", RFC 7230, June 2014.
- [RFC7231] Fielding, R. and J. Reschke, "Hypertext Transfer Protocol (HTTP/1.1): Semantics and Content", RFC 7231, June 2014.
- [RFC7232] Fielding, R. and J. Reschke, "Hypertext Transfer Protocol (HTTP/1.1): Conditional Requests", RFC 7232, June 2014.
- [RFC7233] Fielding, R., Lafon, Y., and J. Reschke, "Hypertext Transfer Protocol (HTTP/1.1): Range Requests", RFC 7233, June 2014.
- [RFC7234] Fielding, R., Nottingham, M., and J. Reschke, "Hypertext Transfer Protocol (HTTP/1.1): Caching", RFC 7234, June 2014.
- [RFC7235] Fielding, R. and J. Reschke, "Hypertext Transfer Protocol (HTTP/1.1): Authentication", RFC 7235, June 2014.

[I-D.ietf-lmap-information-model]
Burbridge, T., Eardley, P., Bagnulo, M., and J.
Schoenwaelder, "Information Model for Large-Scale
Measurement Platforms (LMAP)", draft-ietf-lmap-
information-model-02 (work in progress), August 2014.

10.2. Informative References

[I-D.bagnulo-ippm-new-registry-independent]
Bagnulo, M., Burbridge, T., Crawford, S., Eardley, P., and
A. Morton, "A registry for commonly used metrics.
Independent registries", draft-bagnulo-ippm-new-registry-
independent-01 (work in progress), July 2013.

[I-D.ietf-lmap-framework]
Eardley, P., Morton, A., Bagnulo, M., Burbridge, T.,
Aitken, P., and A. Akhter, "A framework for large-scale
measurement platforms (LMAP)", draft-ietf-lmap-
framework-08 (work in progress), August 2014.

Authors' Addresses

Marcelo Bagnulo
Universidad Carlos III de Madrid
Av. Universidad 30
Leganes, Madrid 28911
SPAIN

Phone: 34 91 6249500
Email: marcelo@it.uc3m.es
URI: <http://www.it.uc3m.es>

Trevor Burbridge
British Telecom
Adastral Park, Martlesham Heath
IPswitch
ENGLAND

Email: trevor.burbridge@bt.com

Sam Crawford
SamKnows

Email: sam@samknows.com

Juergen Schoenwaelder
Jacobs University Bremen
Campus Ring 1
28759 Bremen
Germany

Email: j.schoenwaelder@jacobs-university.de

Vaibhav Bajpai
Jacobs University Bremen
Campus Ring 1
28759 Bremen
Germany

Email: v.bajpai@jacobs-university.de

Network Working Group
Internet-Draft
Intended status: Informational
Expires: April 24, 2014

T. Burbridge
P. Eardley
British Telecom
M. Bagnulo
Universidad Carlos III de Madrid
J. Schoenwaelder
Jacobs University
October 21, 2013

Information Model for Large-Scale Measurement Platforms (LMAP)
draft-burbridge-lmap-information-model-01

Abstract

This Information Model applies to the Measurement Agent within a Large-Scale Measurement Platform. As such it outlines the information that is (pre-)configured on the MA or exists in communications with a Controller or Collector within an LMAP framework. The purpose of such an Information Model is to provide a protocol and device independent view of the MA that can be implemented via one or more Control and Report protocols.

Requirements Language

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].

Status of this Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <http://datatracker.ietf.org/drafts/current/>.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on April 19, 2014.

Copyright Notice

Copyright (c) 2013 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1. Introduction	3
2. LMAP Information Model	3
2.1. Information Structure	4
2.2. Pre-Configuration Information	5
2.3. Configuration Information	5
2.4. Instruction Information	6
2.5. Logging Information	9
2.6. Status Information	10
2.7. Reporting Information	11
2.8. Channels	12
2.9. Timing Information	13
2.9.1. Periodic Timing	14
2.9.2. Calendar Timing	14
2.9.3. One-Off Timing	15
2.9.4. Immediate Timing	15
2.9.5. Timing Randomness	15
3. IANA Considerations	15
4. Security Considerations	16
5. Acknowledgements	16
6. Informative References	16
Authors' Addresses	16

1. Introduction

A large-scale measurement platform is a collection of components that work in a coordinated fashion to perform measurements from a large number of vantage points. The main components of a large-scale measurement platform are the Measurement Agents (hereafter MAs), the Controllers and the Collectors.

The MAs are the elements actually performing the measurements. The MAs are controlled by one or more Controllers and the Collectors gather the results generated by the MAs. In a nutshell, the normal operation of a large-scale measurement platform starts with the Controller instructing a set of MAs to perform a set of measurements at a certain point in time. The MAs execute the instructions from the Controller and once they have done so they report the results of the measurements to the Collector. The overall framework for a Large Measurement platform and the terminology used in this document is described in detail in [I-D.ietf-lmap-framework].

A large-scale measurement platform involves basically three protocols, namely, a Control protocol between the Controller(s) and the MAs, a Report protocol between the MAs and the Collector(s) and several measurement protocols between the MAs used to actually perform the measurements. In addition the some information is required to be provisioned to the MA prior to any communication with the Controller.

This document defines the information model for both the Control and the Report protocol along with pre-configuration information that is required before communicating with the Controller, broadly named as the LMAP Information Model (or LMAP IM for short). The measurement protocols are out of the scope of this document.

As defined in [RFC3444], the LMAP IM defines the concepts involved in a large-scale measurement platform at a high level of abstraction, independently of any specific implementation or actual protocol used to exchange the information. It is expected that the proposed information model can be used with different protocols in different measurement platform architectures and across different types of MA device (e.g. home gateway, smartphone, PC, router etc.).

2. LMAP Information Model

2.1. Information Structure

The information described herein relates to the information stored, received or transmitted by a Measurement Agent as described within the LMAP framework [I-D.ietf-lmap-framework]. As such, some subsets of this information model are applicable to the measurement Controller, Collector and systems that pre-configure the Measurement Agent. The information described in these models will be transmitted across the protocols and interfaces between the Measurement Agent and such systems according to a Data Model.

For clarity the information model is divided into six sections:

1. Pre-Configuration Information. Information pre-configured on the Measurement Agent prior to any communication with other components of the LMAP architecture, specifically detailing how to register with a Controller
2. Configuration Information. Information delivered to the MA on registration with a Controller or updated during a later communication, in particular detailing how to retrieve measurement and reporting instruction information from a Controller along with information specifically about the MA
3. Instruction Information. Information that is received by the MA from the Controller pertaining to the measurement and reporting configuration. This includes measurement configuration, report channel configuration, measurement schedules and measurement suppression information
4. Logging Information. Information transmitted from the MA to the Controller detailing the results of any configuration operations along with error and status information from the operation of the MA
5. Status Information. Information on the general status and capabilities of the MA. For example, the set of measurements that are supported on the device
6. Reporting Information. Information transmitted from the MA to the Collector including measurement results and the context in which they were conducted

In addition the MA may hold further information not described herein, and which may be optionally transferred to or from other systems including the Controller and Collector. One example of information in this category is subscriber or line information that may be reported by the MA as optional fields in the reporting communication

to the Collector.

2.2. Pre-Configuration Information

This information is the minimal information that needs to be pre-configured to the MA in order for it to successfully communicate with a Controller during the registration process.

This pre-configuration information needs to include an URL of the Controller where configuration information can be retrieved along with the security information required for the communication including the certificate of the Controller (or the certificate of the Certification Authority which was used to issue the certificate for the Controller) as well as the timing for that communication. All this is expressed as the Configuration Channel. In addition to the Configuration Channel information, the MA's security information is configured which can be either a certificate and a private key or a password, depending on the security solution used.

Detail of the information model elements:

1. MA MAC: MAC Address
2. Configuration Channel: Channel
3. MA Certificate: Certificate (optional)
4. MA ID: random UUID (optional)
5. MA password: string (optional)

The detail of the Channel object is described later since it is common to several parts of the information model.

2.3. Configuration Information

During registration or at any later point at which the MA contacts the Controller, the choice of Controller and details for the timing of communication with the Controller can be changed. For example the pre-configured Controller may be replaced with a specific Controller that is more appropriate to the MA device type, location of characteristics of the network (e.g. access technology type or broadband product). The initial communication timing object may also be replaced with one more relevant to routine communications between the MA and the Controller.

In addition the MA will be given further items of information that relate specifically to the MA rather than the measurements it is to

conduct or how to report results. The assignment of an ID to the MA is mandatory. Optionally a Group ID may also be given which identifies a group of interest to which that MA belongs. For example the group could represent an ISP, broadband product, technology, market classification, geographic region, or a combination of multiple such characteristics. Where the Measurement Group ID is set an additional flag (the Report MA ID flag) is required to control whether the Measurement Agent ID is to be reported. This allows the MA to remain anonymous which may be particularly useful to prevent tracking of mobile MA devices.

The configuration information will also contain information about different communication channels that the MA will have with different elements of the infrastructure. Each channel specifies a URL, security information and timing information for the communication.

Detail of the additional information model elements:

1. Measurement Agent ID: UUID
2. Measurement Group ID (optional): String
3. Report MA ID flag (optional): Boolean
4. Instruction Channel: Channel (DISCUSSION: shouldn't we split this into 4 different channels i.e. the Measurement Task Configuration channel, the Report Channel channel, the Measurement Schedules channel and the Measurement Suppression channel?)
5. Status Channel: Channel
6. Logging Channel: Channel

2.4. Instruction Information

The Instruction information model has four sub-elements:

1. Measurement Task Configurations: Set
2. Report Channels: Set
3. Measurement Schedules: Set
4. Measurement Suppression: Object

Conceptually each Measurement Task Configuration defines the parameters of a Measurement Task that the Measurement Agent (MA) may perform at some point in time. It does not by itself actually

instruct the MA to perform them at any particular time (this is done by a Measurement Schedule).

Example: A Measurement Task Configuration may configure a single Measurement Task for measuring UDP latency. The Measurement Task Configuration could define the destination port and address for the measurement as well as the duration, internal packet timing strategy and other parameters (for example a stream for one hour and sending one packet every 500 ms). It may also define the output type and possible parameters (for example the output type can be the 95th percentile mean) where the measurement task accepts such parameters. It does NOT define when the task starts (this is defined by the Measurement Schedule element), so it does not by itself instruct the MA to actually perform this measurement task.

The Measurement Task Configuration will include a local short name for reference by the Measurement Schedule, along with a registry entry [I-D.bagnulo-ippm-new-registry] that defines the Measurement Task. The MA itself will resolve the registry entry to a local executable program. In addition the Measurement Task is specialised through a set of configuration Options. The nature and number of these Options will depend upon the Measurement Task and will be defined in the Measurement Task Registry. In addition the Measurement Task Configuration may optionally also be given a Measurement Cycle ID. The purpose of this ID is to easily identify a set of measurement results that have been produced by Measurement Tasks with comparable Options. This ID is manually incremented when an Option change is implemented which could mean that two sets of results should not be directly compared.

A Report Channel defines how to report results to a single Collector. Several Report Channels can be defined to enable results to be split or duplicated across different report intervals or destinations. E.g. a single Collector may have three Report Channels, one reporting hourly, another reporting daily and a third on which to send immediate results for on-demand measurement tasks. The details of the Channel element is described later as it is common to several objects.

A Measurement Schedule contains the instruction from the Controller to the MA to execute a single or repeated series of Measurement Tasks. Each Measurement Schedule contains basically three elements: a reference to a list of Measurement Task Configuration, a reference to a set of one or more Report Channels, and a timing object for the schedule. The schedule basically states what measurement task to run, how to report the results, and when to run the measurement task. Multiple measurement tasks in the list will be executed in order with

minimal gaps. Note that the Controller can instruct the MA to report to several Collectors by specifying several Report Channels.

Example: a Measurement Schedule references a single Measurement Task Configuration for the UDP latency defined in the previous example. It references the Report Channel in the previous example to send results immediately as available to the specified Collector. The timing is specified to run the configured Measurement Task Configuration every hour at 23 minutes past the hour.

Measurement Suppression information is used to over-ride the Measurement Schedule and stop measurements from the MA for a defined or indefinite period. While conceptually measurements can be stopped by simply removing them from the Measurement Schedule, splitting out separate information on Measurement Suppression allows this information to be updated on the MA on a different timing cycle or protocol implementation to the Measurement Schedule.

The goal when defining these four different elements is to allow each part of the information model to change without affecting the other three elements. For example it is envisaged that the Report Channels and the set of Measurement Tasks Configurations will be relatively static. The Measurement Schedule on the other hand is likely to be more dynamic as the measurement panel and test frequency are changed for various business goals. Another example is that measurements can be suppressed with a Measurement Suppression command without removing the existing Measurement Schedules that would continue to apply after the Measurement Suppression expires or is removed. In terms of the Controller-MA communication this can reduce the data overhead. It also encourages the re-use of the same standard Measurement Task Configurations and Reporting Channels to help ensure consistency and reduce errors.

Definition of the information model elements:

1. Measurement Task Configurations: Set

1. Measurement Task Configuration: Object

1. Task Name (used for referral from the Measurement Schedules): String
2. Registry Entry: URN
3. Options: Set (optional)
 1. Interface name (reference by name to one of the Interfaces defined in the Status information): String

4. Measurement Cycle ID: String (optional)
 2. Report Channels: Set
 1. Report Channel: Channel
 3. Measurement Schedules: Set
 1. Measurement Schedule: Object
 1. Schedule Name: String
 2. Measurement Task Configuration Names (reference by Name to one of the measurement tasks defined in the Measurement Task Configuration set): List
 1. Task Name: String
 3. Report Channel Names (reference by Name to one of the measurement tasks defined in the Measurement Task Configuration set): Set
 1. Channel Name: String
 4. Measurement Timing: Timing
 4. Measurement Suppression: Object (optional)
 1. Start: datetime
 2. End: datetime
 3. Set of Measurement Task Configuration Names (optional - default all)
 1. Task Name: String
- 2.5. Logging Information
- The MA will report back success/failure and status information to the Controller. These messages will fall into a number of different categories:
1. Success/failure messages in response to information updates from the Controller. For example:
 - * "Report Channel 'hourly db' configured"

- * "Measurement Schedule does not conform to schema, Row 211"
- 2. Status updates from the operation of the MA. For example:
 - * "out of memory: cannot record result"
 - * "Collector 'collector.example.com' not responding"

Each log message will have the following Information model elements:

1. Log Time: datetime
2. Log Event: Object

2.6. Status Information

In addition to the information reported by the MA through the logging information, the MA will hold further status information that can be retrieved by a Controller. One category of additional information that has not been defined in earlier sections is the availability of Measurement Tasks on that MA.

MA Status information model elements:

1. MA ID: String
2. MA Device: String
3. MA hardware: String (optional)
4. MA firmware: String (optional)
5. MA software: String (optional)
6. MA Interfaces: set
 1. If name: String
 2. If type: String (one of eth, wlan, TBC)
 3. If speed: Integer (expressed in Mbps)
 4. Link Layer Address: String
 5. IP address: Set
 1. Protocol: String (one of v4, v6)

- 2. Address: String
- 6. Gateway: Set (optional)
 - 1. Protocol: String (one of v4, v6)
 - 2. Address: String
- 7. DNS server: Set (optional)
 - 1. Protocol: String (one of v4, v6)
 - 2. Address: String
- 7. Last Measurement: datetime
- 8. Last Report: datetime
- 9. Last Instruction: datetime
- 10. Last Configuration: datetime
- 11. Supported Measurements: Set
 - 1. Registry Entry: URN
 - 2. Version: String (optional)

2.7. Reporting Information

At a point in time specific by the Report Channel, the MA will communicate a set of measurement results to the Collector. These measurement results should be communicated within the context in which they were collected.

The report is structured hierarchically to avoid repetition of report, Measurement Agent and Measurement Task Configuration information. The report starts with the timestamp of the report generation on the MA and details about the MA including the optional Measurement Agent ID and Group ID (controlled by the Configuration Information). In addition optional further MA context information can be included at this point such as the line sync speed or ISP and product if known by the MA.

After the MA information the results are reported grouped into the different Measurement Tasks. Each Task starts with replicating the Measurement Task Configuration information before the result headers (titles for data columns) and the result data rows.

Information model elements:

1. Report Date: datetime
2. Measurement Agent ID: String (optional)
3. Measurement Group ID: String (optional)
4. MA Context: Set (optional)
 1. Context Item: Object
5. Measurement Task: Set
 1. Measurement Task Configuration: Object
 2. Result Headers: List
 1. Column Name: String
 3. Result Data: List
 1. Result Row: Object
 1. Measurement Time: datetime
 2. Cross-traffic: Integer (optional)
 3. Result Columns: List
 1. Column Data

2.8. Channels

A Channel defines a communication channel between the MA and other element of the measurement framework i.e. with the Collector to report results back, to Controller to retrieve Instructions or other information exchanged between the parties. Several Channels can be defined to enable results to be split or duplicated across different report intervals or destinations. E.g. a single Collector may have three Report Channels, one reporting hourly, another reporting daily and a third on which to send immediate results for on-demand measurement tasks.

Each Channel contains the details of the target (including location and security information such as the certificate), and the timing for the communication i.e. when to establish the communication. The certificate can be the digital certificate associated to the FQDN in

the URL or it can be the certificate of the Certification Authority that was used to issue the certificate for the FQDN of the target URL (which will be retrieved later on using a communication protocol such as SSL). The Channel can use the same timing information object as a Measurement Schedule and the Controller Communication Timing defined earlier. There are several options, such as immediately after the results are obtained or at a given interval or calendar based cycle). As with the Measurement task Configuration, each Channel is also given a local short name by which it can be referenced from a Measurement Schedule or other elements.

Example: A Channel using for reporting results may specify that results are to be sent to the URL (<https://collector.foo.org/report/>), using the appropriate digital certificate to establish a secure channel. The Channel specifies that the results are to be sent immediately as available and not batched.

Channel: Object

1. Channel Name (used for referral from other objects): String
2. Target: URL
3. Certificate: X.509 Certificate
4. Communication Timing: Timing

2.9. Timing Information

The Timing information object used throughout the information models can take one of four different forms:

1. Periodic. Specifies a start, end and interval time in milliseconds
2. Calendar: Specifies a calendar based pattern - e.g. 22 minutes past each hour of the day on weekdays
3. One Off: A single instance occurring at a specific time
4. Immediate: Should occur as soon as possible

Optionally each of the first three options may also specify a randomness that should be evaluated and applied separately to each indicated event.

2.9.1. Periodic Timing

Information model elements:

1. 1. Timing Name: String
2. 2. Start: datetime (optional)
3. 3. End: datetime (optional)
4. 4. Interval: Integer (in milliseconds)
5. 5. Randomness: Timing Randomness (optional)

2.9.2. Calendar Timing

Information model elements:

1. Timing Name: String
2. Start: datetime (optional)
3. End: datetime (optional)
4. Months: Set (optional - default [1-12])
 1. Month: Integer
5. Weekdays: Set (optional - default [Mon-Sun])
 1. Weekday: String (one off Mon, Tue, Wed, Thu, Fri, Sat Sun)
6. Days: Set (optional - default [1-31])
 1. Day: Integer
7. Hours: Set (optional - default [1-24])
 1. Hour: Integer
8. Minutes: Set (optional - default [1-60])
 1. Minute: Integer
9. Seconds: Set (optional - default [1-60])
 1. Second: Integer

10. Randomness: Timing Randomness (optional)

2.9.3. One-Off Timing

Information model elements:

1. Time: datetime
2. Randomness: Timing Randomness (optional)

2.9.4. Immediate Timing

The immediate timing object has no further information elements. The measurement or report is simply to be done as soon as possible.

2.9.5. Timing Randomness

The Timing randomness object specifies a random distribution that can be applied to any scheduled execution event such as a measurement or report. The intention is to be able to spread the load on the Controller, Collector and network in an automated manner for a large number of Measurement Agents. The randomness is expressed as a distribution (e.g. Poisson, Normal, Uniform etc.) along with the spread over which the distribution should be applied. In addition optional upper and lower bounds can be applied to control extreme spread of timings.

Information model elements:

1. Distribution: String
2. Upper Cut: Integer (optional)
3. Lower Cut: Integer (optional)
4. Spread: Integer

3. IANA Considerations

This document makes no request of IANA .

Note to RFC Editor: this section may be removed on publication as an RFC.

4. Security Considerations

This Information Model deals with information about the control and reporting of the Measurement Agent. There are broadly two security considerations for such an Information Model. Firstly the Information Model has to be sufficient to establish secure communication channels to the Controller and Collector such that other information can be sent and received securely. The second consideration is that no mandated information items pose a risk to confidentiality or privacy given such secure communication channels. For this latter reason items such as the MA context and MA ID are left optional and can be excluded from some deployments. This would, for example, allow the MA to remain anonymous and for information about location or other context that might be used to identify or track the MA to be omitted or blurred.

5. Acknowledgements

6. Informative References

- [I-D.bagnulo-ippm-new-registry]
Bagnulo, M., Burbridge, T., Crawford, S., Eardley, P., and A. Morton, "A registry for commonly used metrics", draft-bagnulo-ippm-new-registry-00 (work in progress), January 2013.
- [I-D.ietf-lmap-framework]
Eardley, P., Morton, A., Bagnulo, M., Burbridge, T., Aitken, P., and A. Akhter, "A framework for large-scale measurement platforms (LMAP)", draft-ietf-lmap-framework-00 (work in progress), October 2013.
- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.
- [RFC3444] Pras, A. and J. Schoenwaelder, "On the Difference between Information Models and Data Models", RFC 3444, January 2003.

Authors' Addresses

Trevor Burbridge
British Telecom
Adastral Park, Martlesham Heath
Ipswich, IP5 3RE
UK

Phone:
Fax:
Email:
URI:

Philip Eardley
British Telecom
Adastral Park, Martlesham Heath
Ipswich, IP5 3RE
UK

Phone:
Fax:
Email:
URI:

Marcelo Bagnulo
Universidad Carlos III de Madrid
Av. Universidad 30
Leganes, Madrid, 28911

Phone:
Fax:
Email:
URI:

Juergen Schoenwaelder
Jacobs University

Phone:
Fax:
Email:
URI:

Network Working Group
Internet-Draft
Intended status: Informational
Expires: January 16, 2014

P. Eardley
T. Burbridge
BT
A. Morton
AT&T Labs
July 15, 2013

A framework for large-scale measurements
draft-eardley-lmap-framework-02

Abstract

Measuring broadband service on a large scale requires standardisation of the logical architecture and a description of the key protocols that coordinate interactions between the components. The document presents an overall framework for large-scale measurements and discusses which elements could be standardised in the IETF. It is intended to assist the work of the LMAP working group.

Status of this Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <http://datatracker.ietf.org/drafts/current/>.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on January 16, 2014.

Copyright Notice

Copyright (c) 2013 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must

include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1. Introduction	3
2. Outline of framework	4
3. Constraints	6
3.1. Measurement system is under the direction of a single organisation	6
3.2. Each MA may only have a single Controller at any point in time	7
3.3. A Measurement Agent acts autonomously	7
4. Work items for LMAP WG	8
4.1. Information Model	9
4.2. Control Protocol	10
4.3. Report Protocol	10
5. Related work required but out of scope of LMAP	10
5.1. Standard measurement tests	10
5.2. Characterisation plan	11
5.3. Other elements	11
6. IANA Considerations	12
7. Security Considerations	12
8. Acknowledgements	13
9. Changes	13
9.1. from -00 to -01	13
10. Informative References	14
Authors' Addresses	14

1. Introduction

The Large-Scale Measurement of Broadband Performance (LMAP) working group standardizes the LMAP measurement system for performance measurements of broadband access devices such as home and enterprise edge routers, personal computers, mobile devices, set top box, whether wired or wireless. Measuring portions of the Internet on a large scale is essential for accurate characterizations of performance over time and geography.

[use-cases] discusses several use cases have been proposed for large-scale measurements:

- o Operators: to help plan their network and identify faults
- o End Users: to run diagnostic checks, such as a network speed test
- o Regulators: to benchmark several network operators and support public policy development

The LMAP framework should be useful for all these.

The goal is to have the measurements (made using the same metrics and mechanisms) for a large number of points on the Internet, and to have the results collected and stored in the same form.

There are existing measurement systems. However, they typically lack some of the desirable features for a large-scale measurement system:

- o Standardised - in terms of the tests that they perform, the components, the data models and protocols for transferring information between the components. For example so that it is meaningful to compare measurements made of the same metric at different times and places. For example so that the operator of a measurement system can buy the various components from different vendors. Today's systems are proprietary in some or all of these aspects.
- o Extensible - it should be easy to add or modify tests, for example an improved test methodology or to measure a performance metric not previously considered important (e.g., bufferbloat).
- o Large-scale - [use-cases] envisages Measurement Agents in every home gateway and edge device such as set-top-boxes and tablet computers. Existing systems have up to a few thousand Measurement Agents (without judging how much further they could scale).

- o Diversity - a measurement system should handle different types of Measurement Agent - for example Measurement Agents may come from different vendors, be in wired and wireless networks and be on devices with IPv4 or IPv6 addresses.

2. Outline of framework

The LMAP framework for large-scale measurements has four elements:

- o Measurement Agent (MA)
- o Measurement Peer
- o Controller
- o Collector

In addition there are some components that are outside LMAP but useful within the context of a large scale measurement system:

- o Initialiser
- o Subscriber Parameter Database
- o Results Database
- o Data Analysis Tools
- o Operator's OAM (Operations Administration and Management)

a large-scale measurement system essentially has three sets of communications:

- o several measurement protocols between a Measurement Agent and a Measurement Peer
- o a Control Protocol between a Controller and a MA
- o a Report Protocol between a MA and a Collector.

A Measurement Agent and a Measurement Peer jointly perform an active measurement test, by generating test traffic and measuring some metric associated with its transfer over the path from one to the other; for example the time taken to transfer a 'test file'. A MA may also conduct passive testing through the observation of traffic (i.e. without the involvement of a Measurement Peer); for example an end user's mix of applications.

The MA interacts with the Controller and Collector, and a Measurement Peer only takes part in active tests (and does not interact with the Controller and Collector).

The MA functions are implemented either in specialised hardware or as code on general purpose devices like a PC, tablet or smartphone. The Measurement Peer may be an LMAP device or a normal, non-LMAP device (for example if the MA measures the time for a DNS response or a webpage download from www.example.com).

The Controller manages a MA by instructing it which tests it should perform and when. For example it may instruct a MA at a home gateway: "Run the 'download speed test' with the test server at the end user's first IP point in the network; if the end user is active then delay the test and re-try 1 minute later, with up to 3 re-tries; repeat every hour at $xx.05 + \text{Unif}[0,180]$ seconds". The Controller also manages a MA by instructing it how to report the test results, for example: "Report results once a day in a batch at 4am + $\text{Unif}[0,180]$ seconds; if the end user is active then delay the report 5 minutes". As well as regular tests, a Controller can initiate a one-off test ("Do test now", "Report as soon as possible"). These are called the Test and Report Schedule.

The Collector accepts a Report from a MA with the results from its tests. It may also do some processing on the results - for instance to eliminate outliers, as they can severely impact the aggregated results.

Therefore the MA is a LMAP-specific device that initiates the test, gets instructions from the Controller and reports to the Collector.

It is possible that communications between two Collectors, two Controllers and a Controller and Collector may be useful in some use cases, perhaps to help a measurement system scale. Work on such a protocol is out of scope of LMAP (?)

The Initialiser, Subscriber Parameter Database, Results Database, Data Analysis Tools and OAM are out of scope of LMAP. They may be provided through existing protocols or applications and are likely to be part of a complete large-scale measurement system. See Section 5 for further discussion.

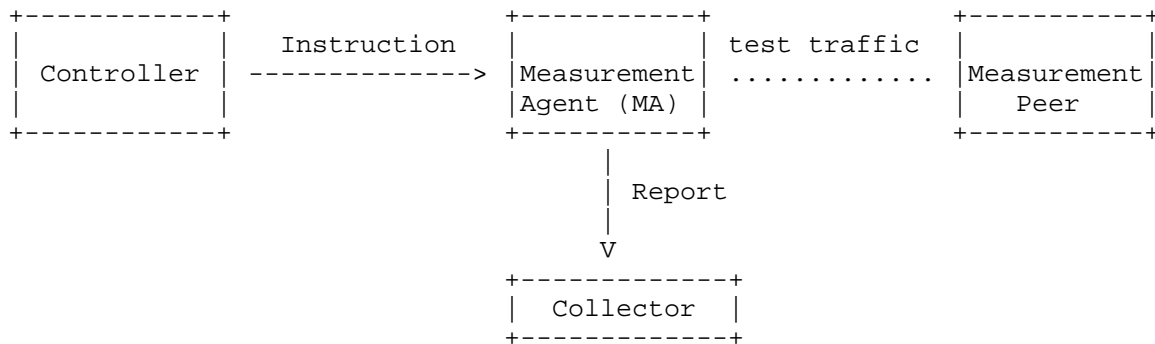


Figure 1: Schematic of main elements of LMAP framework

3. Constraints

3.1. Measurement system is under the direction of a single organisation

In the LMAP framework (as defined in the WG's charter) the measurement system is under the direction of a single organisation that is responsible both for the data and the quality of experience delivered to its users. Clear responsibility is critical given that a misbehaving large-scale measurement system could potentially harm user experience, user privacy and network security.

However, the components of an LMAP measurement system can be deployed in administrative domains that are not owned by the measuring organisation. Thus, the system of functions deployed by a single organisation constitutes a single LMAP domain which may span ownership or other administrative boundaries.

Note that different LMAP measurement systems may overlap, in the sense that the active measurement packets of one measurement system appear along with normal user traffic to another measurement system. For instance, imagine an operator with an MA on the home gateway and an end user with an MA on their laptop. Rather than making separate measurements, an organisation might share its measurement data, or a suitably anonymised version of it, with another organisation. However, any form of coordination between different organisation involves difficult commercial and technical issues and so, given the novelty of large-scale measurement efforts, any form of inter-organisation coordination is outside the scope of LMAP.

3.2. Each MA may only have a single Controller at any point in time

The constraint avoids different Controllers giving a MA conflicting instructions and so means that the MA does not have to manage contention between multiple Test (or Report) Schedules. This simplifies the design of MAs (critical for a large-scale infrastructure) and allows a Test Schedule to be tested on specific types of MA before deployment to ensure that the home user experience is not impacted (due to CPU, memory or broadband-product constraints).

An operator may have several Controllers, perhaps with a Controller for different types of MA (home gateways, tablets) or location (Ipswich, Edinburgh).

To avoid problems with NAT and firewalls, it is likely that the MA 'pulls' the configuration from its Controller, as identified by the Initialiser.

- o Open issue: Should there be negotiation between a Controller and its MA, or should the Controller simply instruct the MA by sending its Test and Report Schedules?
 - * The argument for negotiation is that occasionally the MA may be updated with enhanced with versions of existing tests. It is easier for the Controller to learn the MAs capabilities directly from the MA than from a management system. It avoids any mis-synchronisation.
 - * The argument against negotiation is that it makes the Controller-MA protocol more complicated, increases the MAs resource requirements and increases the complexity of the Controller when it decides how to schedule tests across numerous MAs or when it deploys a new Test Schedule to potentially millions of MAs.
- o Open issue: what happens if a Controller fails, how is the MA is homed onto a new one?

3.3. A Measurement Agent acts autonomously

Once the MA gets its Test and Report Schedules from its Controller then it acts autonomously, in terms of operation of the tests and reporting of the result.

Firstly, this means that the MA initiates Measurement Tasks. For the typical case where the MA is on a home gateway or edge device, this means that the MA initiates a 'download speed test' by asking a

Measurement Peer to send the file. The main rationale is that, for a test that should be performed when there is no user traffic on the link, the MA knows whether the end user is active and therefore whether to start the test or delay it. Having the Schedule on the MA also avoids it having to check frequently with the Controller. Further, if the MA is behind a NAT then the Measurement Peer naturally learns its public-facing IP address.

Secondly, it is useful for the MA and perhaps the Measurement Peer to make some 'admission control' checks at the initiation of the Measurement Task to ensure that desired test conditions are present. The exchange of initialization packets between the MA and Measurement Peer ensures basic connectivity between them. Also, the MA may delay Measurement Task may if the associated subscriber is active, or the Measurement Peer may reject a testing request if it is overloaded. It has also been suggested that, in extremis, the Controller may want the ability to send a Measurement Suppression message to an MA, which causes the Measurement Tasks to be temporarily stopped.

Last, it is easier to secure the reporting process, for example with a unique certificate for each MA-Collector pair, so that the Collector is confident the results really do originate from the MA. All measurement results are sent from the MA.

4. Work items for LMAP WG

This Section considers the work that the LMAP working group needs to tackle. Section 5 considers other work that needs doing that would be beyond the scope of the LMAP WG.

The main work items are:

- o Information Model, the abstract definition of the information carried from the Controller to the MA and the information carried from the MA to the Collector.
- o Control protocol and the associated data model: The definition of how instructions are delivered from a Controller to a MA; this includes a Data Model consistent with the Information Model plus a transport protocol.
- o Report protocol and the associated data model: The definition of how the Report is delivered from a MA to a Collector; this includes a Data Model consistent with the Information Model plus a transport protocol.

4.1. Information Model

The Information Model provides a protocol and device independent view of the information carried from the Controller to the MA and the information carried from the MA to the Collector. It can be implemented via a Control Protocol and Report Protocol, as defined by the LMAP WG. It is also possible that other Control and Report Protocols could be defined by other standards bodies or proprietary, however it is important that they all implement the same Information Model, in order to ease the definition, operation and interoperability of large-scale measurement systems.

The Information Model also includes information that is pre-configured on the MA in order that it can start communicating with a Controller.

An initial proposal for the Information Model is in [information-model].

The Information Model is divided into two main parts, each of which may be broken down into sub-parts:

- o information about the Instruction: Information that is received by the MA from the Controller pertaining to the measurement and reporting configuration. This includes:
 - * what measurements to do: the Measurement Task could be defined by reference to a registry entry, along with any parameters that need to be set (such as the address of the Measurement Peer) and any Environmental Constraint (such as, delay the test if the end user is active)
 - * when to do them: the Measurement Schedule details the timings of regular tests, one-off tests, and if regularly tests should be temporarily suppressed
 - * how to report the Measurement Results: via Reporting Channel(s), each of which defines a target Collector and Report Schedule
- o information about the Report: Information transmitted from the MA to the Collector including measurement results and the context in which they were conducted. This includes:
 - * the MAs identifier, or perhaps a Group-ID to anonymise results
 - * the actual Measurement Results, including the time they were measured

- * the details of the Measurement Task (to avoid the Collector having to ask the Controller for this information later)

It is important to consider how to divide the Information Model into (sub-)parts, so that each (sub-)part can be updated independently at different times and regularities, as discussed in [information-model]

4.2. Control Protocol

The Control protocol and its associated data model define how instructions are delivered from a Controller to a MA; this includes a Data Model consistent with the Information Model plus a transport protocol. This may be a simple instruction - response protocol, and LMAP will specify how it operates over an existing protocol (to be selected, perhaps REST-style HTTP(s) or NETCONF).

4.3. Report Protocol

The Report protocol and the associated data model: The definition of how the Report is delivered from a MA to a Collector; this includes a Data Model consistent with the Information Model plus a transport protocol (to be selected, perhaps REST-style HTTP(s) or IPFIX).

5. Related work required but out of scope of LMAP

This section considers the items that need to be agreed between deployers of large-scale measurement systems, but that are out of scope of the LMAP WG (Section 4 considers items within its scope).

5.1. Standard measurement tests

Standardised methods are needed for each metric that is measured. A registry for commonly-used metrics [registry] is also required, so that a test can be defined simply by its identifier in the registry. The methods and registry would hopefully also be referenced by other standards organisations.

- o Such activities are in scope of the IPPM working group (possibly re-chartered) and not LMAP.

A new (or revised) test may need to be uploaded to MAs. How this is done is out of scope of the IETF; it could be as a firmware upgrade for a home hub, or new app for a PC, etc and may be device-specific.

5.2. Characterisation plan

Each organisation operating an LMAP system and collecting measurements for comparison purposes needs to conduct the same measurements according to the same sampling plan (ie size and schedule) and make the results available in the same format. The scope of comparison determines the set of organisations needing to agree on the common characterisation plan; for example those falling within the same regulatory environment in a particular country or region. Such agreements are certainly facilitated by IETF's work, but the details of the plan are beyond the scope of work in IETF.

5.3. Other elements

Other elements may be useful within the context of a large scale measurement system and worthy of standardisation, but are outside the scope of the LMAP WG: Initialiser, Subscriber Parameter Database, Results Database, Data Analysis Tools and operator's OAM.

An Initialiser configures a MA with details about its Controller, including authentication credentials. A bootstrap protocol is likely to be technology specific and so for different types of device could be defined by the Broadband Forum, DOCSIS or IEEE. Possible protocols are SNMP, NETCONF or (for Home Gateways) CPE WAN Management Protocol (CWMP) from the Auto Configuration Server (ACS) (as specified in TR-069).

A Subscriber Parameter Database contains information about the line, for example the customer's broadband contract (2, 40 or 80Mb/s), the line technology (DSL or fibre), the time zone where the MA is located, and the type of home gateway and MA. These are all factors which may affect the choice of what Measurement Tasks to run and how to interpret the Measurement Results. For example, a download test suitable for a line with an 80Mb/s contract may overwhelm a 2Mb/s line. Another example is if the Controller wants to run a one-off test to diagnose a fault, then it should understand what problem the customer is experiencing and what tests have already been run. The subscribers' service parameters are already gathered and stored by existing operations systems.

A Results Database records all measurements in an equivalent form, for example an SQL database [schulzrinne], so that they can be easily accessed by the Data Analysis Tools whilst the LMAP system implementor can choose local solutions for each component. The Data Analysis Tools also need to understand subscriber service information, for example the broadband contract.

The Data Analysis Tools receive the results from the Collector or via

the Results Database. They might visualise the data or identify which component or link is likely to be the cause of a fault or degradation.

The operator's OAM (Operations, Administration and Management) uses the results from the tools.

6. IANA Considerations

This document makes no request of IANA.

Note to RFC Editor: this section may be removed on publication as an RFC.

7. Security Considerations

The security of the LMAP framework should protect the interests of the measurement operator(s), the network user(s) and other actors who could be impacted by a compromised measurement deployment.

We assume that each Measurement Agent will receive test configuration, scheduling and reporting instructions from a single organisation (operator of the Controller). These instructions must be authenticated (to ensure that they come from the trusted Controller), checked for integrity (to ensure no-one has tampered with them) and be prevented from replay. If a malicious party can gain control of the Measurement Agent they can use the MA capabilities to launch DoS attacks at targets, reduce the network user experience and corrupt the measurement results that are reported to the Collector. By altering the tests that are operated and/or the Collector address they can also compromise the confidentiality of the network user and the MA environment (such as information about the location of devices or their traffic).

The reporting of the MA must also be secured to maintain confidentiality. The results must be encrypted such that only the authorised Collector can decrypt the results to prevent the leakage of confidential or private information. In addition it must be authenticated that the results have come from the expected MA and that they have not been tampered with. It must not be possible to spoof an MA to inject falsified data into the measurement platform or to corrupt the results of a real MA.

Availability should also be considered. While the loss of some MAs may not be considered critical, the unavailability of the Collector could mean that valuable business data or data critical to a

regulatory process is lost. Similarly, the unavailability of a Controller could mean that the MAs continue to operate an incorrect test schedule or fail to initiate.

A malicious party could "game the system". For example, where a regulator is running a measurement system in order to benchmark operators, an operator could try to identify the broadband lines that the regulator was measuring and prioritise that traffic. This potential issue is currently handled by a code of conduct. It is outside the scope of the LMAP WG to consider the issue.

Concerning privacy and data protection, the role of the LMAP framework should be to ensure that only authorised data is collected and that this data is returned securely to the framework operator. Data should be stored securely and onward sharing of data to other parties should be controlled according to local data protection regulations. Depending upon the ownership/placement of the MA, local data protection laws, the tests being operated and existing user agreements, it is possible that additional consent may need to be secured from parties such as the home broadband user. Having the measurement system under the direction of a single organisation clarifies the responsibility for data protection.

The next versions of [lmap-yang] and [lmap-ipfix] will also include further consideration of security.

8. Acknowledgements

Thanks to numerous people for much discussion, directly and on the LMAP list. This document tries to capture the current conclusions.

Philip Eardley and Trevor Burbridge work in part on the Leone research project, which receives funding from the European Union Seventh Framework Programme [FP7/2007-2013] under grant agreement number 317647.

9. Changes

9.1. from -00 to -01

aligned with terminology in draft-eardley-lmap-terminology

introduced aspects mentioned in the LMAP WG charter

introduced aspects from the Information model in draft-burbridge-lamp-information-model

10. Informative References

- [RFC6241] "Network Configuration Protocol (NETCONF)",
<<http://tools.ietf.org/html/rfc6241>>.
- [information-model]
Burbridge, T., Eardley, P., Bagnulo, M., and J.
Schoenwaelder, "Information Model for Large-Scale
Measurement Platforms (LMAP)", <<http://tools.ietf.org/html/draft-burbridge-lmap-information-model>>.
- [lmap-ipfix]
"An LMAP application for IPFIX",
<<http://tools.ietf.org/html/draft-bagnulo-lmap-ipfix>>.
- [lmap-netconf]
"Considerations on using NETCONF with LMAP Measurement
Agents",
<<http://tools.ietf.org/html/draft-schoenw-lmap-netconf>>.
- [lmap-yang]
"A YANG Data Model for LMAP Measurement Agents",
<<http://tools.ietf.org/html/draft-schoenw-lmap-yang>>.
- [registry]
"A registry for commonly used metrics. Independent
registries", <<http://tools.ietf.org/html/draft-bagnulo-ippm-new-registry-independent>>.
- [schulzrinne]
"Large-Scale Measurement of Broadband Performance: Use
Cases, Architecture and Protocol Requirements", <<http://tools.ietf.org/html/draft-schulzrinne-lmap-requirements>>.
- [use-cases]
"Large-Scale Broadband Measurement Use Cases",
<<http://tools.ietf.org/html/draft-linsner-lmap-use-cases>>.
- [yang-api]
"YANG-API Protocol", <<http://tools.ietf.org/html/rfc6241>>.

Authors' Addresses

Philip Eardley
British Telecom
Adastral Park, Martlesham Heath
Ipswich
ENGLAND

Email: philip.eardley@bt.com

Trevor Burbridge
British Telecom
Adastral Park, Martlesham Heath
Ipswich
ENGLAND

Email: trevor.burbridge@bt.com

Al Morton
AT&T Labs
200 Laurel Avenue South
Middletown, NJ
USA

Email: acmorton@att.com

Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: January 12, 2014

P. Eardley
BT
A. Morton
AT&T Labs
M. Bagnulo
UC3M
T. Burbridge
BT
July 11, 2013

Terminology for Large MeAsurement Platforms (LMAP)
draft-eardley-lmap-terminology-02

Abstract

This documents defines terminology for Large Scale Measurement Platforms.

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <http://datatracker.ietf.org/drafts/current/>.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on January 12, 2014.

Copyright Notice

Copyright (c) 2013 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of

the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1. Introduction	2
2. Summary	2
3. LMAP Terminology	4
3.1. Other potentially useful terminology	6
4. Commentary and notes	6
5. Security considerations	9
6. IANA Considerations	9
7. Acknowledgments	9
8. History	9
8.1. from -00 to -01:	9
8.2. from -01 to -02	10
9. Informative References	10
Authors' Addresses	10

1. Introduction

This document, in Section 3, defines terminology for LMAP. Since 'raw' terminology is reader-unfriendly, Section 2 provides an initial idea of the terminology by explaining how LMAP works whilst using the terms. Section 4 provides some commentary on the terminology, including a comparison with that in [RFC2330].

Please note that defined terms are capitalized.

2. Summary

A Measurement Task is an act that yields a single Measurement Result. An Active Measurement Task involves (for example) a Measurement Agent injecting test packet(s) into the network destined for a Measurement Peer and measuring some performance or reliability parameter associated with the transfer. The generic version of the Measurement Task is the Measurement Method; in other words the Measurement Task is the instantiation of the Measurement Method at a specific time and place.

For example, a Measurement Method might be the injection of a UDP packet by a Measurement Agent destined for a Measurement Peer, which immediately reflects the UDP packet back to the Measurement Agent, which measures the round trip latency. The associated Measurement Task might be: the injection of a UDP packet by the Measurement Agent at 192.0.2.0 destined for the Measurement Peer at 198.51.100.0 at UTC 13:01 and 58.6 seconds on 2013-06-15, with the Measurement Peer immediately reflecting the UDP packet back to the source, which

measures the associated round trip latency (using a second timestamp associated with arrival).

A Metric is a parameter of interest that is related to the performance and reliability of the Internet. For example, "UDP latency". Typically the value of a Metric is assessed as simply the average of several Measurement Results. However a Derived Metric consists of some combination of various Measurement Results. For example, a path delay might be assessed by adding several component delays, or the bulk transport capacity might be assessed by combining several different parameters as suggested in [I-D.mathis-ippm-model-based-metrics].

How and when to perform the Measurement Task and report the Measurement Result is defined by the Instruction, which the Controller sends to the Measurement Agent. Whilst the Instruction may define a single Measurement Task, more typically it defines a series of Measurement Tasks, all based on the same Measurement Method and carried out at regular times according to a Measurement Schedule. The Measurement Result of the former is likely to be reported immediately, whilst Measurement Results of the latter will be sent at regular time intervals, as defined by the Report Schedule. The Instruction consists of the following items (which effectively define a series of Measurement Tasks):

1. The Measurement Method: typically this is defined by a reference in a well-known registry (for example, 'how to measure UDP latency')
2. The configuration of parameters left open by the Measurement Method (for example, the addresses of the Measurement Agent and Measurement Peer)
3. The Measurement Schedule (for example, start at 0400 UTC, repeat every 500 ms, end at 0403 UTC)
4. Any environmental constraints (for example, do not perform the Measurement Task if there is cross-traffic)
5. How and when to send a report:
 - a. The definition of the Report. Typically the Report includes every single Measurement Result (since the last Report), but it may instead be a statistic (such as their average). Typically the Report also includes other relevant information, for example an 'echo' of the Measurement Method, configuration parameters and schedule.

- b. The configuration of parameters associated with the Report (for example, the address of the Collector to which the Report is sent)
- c. The Report Schedule (for example, send once a day at 01:00 hours)

The Control Protocol and Report Protocol define the delivery of the Instruction and the Report (respectively); they consist of a Data Model (the semantics and structure of the information, in a particular data modeling language such as a JSON schema language or YANG) and a transport protocol (such as HTTP or NETCONF).

3. LMAP Terminology

Active Measurement Method (Task): A type of Measurement Method (Task) that involves a Measurement Agent and a Measurement Peer (or possibly Peers), where either the Measurement Agent or the Measurement Peer injects test packet(s) into the network destined for the other, and which involves one of them measuring some performance or reliability parameter associated with the transfer of the packet(s).

Bootstrap Protocol: A protocol that initialises a Measurement Agent with the information necessary to talk to a Controller.

Collector: A function that receives a Report from a Measurement Agent. Colloquially, a Collector is a physical device that performs this function.

Controller: A function that provides a Measurement Agent with Instruction(s). Colloquially, a Controller is a physical device that performs this function.

Control Protocol: The protocol delivering Instruction(s) from a Controller to a Measurement Agent.

Data Model: The implementation of an Information Model in a particular data modelling language.

Derived Metric: A Metric that is a combination of other Metrics, and/or a combination of the same Metric measured over different parts of the network, or at different times.

Information Model: The protocol-neutral definition of the semantics of either the Instruction or the Report.

Instruction: The description of Measurement Tasks to perform and the details of the Report to send. The Instruction is sent by a Controller to a Measurement Agent.

Measurement Agent (MA): The function that receives Instructions from a Controller, performs Measurement Tasks (perhaps in concert with a Measurement Peer) and reports Measurement Results to a Collector. Colloquially, a Measurement Agent is a physical device that performs this function.

Measurement Method: The process for assessing the value of a Metric; the process of measuring some performance or reliability parameter; the generalisation of a Measurement Task.

Measurement Peer: The function that receives control messages and test packets from a Measurement Agent and may reply to the Measurement Agent as defined by the Measurement Method.

Measurement Result: The output of a single Measurement Task (the value obtained for the parameter of interest, or Metric).

Measurement Schedule: the schedule for performing a series of Measurement Tasks.

Measurement Task: The act that yields a single Measurement Result; the act consisting of the (single) operation of the Measurement Method at a particular time and with all its parameters set to specific values.

Metric: The quantity related to the performance and reliability of the Internet that we'd like to know the value of, and that is carefully specified.

Passive Measurement Method (Task): A Measurement Method (Task) in which a Measurement Agent observes existing traffic at a specific measurement point, but does not inject test packet(s).

Report: The Measurement Results and other associated information (as defined by the Instruction); a specific instance of the Data Model. The Report is sent by a Measurement Agent to a Collector.

Report Protocol: The protocol delivering Report(s) from a Measurement Agent to a Collector.

Report Schedule: the schedule for sending a series of Reports to a Collector.

3.1. Other potentially useful terminology

The following terms have also been suggested and will be included above, assuming they prove useful during the early stages of the LMAP work.

Cycle-ID: A tag that is sent by the Controller in an Instruction and echoed by the MA in its Report; Measurement Results with the same Cycle-ID are expected to be comparable.

Environmental Constraint: A parameter that is measured as part of the Measurement Task, its value determining whether the rest of the Measurement Task proceeds.

Group-ID: An identifier of a group of MAs.

Measurement Parameter: A parameter whose value is left open by the Measurement Method.

Measurement Suppression: a type of Instruction that stops (suppresses) Measurement Tasks.

Report Channel: a specific Report Schedule and Collector

4. Commentary and notes

To avoid confusion the word 'Measurement' is only used as an adjective.

It is worth explaining how the terms defined here compare with those in [RFC2330], "Framework for IP Performance Metrics". The definition of Metric is taken from RFC2330. The definition of Measurement Method is (we believe) equivalent in RFC2330's terms to a measurement methodology for a singleton metric. A set of Measurement Tasks defined by a Measurement Schedule relates to RFC2330's concept of a sample metric.

If a Measurement Method is used multiple times under identical or similar conditions, it should result in a consistent value for the Metric.

A Measurement Method may be a more specific version of another Measurement Method. For example, [I-D.bagnulo-ippm-new-registry-independent] defines UDP latency as a round trip delay [RFC2681] with the packet type set to UDP.

A registry, as proposed in [I-D.bagnulo-ippm-new-registry-independent], would be a registry of

Measurement Methods and their associated Metrics. A Passive Measurement Method (Task) involves only a Measurement Agent; for example, it measures the mix of applications. An Active Measurement Method (Task) also involves a Measurement Peer. It is possible that some Active Measurement Methods (Tasks) involve additional Measurement Agent(s) or Measurement Peer(s); for example, one way to measure 'latency under load' may be to send test traffic between a Measurement Agent and Measurement Peer whilst a second Measurement Peer generates the load (cross-traffic).

The consensus is that the LMAP working group should assume that a Measurement Agent receives Instruction from only a single Controller at any point in time (however it may Report to more than one Collector).

By definition a Measurement Peer does not interact with a Controller or Collector. A Measurement Peer will typically respond to the test packet(s) from the Measurement Agent. For example, it may echo a UDP packet, or measure the amount of loss of the test packets and then send the Measurement Results to the Measurement Agent.

The Measurement Agent is implemented either in specialised hardware or as code on general purpose devices like a PC, tablet or smartphone. Note that a Measurement Peer may not have specific LMAP or IPPM functionality. For example, to assess DNS response time a Measurement Agent sends DNS requests to a standard DNS server.

A Controller can send an Instruction for immediate action, containing a one-off Measurement Task. This is in addition to the more typical scenario of a series of Measurement Tasks carried out on a regular schedule, with the Measurement Results reported periodically.

It may be sensible for an Instruction to be able to refer to more than one Measurement Method. This is for further study.

[RFC3444] discusses the difference between an Information Model and Data Model. An Informational Model "model[s] managed objects at a conceptual level, independent of any specific implementations or protocols used to transport the data ... it defines relationships between managed objects". A Data Model is "defined at a lower level of abstraction and includes many details ... and include[s] protocol-specific constructs." Multiple Data Models can be derived from a single Information Model, since a conceptual/abstract model can be implemented in different ways. An Informational Model is for designers and operators, whilst a Data Model is for implementors.

An Information Model can be divided into different parts and sub-parts, to allow the values in each part and sub-part to be updated

independently. For example, one part could contain the Instruction Information and another the Reporting Information. The Instruction could contain sub-parts for configuring Measurement Tasks and for setting Measurement Schedules (which may be updated at different times and frequencies). This is discussed in [information-model]

The Control Protocol defines the Data Model and so effectively defines the Instruction. The Instruction includes: the Measurement Method; values for the parameters that the Measurement Method leaves open (configuration); when to perform the Measurement Tasks (the Measurement Schedule); any environmental conditions (such as "don't perform the Measurement Task if there is end user traffic present"); the Report Protocol, which includes its Data Model; when to send a Report (the Report Schedule); where to send the Report (the address of the Collector) and values for any other parameters that the Report Protocol leaves open (configuration). This is for discussion.

Typically the Report includes every single Measurement Result, but it may instead be a statistic (such as their average). The latter may be useful when the bandwidth between the Measurement Agent and Collector is severely constrained and/or the full set of Measurement Results provides little extra information.

The Report includes: the Measurement Results (or statistic based on them); the details of the Measurement Tasks (essentially a copy of much of the Instruction, for example the Measurement Method, the configuration parameters and the time at which each Measurement Result was obtained); and other relevant information known by the Measurement Agent (such as the line's speed, the version of the Measurement Agent, and the amount of cross-traffic during the measurement). Again this is very much for discussion.

A proposal for a Control Protocol based on HTTP is currently under development. There are already internet drafts describing a Control Protocol based on NETCONF and a Report Protocol based on IPFIX.

The job of a Bootstrap Protocol is to provide an automated way to associate a Measurement Agent to its Controller, including authentication credentials. Similarly, there should be a way to pull the plug on rogue Measurement Agents. The current consensus on the LMAP mailing list is that the working group should define the bootstrap process but not a protocol. The reason is that it could be done in many different ways, depending on the device and the measurement system, for instance: loaded at manufacture, updated locally via USB port, or orchestrated via a protocol (which may be defined by organisations other than the IETF, for example, the Broadband Forum).

The purpose of the Cycle-ID is to allow the data analysis tools to identify easily Measurement Results that are expected to be comparable, typically because the associated Measurement Tasks all operate the same Measurement Method with the same values for its parameters. This set of Measurement Tasks could be termed the Measurement Cycle.

An example of an Environmental Constraint is "no end-user traffic". The Measurement Agent could measure the amount of end-user traffic over the previous 10 seconds; if there is none then it uploads a file to the Measurement Peer, whilst if the end-user is active then it defers the upload.

The Report Channel contains the details of one collector (including location and security information such as the certificate), and the timing for the report (when to report the results). Each Report Channel is also given a local short name by which it can be referenced from a Measurement Schedule.

The Group-ID identifies a group of interest to which a MA belongs. For example the group could represent an ISP, broadband product, technology, market classification, geographic region, or a combination of multiple such characteristics. A MA can remain anonymous by including its Group-ID (and not its own identifier) in the Reports it sends.

Measurement Suppression is used to over-ride the Measurement Schedule. A Controller uses Measurement Suppression to stop a MA making measurements for a defined or indefinite period. For discussion of Measurement Suppression, see [information-model]

5. Security considerations

There are no security considerations in this memo.

6. IANA Considerations

There are no IANA considerations in this memo.

7. Acknowledgments

We thank participants on the LMAP mailing list for their input, especially Juergen Schoenwaelder for his detailed review.

8. History

8.1. from -00 to -01:

'Complete Measurement Agent' replaced by 'Measurement Agent', and
'Remote Measurement Agent' replaced by 'Measurement Peer'.

Bootstrap protocol added

Section 3.1 added, with terms Cycle-ID, Measurement Parameter and
Environmental Constraints

Adjustments to terms for: Active Measurement Method (Task), Control
Protocol, Information Model, Instruction, Report Protocol.

8.2. from -01 to -02

Added to Section 3.1 the terms Group-ID, Measurement Suppression and
Report Channel, as these are used in [information-model]

Other minor clarifications.

9. Informative References

[I-D.bagnulo-ippm-new-registry-independent]

Bagnulo, M., Burbridge, T., Crawford, S., Eardley, P., and
A. Morton, "A registry for commonly used metrics.
Independent registries", draft-bagnulo-ippm-new-registry-
independent-00 (work in progress), January 2013.

[RFC2330] Paxson, V., Almes, G., Mahdavi, J., and M. Mathis,
"Framework for IP Performance Metrics", RFC 2330, May
1998.

[I-D.mathis-ippm-model-based-metrics]

Mathis, M. and A. Morton, "Model Based Internet
Performance Metrics", draft-mathis-ippm-model-based-
metrics-01 (work in progress), February 2013.

[RFC2681] Almes, G., Kalidindi, S., and M. Zekauskas, "A Round-trip
Delay Metric for IPPM", RFC 2681, September 1999.

[information-model]

Burbridge, T., Eardley, P., Bagnulo, M., and J.
Schoenwaelder, "Information Model for Large-Scale
Measurement Platforms (LMAP)", , <[http://tools.ietf.org/
html/draft-burbridge-lmap-information-model](http://tools.ietf.org/html/draft-burbridge-lmap-information-model)>.

Authors' Addresses

Philip Eardley
British Telecom
Adastral Park, Martlesham Heath
Ipswich
ENGLAND

Email: philip.eardley@bt.com

Al Morton
AT&T Labs
200 Laurel Avenue South
Middletown, NJ
USA

Email: acmorton@att.com

Marcelo Bagnulo
Universidad Carlos III de Madrid
Av. Universidad 30
Leganes, Madrid 28911
SPAIN

Phone: 34 91 6249500
Email: marcelo@it.uc3m.es
URI: <http://www.it.uc3m.es>

Trevor Burbridge
British Telecom
Adastral Park, Martlesham Heath
Ipswich
ENGLAND

Email: trevor.burbridge@bt.com

LMAP WG
Internet-Draft
Intended status: Informational
Expires: January 16, 2014

D. Goergen
R. State
University of Luxembourg
V. Gurbani
Bell Labs, Alcatel-Lucent
July 15, 2013

Aggregating large-scale measurements for Application Layer Traffic
Optimization (ALTO) Protocol
draft-goergen-lmap-fcc-00

Abstract

Analyzing and aggregating large-scale broadband measurements is essential to study trends and derive network analytics. These trends and analyses could be made available through well defined protocols such as the Application Layer Traffic Optimization (ALTO) protocol. However, ALTO requires network information to be distilled and abstracted in form of a network map and a cost map. We describe our methodology for analyzing the United States Federal Communication Commission's (FCC) Measuring Broadband America (MBA) dataset to derive required topology and cost maps suitable for consumption by an ALTO server.

Requirements Language

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].

Status of this Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <http://datatracker.ietf.org/drafts/current/>.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on January 16, 2014.

Copyright Notice

Copyright (c) 2013 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1. Introduction	4
2. Challenges in data analysis	5
3. Geo-locating the units	6
4. Conclusions and future work	9
5. IANA Considerations	10
6. Security Considerations	11
7. References	12
7.1. Normative References	12
7.2. Informative References	12
Authors' Addresses	13

1. Introduction

Measuring broadband performance is increasingly important as communications continue to move towards the Internet. Internet service providers (ISP), national agencies and other entities gather broadband data and may provide some, or all, of the dataset to the public for analysis. As [I-D.seedorf-lmap-alto] notes, there are two extremes prevalent for presenting large-scale data. One is in the form of charts, figures, or summarized reports amenable for easy and quick consumption. The other extreme includes releasing raw data in the form of large files containing tables formatted as values separated by a delimiter. While the former is indispensable to acquire a summary view of the dataset, it does not suffice for additional analysis beyond what is presented. Conversely, the problem with the latter option (raw files) is that the unsuspecting user perusing them is lost in the deluge of data.

[I-D.seedorf-lmap-alto] offers the argument that a reasonable medium between the two extremes may be a protocol that allows a constrained set of user-driven ad-hoc queries on the dataset. It further offers that the Application Layer Traffic Optimization (ALTO) protocol [I-D.ietf-alto-protocol] be the protocol of choice that allows such reasoning on the dataset. A necessary prerequisite for using ALTO is abstracting the network information into a form that is suitable for consumption by the protocol. The implication of using ALTO is that data from any large-scale measurement effort must first be distilled in two maps: a topology map and a cost map. Further analysis and ad-hoc queries can be subsequently performed on the normalized dataset.

In the United States, the Federal Communication Commission (FCC) has embarked on a nationwide performance study of residential wireline broadband service [fcc]. Our aim is to use the raw datasets from this study for analysis and to create a topology map and a cost map from this dataset. ALTO queries aimed at these maps will enable users and interested parties to fulfill the use cases listed in Section 2 of [I-D.seedorf-lmap-alto].

2. Challenges in data analysis

The FCC Measuring Broadband America (MBA) study consisted of 7,782 volunteers spread across the United States with adequate geographic diversity. Volunteers opted in for the study, however, each of the volunteers remained anonymous. An opaque integral number (`unit_id`) represented a subscriber in the raw dataset. This `unit_id` remains constant during the duration of the study in the dataset and uniquely identifies a volunteer subscriber, even if the subscriber switches the ISP. More detail about the methodology used is described in [fcc].

The dataset consisted of 12 tables, each table corresponding to the data drawn from a certain performance test. For the analysis we present in this document we focus on the "curr_dns" table, which contains the time taken for the ISP's recursive DNS resolver to return a DNS A RR for a popular website domain name. This test was ran approximately every hour in a 24-hour period, and produced about 75-78 million records per month. This resulted in a typical file size in the range of 6-7 GBytes per month. We note that the "curr_dns" table is one of the smaller tables in the dataset.

The first challenge, therefore, was to arrive at computing resources comparable in scale with respect to the dataset consisting of millions of records spread across gigabyte-sized files. To analyze the volume of data we used a canonical Map-Reduce computational paradigm on a Hadoop cluster (more details on the methodology are outlined in Section 3).

A second, more pressing challenge, was to identify the geographic location of the `unit_ids` generating the data. In order to derive a topological map and impose costs on the links, it is important to know the physical locations of the `unit_ids` that contributed the measurements. However, in the MBA dataset, the population is anonymized and the individual subscriber reporting the measurement data is simply referred to by an opaque integral number. Therefore, an important task was to use the information in the public tables to reveal a coarse location of the subscriber.

We outline the methodology we used to do so in the next section. We stress that this methodology does not identify the specific location of a subscriber, who still remains anonymous. Instead, it simply locates the subscriber in a larger metropolitan region. This level of granularity suffices for our work.

3. Geo-locating the units

To geo-locate the units, we simply note that broadband subscriber devices are likely to be configured using DHCP by their ISP. Besides imparting an IP address to the subscriber device, DHCP also populates the DNS name servers the subscriber devices uses for DNS queries. In most installations, these DNS name servers are located in close physical proximity of the subscriber device. The FCC technical appendix states that the DNS resolution tests were targeted directly at the ISP's recursive resolvers to circumvent caching and users configuring the subscriber device to circumvent the ISP's DNS resolvers. Therefore, a reasonable approximation of a subscribers geo-location could be the geographic location of the DNS name server serving the subscriber. We use this very heuristic to geo-locate a subscriber.

Thus our first, and very simple filter consisted of obtaining a mapping from a unit_id (representing a subscriber) to one or more DNS name servers that the unit_id is sending DNS requests to. It turned out that while this was a necessary condition for advancing, it was not a sufficient one. The raw data would need to be further processed to reduce inconsistencies and remove outliers. A number of interesting artifacts were uncovered during further processing of the data. These artifacts informed the selection of the unit_ids for further analysis.

The artifacts are documented below.

- o A handful of unit_ids were geo-located in areas outside the contiguous United States, such as Ukraine, Poland or the United Kingdom. We theorize that the subscribers corresponding to the unit_ids geo-located outside the contiguous United States had simply configured their devices to use alternate DNS servers, probably located outside the United States. We removed these records before conducting our analysis.
- o We also observed a reasonable number of non-ISP DNS resolvers, especially Google's 8.8.8.8 and 8.8.4.4 and OpenDNS 208.67.222.222 and 208.67.220.220. These 4 public DNS servers are geo-located in California. We removed these records to ensure that the specific location that these resolvers represented was not oversampled.
- o We noticed that a large number of unit_ids were being geo-located in Potwin, Kansas. Intrigued as to why there appeared to be a large population of Internet users being located in a small rural community in Kansas, we investigated further. It appears that Potwin, Kansas is the geographical center of the United States and a number of ISPs have chosen to establish data centers in or

around the Potwin area. These ISPs generally locate their primary or secondary DNS name servers in Potwin-area data centers, thus accounting for the popularity of Potwin as an Internet destination. We continue to further investigate on minimizing the impact of such natural aggregation points that, if not accounted for, will skew our results in an unwarranted direction.

- o We observed some `unit_ids` changing ISPs during the observation period. This is a normal occurrence and to the extent that the `unit_id` is geo-located in the same geographical area after the change in ISP, we do not exclude such `unit_ids` from further analysis.

Subsequent filters extracted the stable `unit_ids` from our dataset. In order to determine which `unit_id` are stable, i.e., remain constant with respect to their geographic location over the observation period from January to December 2012, we extracted for each `unit_id` the IP address of each DNS name server it consulted. This is obtained by applying the map reduce paradigm on the DNS dataset. We extracted for each `unit_id` the triggered DNS servers and obtained the individual DNS servers accessed by a `unit_id`. This was repeated for each month of the observation period. The resulting sets were cleaned up of private IP addresses and other artifacts discussed above. The cleaned set consisted of about 8000 distinct `unit_id`.

In order to determine the stability of each `unit_id` we proceeded to sum up the occurrences of IP addresses over the whole observation period separated in monthly files. If the IP address of a DNS server occurred 12 times this meant that the `unit_id` always accessed the same DNS server and therefore remained stable over the observation period. The obtained stable `unit_ids`, around 1500, will be used for further analysis. Assuming a 99% confidence level and ± 3 point margin of error, we will require a sample of 1494 `unit_ids`. With our stable `unit_id` set of 1500 `unit_ids`, we are now positioned to perform further analysis on the dataset to create the full topology and cost maps.

Table 1 presents a sample of the geographic location data that we have uncovered for `unit_ids`. A complete list of identified units superimposed on the geographical map of the United States is available at <http://cdb.io/13UOHgD>.

Unit ID	City, State	Latitude/Longitude
872	Morganville, NJ	40.35950089,-74.26280212
885	Madison, WI	43.07310104,-89.40119934
898	Foley, AL	30.40660095,-87.68360138
7969	Manteca, CA	37.79740143,-121.2160034
8024	Quincy, MA	42.25289917,-71.00229645

Sample unit identification tuples

Table 1

4. Conclusions and future work

Identification of the geographic location of the unit_ids generating the performance data is essential in order to continue the work. We have presented a methodology and some early results in identifying a geographic location. This location, although coarse, suffices for our future work that will consist of further data mining and analysis to create appropriate ALTO network and cost maps.

5. IANA Considerations

This document does not contain any IANA considerations

6. Security Considerations

There are no security artifacts that have been invalidated due to our analysis. All of our analysis was performed on publicly available data. However, we do note that some privacy may have been lost based on our analysis. In the raw dataset, the unit identifiers are opaque strings with no immediate correlation with a geographic location. After our analysis, while the unit identifiers still remain opaque, they are nonetheless correlated to a specific, though coarse, geographic location.

7. References

7.1. Normative References

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

7.2. Informative References

- [I-D.ietf-alto-protocol]
Alimi, R., Penno, R., and Y. Yang, "ALTO Protocol",
draft-ietf-alto-protocol-17 (work in progress), July 2013.
- [I-D.seedorf-lmap-alto]
Seedorf, J., Gurbani, V., and E. Marocco, "ALTO for
Querying LMAP Results", draft-seedorf-lmap-alto-01 (work
in progress), July 2013.
- [fcc] United States Federal Communications Commission,
"Measuring Broadband America", Accessed July 12,
2013, <http://www.fcc.gov/measuring-broadband-america>.

Authors' Addresses

David Goergen
University of Luxembourg
Email: david.goergen@uni.lu

Radu State
University of Luxembourg
Email: radu.state@uni.lu

Vijay K. Gurbani
Bell Labs, Alcatel-Lucent
Email: vijay.gurbani@alcatel-lucent.com

INTERNET-DRAFT
Intended Status: Standards Track
Expires: December 30, 2013

R. Huang
Huawei
June 28, 2013

Use Case for Large Scale Measurements Used in Data Collection of
Network Management Systems
draft-huang-lmap-data-collection-use-case-00

Abstract

This document augments the use cases of large scale measurement of broadband performance (LMAP). It discusses measurements for a common platform which works for different usages including troubleshooting, performance understanding, quality evaluation and network adjusting.

Status of this Memo

This Internet-Draft is submitted to IETF in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

The list of current Internet-Drafts can be accessed at
<http://www.ietf.org/lid-abstracts.html>

The list of Internet-Draft Shadow Directories can be accessed at
<http://www.ietf.org/shadow.html>

Copyright and License Notice

Copyright (c) 2013 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents

carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1 Introduction 3

2 Terminology 3

3 Use Case for Data Collection of ISP Network Management Systems 3

3 Security Considerations 6

4 IANA Considerations 6

5 References 6

5.1 Normative References 6

5.2 Informative References 6

Authors' Addresses 6

1 Introduction

To support new services in network and provide better service quality, ISPs have to reconstruct or update their network constantly, which makes the network more and more complex and irregular. ISPs eager to have a comprehensive network management system to make the complex network in control to achieve real-time network performance acquisition, rapidly and accurately diagnose network fault with low cost, and SLAs of users satisfaction. Traditional network management systems consist of many different measurement panels, some of which are measurements in isolated network probes to report serious faults and some of which are measurements initiated by end users or dedicated network devices to locate the problem happening in their service paths or to calculate their performance of subscribed services. However, running network is always invisible to either operators or users. Through these measurements, traditional network management system could only interpret some scattered and fragmentary periods and paths while can't draw the whole running network picture for operators. It is also hard for them to deal with some difficult and complicated faults, such as sudden transient performance decrease which won't trigger any alarms.

This document introduces a new use case supplementing the use cases described in [LMAP-USECASE]. It discusses measurements for a common platform which works for different usages including troubleshooting, performance understanding, quality evaluation and network adjusting.

2 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].

3 Use Case for Data Collection of ISP Network Management Systems

Network data collection is essential for effective network management. Precise data collection and reasonable data analysis and processing could provide necessary information for network performance management, fault management and quality of service monitoring. Lmap could be used in network data collection of network management system as a universal method to sufficiently manage data collecting activities. Rather, lmap could be implemented in an intelligent network management platform to obtain accurate sample data for creating visualized simulation network which simulates the real one. By doing this, the intelligent visualized platform could draw a whole running network picture for ISPs and provide the better capability for troubleshooting, monitoring performance, and even

network planning instead of doing all kinds of measurements in the real networks.

This intelligent network management platform using lmap constantly collects traffic information and device states, e.g. queue information, from the network, analyses them, and creates some sampled snapshots of the real network. By some simulation algorithms, these sampled snapshots could form a simulation network which simulates the real network accurately if sufficient information is retrieved. Usages, such as troubleshooting and quality monitoring, could run on the simulation network instead of using extra probes and measurement in the real network. For example, to learn what happened at a certain past time, ISPs could simply use the sampled snapshots of that time to infer the whole network curve of that past specific period. By further investigating, this intelligent platform could easily provide the ability to find the failure reason or discover the pattern of some complicated faults, e.g., one specific router has overflowed queues, which causes network congests.

In this case, MAs are network devices constituting the whole network. Only passive measurements are needed since MAs just monitor the device states, network status and traffic information. No extra payload will be added to the existing network. Network-specific parameters are enough for this usage. Service-specific parameters will only be required in the simulation network, which is not in the scope of LMAP. This use case must consider and alleviate the performance issues caused by sample frequency and heavy measurement results reporting. As we know that the higher the sample frequency is, the closer the sampled network curve is to real network curve. But too frequent sample times will increase the burden of MAs and exhaust the resources of network devices. To solve this problem, some mechanisms, for example, using lower sample frequency when data vary gently while increasing sample frequency when network data change dramatically, should be considered to adjust data collection frequency. Another concern is the MA implementation in the network. The "ideal" situation is MA in each device (e.g. routers, switches) of the network ISP wishes to manage. However it is not feasible because we could envision heavy measurement report traffic disrupting the normal network traffic in large scale case. So only those network nodes arranged in a crisscross pattern and those important network devices to ISPs should be considered.

Due to the high requirements of precise data collection and large scale environment, traditional protocols like SNMP are insufficient to do this kind of work in such huge and continuously expanding networks because of their constraints, e.g., producing plenty of management data which may causes serious traffic congestions, and incoordination among different network devices from different

vendors.

Normally, usages in other use cases described in [LMAP-USECASE] are designed with some corresponding specific measurements. For example, measurements for identifying network problems, or measurements for evaluation the quality experienced by end users, etc. While different from other use cases, lmap used in this case is not dedicated for certain services, usages or end users. Instead, it is used to create a common and universal network management platform for all kinds of usages required by ISPs, including troubleshooting, performance evaluation, and other functions.

The characteristics of large scale measurements emerging from this use case:

1. Passive measurements are needed while active ones aren't.
2. Network device states are also required as well as specific network performance parameters. Metrics of upper layer 3 are not.
3. The data collection frequency of passive measurements could be adjusted adaptively.
4. Results from the tests should not be averaged.
5. Regular scheduled tests are necessary.

3 Security Considerations

TBD

4 IANA Considerations

TBD

5 References

5.1 Normative References

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

[LMAP-USECASE] Linsner, M., "Large-Scale Broadband Measurement Use Cases", draft-linsner-lmap-user-cases-02, February, 2013

5.2 Informative References

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

Authors' Addresses

Rachel Huang
Huawei Technologies Co., Ltd.
101 Software, Yuhua District
Nanjing, Jiangsu, 210012 P.R.China

EMail: rachel.huang@huawei.com

Network Working Group
Internet-Draft
Intended status: Informational
Expires: April 10, 2015

M. Bagnulo
UC3M
T. Burbridge
BT
S. Crawford
SamKnows
P. Eardley
BT
A. Morton
AT&T Labs
October 7, 2014

A Reference Path and Measurement Points for Large-Scale Measurement of
Broadband Performance
draft-ietf-ippm-lmap-path-07

Abstract

This document defines a reference path for Large-scale Measurement of Broadband Access Performance (LMAP) and measurement points for commonly used performance metrics. Other similar measurement projects may also be able to use the extensions described here for measurement point location. The purpose is to create an efficient way to describe the location of the measurement point(s) used to conduct a particular measurement.

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <http://datatracker.ietf.org/drafts/current/>.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on April 10, 2015.

Copyright Notice

Copyright (c) 2014 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1. Introduction	2
1.1. Requirements Language	3
2. Purpose and Scope	3
3. Terms and Definitions	4
3.1. Reference Path	4
3.2. Subscriber	4
3.3. Dedicated Component (Links or Nodes)	5
3.4. Shared Component (Links or Nodes)	5
3.5. Resource Transition Point	5
3.6. Service Demarcation Point	5
3.7. Managed and Un-Managed Sub-paths	5
4. Reference Path	6
5. Measurement Points	7
6. Translation Between Reference Path and Various Technologies	11
7. Example Resource Transition	12
8. Security considerations	13
9. IANA Considerations	14
10. Acknowledgements	14
11. References	14
11.1. Normative References	14
11.2. Informative References	14
Authors' Addresses	15

1. Introduction

This document defines a reference path for Large-scale Measurement of Broadband Access Performance (LMAP) or similar measurement projects. The series of IP Performance Metrics (IPPM) RFCs have developed terms that are generally useful for path description (section 5 of [RFC2330]). There are a limited number of additional terms needing definition here, and they will be defined in this memo.

The reference path (See section 3.1 and Figure 1 of [Y.1541], including the accompanying discussion) is usually needed when attempting to communicate precisely about the components that comprise the path, often in terms of their number (hops) and geographic location. This memo takes the path definition further, by establishing a set of measurement points along the path and ascribing a unique designation to each point. This topic has been previously developed in section 5.1 of [RFC3432], and as part of the updated framework for composition and aggregation, section 4 of [RFC5835]. Section 4.1 of [RFC5835] defines the term "measurement point".

Measurement points and the paths they inhabit are often described in general terms, like "end-to-end", "user-to-user", or "access". These terms alone are insufficient for scientific method: What is an end? Where is a user located? Is the home network included?

As an illustrative example, consider a measurement agent in an LMAP system. When it reports its measurement results, rather than detailing its IP address and that of its measurement peer, it may prefer to describe the measured path segment abstractly (perhaps for privacy reasons). For instance "from a measurement agent at a home gateway to a measurement peer at a DSLAM". This memo provides the definition for such abstract 'measurement points' and therefore the portion of 'reference path' between them.

The motivation for this memo is to provide an unambiguous framework to describe measurement coverage, or scope of the reference path. This is an essential part of the meta-data to describe measurement results. Measurements conducted over different path scopes are not a valid basis for performance comparisons. We note that additional measurement context information may be necessary to support a valid comparison of results.

1.1. Requirements Language

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. Purpose and Scope

The scope of this memo is to define a reference path for LMAP activities with sufficient level of detail to determine the location of different measurement points along a path without ambiguity. These conventions are likely to be useful in other measurement projects as well, and in describing the applicable measurement scope for some metrics.

The connection between the reference path and specific network technologies (with differing underlying architectures) is within the scope of this method, and examples are provided. Both wired and wireless technologies are in-scope.

The purpose is to create an efficient way to describe the location of the measurement point(s) used to conduct a particular measurement so that the measurement result will adequately described in terms of scope or coverage. This should serve many measurement uses, including:

diagnostic: where the same metric would be measured on different sub-paths bounded by measurement points (see Section 4.10 of[RFC5835]), for example to isolate the sub-path contributing the majority of impairment levels observed on a path.

comparison: where the same metric may be measured on equivalent portions of different network infrastructures, for example to compare the performance of wired and wireless home network technologies.

3. Terms and Definitions

This section defines key terms and concepts for the purposes of this memo.

3.1. Reference Path

A reference path is a serial combination of hosts, routers, switches, links, radios, and processing elements that comprise all the network elements traversed by each packet in a flow between the source and destination hosts. A reference path also indicates the various boundaries present, such as administrative boundaries. A reference path is intended to be equally applicable to all IP and link-layer networking technologies. Therefore, the components are generically defined but their functions should have a clear counterpart or be obviously omitted in any network architecture.

3.2. Subscriber

An entity (associated with one or more users) that is engaged in a subscription with a service provider. The subscriber is allowed to subscribe and un-subscribe to services, and to register a user or a list of users authorized to enjoy these services. [Q1741] Both the subscriber and service provider are allowed to set the limits relative to the use that associated users make of subscribed services.

3.3. Dedicated Component (Links or Nodes)

All resources of a Dedicated Component (typically a link or node on the Reference Path) are allocated to serving the traffic of an individual Subscriber. Resources include transmission time-slots, queue space, processing for encapsulation and address/port translation, and others. A Dedicated Component can affect the performance of the Reference Path, or the performance of any sub-path where the component is involved.

3.4. Shared Component (Links or Nodes)

A component on the Reference Path is designated a Shared Component when the traffic associated with multiple Subscribers is served by common resources.

3.5. Resource Transition Point

A point between Dedicated and Shared Components on a Reference Path that may be a point of significance, and is identified as a transition between two types of resources.

3.6. Service Demarcation Point

This is the point where service managed by the service provider begins (or ends), and varies by technology. For example, this point is usually defined as the Ethernet interface on a residential gateway or modem where the scope of a packet transfer service begins and ends. In the case of a WiFi Service, this would be an Air Interface within the intended service boundary (e.g., walls of the coffee shop). The Demarcation Point may be within an integrated endpoint using an Air Interface (e.g., Long-Term Evolution User Equipment, LTE UE). Ownership does not necessarily affect the demarcation point; a Subscriber may own all equipment on their premises, but it is likely that the service provider will certify such equipment for connection to their network, or a third-party will certify standards compliance.

3.7. Managed and Un-Managed Sub-paths

Service providers are responsible for the portion of the path they manage. However, most paths involve a sub-path which is beyond the management of the subscriber's service provider. This means that private networks, wireless networks using unlicensed frequencies, and the networks of other service are designated as Un-managed sub-paths. The Service Demarcation Point always divides Managed and Un-managed sub-paths.

4. Reference Path

This section defines a reference path for Internet communication.

```
Subsc. -- Private -- Private -- Service-- Intra IP -- GRA -- Transit ...
device   Net #1     Net #2   Demarc.   Access   GW     GRA GW
```

```
... Transit -- GRA -- Service -- Private -- Private -- Destination
   GRA GW    GW     Demarc.   Net #n    Net #n+1  Host
```

GRA = Globally Routable Address, GW = Gateway

The following are descriptions of reference path components that may not be clear from their name alone.

- o Subsc. (Subscriber) device - This is a host that normally originates and terminates communications conducted over the IP packet transfer service.
- o Private Net #x - This is a network of devices owned and operated by the Internet Service Subscriber. In some configurations, one or more private networks and the device that provides the Service Demarcation point are collapsed in a single device (and ownership may shift to the service provider), and this should be noted as part of the path description.
- o Intra IP Access - This is the first point in the access architecture beyond the Service Demarc. where a globally routable IP address is exposed and used for routing. In architectures that use tunneling, this point may be equivalent to the Globally Routable Address Gateway (GRA GW). This point could also collapse to the device providing the Service Demarc., in principle. Only one Intra IP Access point is shown, but they can be identified in any access network.
- o GRA GW - the point of interconnection between a Service Provider's administrative domain and the rest of the Internet, where routing will depend on the GRAs in the IP header.
- o Transit GRA GW - If one or more networks intervene between the Service Provider's access networks of the Subscriber and of the Destination Host, then such networks are designated "transit" and are bounded by two Transit GRA GW.

Use of multiple IP address families in the measurement path must be noted, as the conversions between IPv4 and IPv6 certainly influence the visibility of a GRA for each family.

In the case that a private address space is used throughout an access architecture, then the Intra IP Access points must use the same address space as the Service Demarcation point, and the Intra IP Access points must be selected such that a test between these points produces a useful assessment of access performance (e.g., includes both shared and dedicated access link infrastructure).

5. Measurement Points

A key aspect of measurement points, beyond the definition in section 4.1 of [RFC5835], is that the innermost IP header and higher layer information must be accessible through some means. This is essential to measure IP metrics. There may be tunnels and/or other layers which encapsulate the innermost IP header, even adding another IP header of their own.

In general, measurement points cannot always be located exactly where desired. However, the definition in [RFC5835] and the discussion in section 5.1 of [RFC3432] indicate that allowances can be made: for example, it is nearly ideal when there are deterministic errors that can be quantified between desired and actual measurement point.

The Figure below illustrates the assignment of measurement points to selected components of the reference path.

Subsc.	--	Private	--	Private	--	Service--	Intra IP	--	GRA	--	Transit ...
device		Net #1		Net #2		Demarc.	Access		GW		GRA GW
mp000						mp100	mp150		mp190		mp200
...											
...	Transit	--	GRA	--	Service	--	Private	--	Private	--	Destination
	GRA GW		GW		Demarc.		Net #n		Net #n+1		Host
	mpX90		mp890		mp800						mp900

GRA = Globally Routable Address, GW = Gateway

Figure 1

Each measurement point on a specific reference path MUST be assigned a unique number. To facilitate interpretation of the results, the measuring organisation (and whoever it shares results with) MUST have an unambiguous understanding of what path or point was measured. In order to achieve this, a set of numbering recommendations follow.

When communicating the results of measurements, the measuring organization SHOULD supply a diagram similar to Figure 1 (with the technology-specific information in examples that follow), and MUST supply it when additional measurement point numbers have been defined and used, with sufficient detail to identify measurement locations in the path.

Ideally, the consumer of measurement results would know the location of a measurement point on the reference path from the measurement point number alone, and the recommendations below provide a way to accomplish this goal. Although the initial numbering may be fully compliant with this system, network growth, consolidation, and re-arrangement, or circumstances such as ownership changes, could cause gaps in network numbers or non-monotonic measurement point number assignments along the path over time. These are examples of reasonable causes for numbering deviations which must be identified on the reference path diagram, as required above.

Whilst the numbering of a measurement point is in the context of a particular path, for simplicity the measuring organisation SHOULD use the same numbering for a device (playing the same role) on all the measurement paths through it. Similarly, whilst the measurement point numbering is in the context of a particular measuring organisation, organizations with similar technologies and architectures are encouraged to coordinate on local numbering and diagrams.

The measurement point numbering system, mpXnn, has two independent parts:

1. The X in mpXnn indicates the network number. The network with the Subscriber's device is network 0. The network of a different organization (administrative or ownership domains) SHOULD be assigned a different number. Each successive network number SHOULD be one greater than the previous network's number. Two circumstances make it necessary to designate X=9 in the Destination Host's network and X=8 for the Service Provider network at the Destination:
 - A. The number of Transit networks is unknown.
 - B. The number of Transit networks varies over time.
2. The nn in mpXnn indicates the measurement point and is locally-assigned by network X. The following conventions are suggested:

- A. 00 SHOULD be used for a measurement point at the Subscriber's device and at the Service Demarcation point or GW nearest to the Subscriber's device for Transit Networks.
- B. 90 SHOULD be used for a measurement point at the GW of a network (opposite from the Subscriber's device or Service Demarc.).
- C. In most networks, measurement point numbers SHOULD monotonically increase from the point nearest the Subscriber's device to the opposite network boundary on the path (see below).
- D. When a Destination host is part of the path, 00 SHOULD be used for a measurement point at the Destination host and at the Destination's Service Demarcation point. Measurement point numbers SHOULD monotonically increase from the point nearest the Destination's host to the opposite network boundary on the path ONLY in these networks. This directional numbering reversal allows consistent 00 designation for end hosts and Service Demarcs.
- E. 50 MAY be used for an intermediate measurement point of significance, such as a Network Address Translator (NAT).
- F. 20 MAY be used for a traffic aggregation point such as a DSLAM within a network.
- G. Any other measurement points SHOULD be assigned unused integers between 01 and 99. The assignment SHOULD be stable for at least the duration of a particular measurement study, and SHOULD avoid numbers that have been assigned to other locations within network X (unless the assignment is considered sufficiently stale). Sub-networks or domains within a network are useful locations for measurement points.

When supplying a diagram of the reference path and measurement points, the operator of the measurement system MUST indicate: the reference path, the numbers (mpXnn) of the measurement points, and the technology-specific definition of any measurement point other than X00 and X90 with sufficient detail to clearly define its location (similar to the technology-specific examples in Section 6 of this document).

If the number of intermediate networks (between the source and destination) is not known or is unstable, then this SHOULD be indicated on the diagram and results from measurement points within those networks need to be treated with caution.

Notes:

- o The terminology "on-net" and "off-net" is sometimes used when referring to the Subscriber's Internet Service Provider (ISP) measurement coverage. With respect to the reference path, tests between mp100 and mp190 are "on-net".
- o Widely deployed broadband Internet access measurements have used pass-through devices[SK] (at the subscriber's location) directly connected to the service demarcation point: this would be located at mp100.
- o The networking technology must be indicated for the measurement points used, especially the interface standard and configured speed (because the measurement connectivity itself can be a limiting factor for the results).
- o If it can be shown that a link connecting to a measurement point has reliably deterministic performance or negligible impairments, then the remote end of the connecting link is an equivalent point for some methods of measurement (although those methods should describe this possibility in detail; it is not in-scope to provide such methods here). In any case, the presence of a link and claimed equivalent measurement point must be reported.
- o Some access network architectures may have an additional traffic aggregation device between mp100 and mp150. Use of a measurement point at this location would require a local number and diagram.
- o A Carrier Grade NAT (CGN) deployed in the Service Provider's access network would be positioned between mp100 and mp190, and the egress side of the CGN may be designated mp150. mp150 is generally an intermediate measurement point in the same address space as mp190.
- o In the case that private address space is used in an access architecture, then mp100 may need to use the same address space as its "on-net" measurement point counterpart, so that a test between these points produces a useful assessment of network performance. Tests between mp000 and mp100 could use a different private address space, and when the globally-routable side of a CGN is at mp150, then the private address side of the CGN could be designated mp149 for tests with mp100.
- o Measurement points at Transit GRA GWs are numbered mpX00 and mpX90, where X is the lowest positive integer not already used in the path. The GW of the first transit network is shown, with point mp200 and the last transit network GW with mpX90.

6. Translation Between Reference Path and Various Technologies

This section and those that follow are intended to provide example mappings between particular network technologies and the reference path.

We provide an example for 3G Cellular access below.

Subscriber	--	Private	---	Service	-----	GRA	---	Transit	...
device		Net #1		Demarc.		GW		GRA GW	
mp000				mp100		mp190		mp200	

	_____UE_____		_____RAN+Core_____		_____GGSN_____	
	_____Un-managed sub-path_____		_____Managed sub-path_____			

GRA = Globally Routable Address, GW = Gateway, UE = User Equipment,
RAN = Radio Access Network, GGSN = Gateway GPRS Support Node.

We next provide an example of DSL access. Consider the case where:

- o The Customer Premises Equipment (CPE) has a NAT device that is configured with a public IP address.
- o The CPE is a home router that has also incorporated a WiFi access point and this is the only networking device in the home network, all endpoints attach directly to the CPE through the WiFi access.

We believe this is a fairly common configuration in some parts of the world and fairly simple as well.

This case would map into the defined reference measurement points as follows:

Subsc.	--	Private	--	Private	--	Service	--	Intra IP	--	GRA	--	Transit	...
device		Net #1		Net #2		Demarc.		Access		GW		GRA GW	
mp000						mp100		mp150		mp190		mp200	

	--UE--		-----CPE/NAT-----		-----		-BRAS-		-----	
					-----DSL Network---					
	_____Un-managed sub-path_____		_____Managed sub-path_____							

GRA = Globally Routable Address, GW = Gateway, BRAS = Broadband Remote Access Server

Consider next another access network case where:

- o The Customer Premises Equipment (CPE) is a NAT device that is configured with a private IP address.
- o There is a Carrier Grade NAT (CGN) located deep in the Access ISP network.
- o The CPE is a home router that has also incorporated a WiFi access point and this is the only networking device in the home network, all endpoints attach directly to the CPE through the WiFi access.

We believe this is becoming a fairly common configuration in some parts of the world.

This case would map into the defined reference measurement points as follows:

Subsc. device	-- Private Net #1	-----	Service-- Demarc.	Intra IP Access	-- GRA GW	Transit ... GRA GW
mp000			mp100	mp150	mp190	mp200
--UE--	-----CPE/NAT-----		-----	-CGN-	-----	
				--Access Network--		
	_____Un-managed sub-path_____			_Managed sub-path_		

GRA = Globally Routable Address, GW = Gateway

7. Example Resource Transition

This section gives an example of Shared and Dedicated portions with the reference path. This example shows two Resource Transition Points.

Consider the case where:

- o The CPE consists of a wired Residential GW and modem (Private Net#2) connected to a WiFi access point (Private Net#1). The Subscriber device (UE) attaches to the CPE through the WiFi access.
- o The WiFi subnetwork (Private Net#1) shares unlicensed radio channel resources with other WiFi access networks (and potentially other sources of interference), thus this is a Shared portion of the path.
- o The wired subnetwork (Private Net#2) and a portion of the Service Provider's Network are Dedicated Resources (for a single Subscriber), thus there is a Resource Transition Point between (Private Net#1) and (Private Net#2).

- o Subscriber traffic shares common resources with other subscribers upon reaching the Carrier Grade NAT (CGN), thus there is a Resource Transition Point and further network components are designated as Shared Resources.

We believe this is a fairly common configuration in parts of the world.

This case would map into the defined reference measurement points as follows:

```

Subsc. -- Private -- Private -- Access -- Intra IP -- GRA -- Transit ...
device      Net #1      Net #2      Demarc.      Access      GW      GRA GW
mp000
|--UE--|-----CPE/NAT-----|-----| -CGN- |-----|
      |   WiFi   | 1000Base-T | --Access Network--|
      |
      | -Shared--|RT|-----Dedicated-----| RT |-----Shared-----...
      |_____Un-managed sub-path_____||_Managed sub-path_|

```

GRA = Globally Routable Address, GW = Gateway, RT = Resource Transition Point

8. Security considerations

Specification of a Reference Path and identification of measurement points on the path represent agreements among interested parties, and they present no threat to the implementors of this memo, or to the Internet resulting from implementation of the guidelines provided here.

Attacks at end hosts or identified measurement points are possible. However, there is no requirement to include IP addresses of hosts or other network devices in a reference path with measurement points that is compliant with this memo. As a result, the path diagrams with measurement point designation numbers do not aid such attacks.

Most network operators' diagrams of reference paths will bear a close resemblance to similar diagrams in relevant standards or other publicly available documents. However, when an operator must include atypical network details in their diagram, e.g., to explain why a longer latency measurement is expected, then the diagram reveals some topological details and should be marked as confidential and shared with others under a specific agreement.

When considering privacy of those involved in measurement or those whose traffic is measured, there may be sensitive information

communicated to recipients of the network diagrams illustrating paths and measurement points described above. We refer the reader to the privacy considerations described in the Large Scale Measurement of Broadband Performance (LMAP) Framework [I-D.ietf-lmap-framework], which covers active and passive measurement techniques and supporting material on measurement context. For example, the value of sensitive information can be further diluted by summarising measurement results over many individuals or areas served by the provider. There is an opportunity enabled by forming anonymity sets described in [RFC6973] based on the reference path and measurement points in this memo. For example, all measurements from the Subscriber device can be identified as "mp000", instead of using the IP address or other device information. The same anonymisation applies to the Internet Service Provider, where their Internet gateway would be referred to as "mpl90".

9. IANA Considerations

This memo makes no requests for IANA consideration.

10. Acknowledgements

Thanks to Matt Mathis, Charles Cook, Dan Romascanu, Lingli Deng, and Spencer Dawkins for review and comments.

11. References

11.1. Normative References

- [RFC2330] Paxson, V., Almes, G., Mahdavi, J., and M. Mathis, "Framework for IP Performance Metrics", RFC 2330, May 1998.
- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.
- [RFC3432] Raisanen, V., Grotefeld, G., and A. Morton, "Network performance measurement with periodic streams", RFC 3432, November 2002.
- [RFC5835] Morton, A. and S. Van den Berghe, "Framework for Metric Composition", RFC 5835, April 2010.

11.2. Informative References

- [I-D.ietf-lmap-framework]
Eardley, P., Morton, A., Bagnulo, M., Burbridge, T.,
Aitken, P., and A. Akhter, "A framework for large-scale
measurement platforms (LMAP)", draft-ietf-lmap-
framework-08 (work in progress), August 2014.
- [RFC6973] Cooper, A., Tschofenig, H., Aboba, B., Peterson, J.,
Morris, J., Hansen, M., and R. Smith, "Privacy
Considerations for Internet Protocols", RFC 6973, July
2013.
- [SK] Crawford, Sam., "Test Methodology White Paper", SamKnows
Whitebox Briefing Note
<http://www.samknows.com/broadband/index.php>, July 2011.
- [Q1741] Q.1741.7, , "IMT-2000 references to Release 9 of GSM-
evolved UMTS core network",
<http://www.itu.int/rec/T-REC-Q.1741.7/en>, November 2011.
- [Y.1541] Y.1541, , "Network performance objectives for IP-based
services", <http://www.itu.int/rec/T-REC-Y.1541/en>,
November 2011.

Authors' Addresses

Marcelo Bagnulo
Universidad Carlos III de Madrid
Av. Universidad 30
Leganes, Madrid 28911
SPAIN

Phone: 34 91 6249500
Email: marcelo@it.uc3m.es
URI: <http://www.it.uc3m.es>

Trevor Burbridge
BT
Adastral Park, Martlesham Heath
Ipswich
ENGLAND

Email: trevor.burbridge@bt.com

Sam Crawford
SamKnows

Email: sam@samknows.com

Phil Eardley
BT
Adastral Park, Martlesham Heath
Ipswich
ENGLAND

Email: philip.eardley@bt.com

Al Morton
AT&T Labs
200 Laurel Avenue South
Middletown, NJ
USA

Email: acmorton@att.com

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

This Internet-Draft is submitted to IETF in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

The list of current Internet-Drafts can be accessed at
<http://www.ietf.org/lid-abstracts.html>

The list of Internet-Draft Shadow Directories can be accessed at
<http://www.ietf.org/shadow.html>

Copyright and License Notice

Copyright (c) 2013 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1	Introduction	3
1.1	Terminology	3
2	Use Cases	3
2.1	Internet Service Provider (ISP) Use Case	3
2.2	Regulators	4
2.2.1	Measurement Providers	5
2.2.2	Benchmarking and competitor insight	5
2.3	Fixed and Mobile Service	6
3	Details of ISP Use Case	6
3.1	Existing Capabilities and Shortcomings	6
3.2	Understanding the quality experienced by customers	7
3.3	Understanding the impact and operation of new devices and technology	8
3.4	Design and planning	9
3.5	Identifying, isolating and fixing network problems	10
3.6	Comparison with the regulator use case	12
3.7	Conclusions	13
4	Security Considerations	14
5	IANA Considerations	14
	Appendix A. End User Use Case	14
	Contributors	15
	Normative References	15
	Authors' Addresses	15

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 document 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 an 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 statistically 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 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 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:

- o 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.
- o 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.
- o 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.
- o 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.
- o 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

Network Working Group
Internet-Draft
Intended status: Informational
Expires: January 16, 2014

K. Nagami
Intec
S. Kamei
NTT Communications
K. Koita
IID
T. Jitsuzumi
Kyushu Univ.
I. Mizukoshi
NTT East
July 15, 2013

Use Case from a measurement provider perspective for LMAP
draft-nagami-lmap-use-case-measurement-provider-00

Abstract

This document describes an example of the use cases of measurement provider for LMAP.

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <http://datatracker.ietf.org/drafts/current/>.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on January 16, 2014.

Copyright Notice

Copyright (c) 2013 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect

to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

1. Introduction

In Japan, a high-speed internet of fixed network is used to a lot of people. The use of smartphones and tablet devices increases, the mobile network is also faster. Users know the maximum bandwidth of last mile for fixed network (100Mbps to 2Gbps in FTTH) and mobile network (37.5Mbps to 112.5Mbps in LTE). However, users do not know the performance of the actual network. In order to know them, it is necessary to measure the actual performance of the network. In addition, users need to know the performance of other network providers as information for selecting the network.

Measurement providers measure a network performance from users to multiple content providers using a dedicated hardware or software. They will show the result with measurement methods and measurement parameters.

This document introduces a new use cases of measurement provider for LMAP described in [LMAP-USE-CASES].

2. Use Case for 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 their own. 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.

The following is a case that is currently implemented.

2.1. Measurements for Fixed Networks with dedicated hardware

We have measured a network using dedicated hardware fixed network. We put a dedicated hardware to multiple locations in Japan.

Measurement provider distributes dozens of dedicated hardwares, Linux box with ARM 600MHz ARM Processor and 512M memories.

With major fixed broadband access service in Japan, we can select ISP services using one FTTH access line. Therefore we can measure many ISPs with one FTTH spot. In our trial, we measured 117 ISP-location sets by 9 Linux boxes and 13 ISP accounts.

The architecture of measurement system is composed of Linux box MAS and the control Linux server. They act as follows: Before measurement, we made measurement plan file and set it on control server beforehand.

1. Distributing MAS to some locations.
2. MAS are powered on and connected to FTTH access line.
3. MAS set up PPPoE connection to ISP service.
4. After getting an IP address, MAS set up VPN connection to the control server.
5. MAS fetch measurement plan file from control server, interpret and execute it.
6. MAS put measurement results to control the server.
7. MAS check plan file again, back to step 5.

The plan file is described by the shell script. It can set crontab and can execute any active measurement commands. We also send measurement results from plan file. MAS can act autonomously and can perform without control server.

In our measurement, measuring metrics are RTT, traceroute, HTTP GET (through-put) from/to OTT services.

There are some problems of measurement architecture from this experiment.

- o Scalability is not enough in openvpn tunnel, upload results and checking plan file.
- o It is necessary to consider security risks for plan file in control server.

2.2. Measurements for Fixed Networks with software

The measurement is designed to ask users visit the measurement site. When the user visits the measurement site using a browser, measurement is performed. Because it is measured using the software, we can perform the measurement from the many MAs.

Since the measurement was performed when the end users accessed to the designated measurement site, each individual result was readily available to each participating user himself.

There are two performance measurements. One is performed by an academic researcher and another is performed by a company.

In academic research, the overall summary of all the participants was presented at several conferences that the author attended, and is also available through one of the author's webpage [JITSUZUMI-WEB] and a slideshare site [JITSUZUMI-SLIDESHARE]. These measurements were part of academic researches, and therefore all of the costs were covered by the research grants from the government.

The measurement company shows individual results and a summary of the measurements to participating users, free of charge. The cost of the measurement is covered by offering detailed results, after proper anonymization, to interested companies for fee.

2.3. Measurements for Mobile Networks with smartphone

Measurement provider distributes a measurement applications to users for smart phones. By using this application, the user performs the measurement to the servers from a smartphone.

There are two type of measurements.

1. Run a measurement when users push a measurement button. We have been measured to the servers from 500,000 MAs.
2. Run a measurement periodically. Measurement application is executed periodically in the background. We have been measured to the servers from 2,000 MAs.

3. IANA Considerations

There are no IANA considerations in this memo.

4. Security Considerations

TBD

5. Informative References

[JITSUZUMI-SLIDESHARE]

Jitsuzumi, T., "Jitsuzumi's Slideshare", , <<http://www.slideshare.net/toshiyajitsuzumi/report-actual-qosinjapan>>.

[JITSUZUMI-WEB]

Jitsuzumi, T., "Jitsuzumi's Web Page", , <<http://www.facebook.com/toshiya.jitsuzumi>>.

[LMAP-TERMINOLOGY]

Eardley, P., Morton, A., Bagnulo, M., and T. Burbridge, "Terminology for Large MeAsurement Platforms", draft-eardley-lmap-terminology-02 , July 2013.

[LMAP-USE-CASES]

Linsner, M., Eardley, P., and T. Burbridge, "Large-Scale Broadband Measurement Use Cases", draft-linsner-lmap-use-cases-02 , February 2013.

Authors' Addresses

Kenichi Nagami
Intec

Email: nagami@inetcore.com

Satoshi Kamei
NTT Communications

Email: skame@nttv6.jp

Kenji Koita
IID

Email: k-koita@iid.co.jp

Toshiya Jitsuzumi
Kyushu Univ.

Email: jitsuzumi@econ.kyushu-u.ac.jp

Ichiro Mizukoshi
NTT East

Email: i.mizukoshi@east.ntt.co.jp

LMAP
Internet-Draft
Intended status: Informational
Expires: April 24, 2014

J. Seedorf
NEC
D. Goergen
R. State
University of Luxembourg
V. Gurbani
Bell Labs, Alcatel-Lucent
E. Marocco
Telecom Italia
October 21, 2013

ALTO for Querying LMAP Results
draft-seedorf-lmap-alto-02

Abstract

In the context of Large-Scale Measurement of Broadband Performance (LMAP), measurement results are currently made available to the public either at the finest granularity level (e.g. as a list of results of all individual tests), or in a very high level human-readable format (e.g. as PDF reports). This document argues that there is a need for an intermediate way to provide access to large-scale network measurement results, flexible enough to enable querying of specific and possibly aggregated data. The Application-Layer Traffic Optimization (ALTO) Protocol, defined with the goal to provide applications with network information, seems a good candidate to fulfill such a role. Finally, we describe our methodology for analyzing the United States Federal Communication Commission's (FCC) Measuring Broadband America (MBA) dataset to derive required topology and cost maps suitable for consumption by an ALTO server.

Status of this Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <http://datatracker.ietf.org/drafts/current/>.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on April 24, 2014.

Copyright Notice

Copyright (c) 2013 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1. Introduction	3
2. Example Use Cases	5
3. Advantages of using ALTO	6
4. Examples	7
4.1. Download speeds	7
4.1.1. Network map	8
4.1.2. Cost map	9
5. Discussion of Useful ALTO Extensions	10
6. Case study: Analyzing a large-scale dataset	11
6.1. Challenges in data analysis	11
6.2. Geo-locating the units	12
7. Security considerations	15
8. IANA considerations	16
9. Conclusion	17
10. References	18
10.1. Normative References	18
10.2. Informative References	18
Appendix A. Acknowledgment	19
Authors' Addresses	20

1. Introduction

Recently, there is a discussion on standardizing protocols that would allow measurements of broadband performance on a large scale (LMAP [I-D.schulzrinne-lmap-requirements]). In principle, the vision is that "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."

Apart from protocols that can be used to gather measurement data and to upload such data to dedicated servers, there is also a need for protocols to retrieve - potentially aggregated - measurement results for a certain network (or part of a network), possibly in an automated way. Currently, two extremes are being used to provide access to large-scale measurement results: On the one hand, highly aggregated results for certain networks may be made available in the form of PDFs or figures. Such presentations may be suitable for certain use cases, but certainly do not allow a user (or entity such as a service provider) to select specific criteria and then create corresponding results. On the other hand, complete and detailed results may be made available in the form of comma-separated-values (csv) files. Such data sets typically include the complete results being measured on a very fine-grained level and usually imply large file sizes (of result data sets). Such detailed result data sets are very useful e.g. for the scientific community because they enable to execute complex data analytics algorithms or queries to analyse results.

Considering the two extremes discussed above, this document argues that there is a need for an intermediate way to provide access to large-scale network measurement results: It must be possible to query for specific, possibly aggregated, results in a flexible way. Otherwise, entities interested in measurement results either cannot select what kind of result aggregation they desire, or must always fetch large amounts of detailed results and process these huge datasets themselves. The need for a flexible mechanism to query for dedicated, partial results becomes evident when considering use cases where a service provider or a process wants to use certain measurement results in an automated fashion. For instance, consider a video streaming service provider which wants to know for a given end-user request the average download speed by the end user's access provider in the end user's region (e.g. to optimize/parametrize its http adaptive streaming service). Or consider a website which is interested in retrieving average connectivity speeds for users depending on access provider, region, or type of contract (e.g. to be able to adapt web content on a per-request basis according to such statistics).

This document argues that use cases as described above may enhance the value of measurements of broadband performance on a large scale (LMAP), given that it is possible to query for selected results in an automated fashion. Therefore, in order to facilitate such use cases, a protocol is needed that enables to query LMAP measurements results while allowing to specify certain parameters that narrow down the particular data (i.e. measurement results) the issuer of the query is interested in. This document argues that ALTO [RFC5693] [I-D.ietf-alto-protocol] could be a suitable candidate for such a flexible LMAP result query protocol.

2. Example Use Cases

To motivate the usefulness of ALTO for querying LMAP results, consider some key use cases:

- o Video Streaming Service Provider: For HTTP adaptive streaming, it may be very useful to be able to query for average measurement values regarding a particular end user's access network provider. For instance, consider a video streaming service provider that queries LMAP measurement results to retrieve for a given end-user request the average download speed by the end user's access provider in the end user's region. Such data could help the service provider to optimize/parametrize its HTTP adaptive streaming service.
- o Website Front End Optimization: A website might be interested in statistics about average connectivity types or download speeds for a given end user request in order to dynamically adapt HTML/CSS/JavaScript content depending on such information (sometimes referred to as "Front End Optimization"). For instance, image compression may or may not be employed depending on the average connectivity type/speed of a user in a given region or with a given access network provider.
- o Display estimation of service quality or total download time to users: A webservice could use statistics about average download speeds for a given ISP and/or region to estimate Quality-of-Service for provided services (e.g. to indicate to the user what Quality-of-Experience to expect when clicking on a given link) or to estimate (and display to the user) the total download time for given content.
- o Troubleshooting: In general, any service on the Internet may be interested in LMAP data for troubleshooting. In case a service does not work as expected (e.g. low throughput, high packet loss, ...), it may be of value for the service provider to retrieve (fairly) recent measurement data regarding the host that is requesting the service.
- o TBD: add more use cases

3. Advantages of using ALTO

The ALTO protocol [I-D.ietf-alto-protocol] specifies a very lightweight JSON-based encoding for network information and can play an important role in querying the measurement results as we argue in Section 2.

ALTO is designed on two abstractions that are useful here. First is the abstraction of the physical network topology into an aggregated but logical topology. In this abstract topological view, referred to as "network map", individual hosts are aggregated into a well defined network location identifier called a PID. Hosts could be aggregated into the PID depending on certain identifying characteristics such as geographical location, serving ISP, network mask, nominal access speed, or any mix of them. The "network map" abstraction is essential for exporting network information in a scalable and privacy-preserving way.

The second abstraction that is useful for LMAP is the notion of a "cost map". Each PID identified in the network map can, in a sense, become a vertex in a cost map, and each edge joining adjacent vertices can have an associated cost. The cost can be defined by the measurement server and can indicate routing hops, the financial cost of sending data over the link, available bandwidth on the link with bottlenecked links increasingly showing a smaller value, or a user-defined cost attribute that allows arbitrary reasoning.

The ALTO protocol defines several basic services based on such abstractions, but additional ones can be easily defined as extensions.

There are other advantages to using ALTO as well. The protocol is defined as a set of REST APIs on top of HTTP. The data carried by the protocol is encoded as JSON. Queries can be performed by clients locally after downloading the entire topological and cost maps or clients can send filtered requests to the ALTO server such that the ALTO server performs the required computation and returns the results. The protocol supports a set of atomic constraints related to equality that can be used to filter results and only obtain a set of interest to the query.

Additionally, protocol extensions that could also be useful for the LMAP usage scenario (e.g. extensions for incremental updates, for asynchronous change notifications and for encoding of multiple costs within the same cost map) have been proposed and are currently being discussed in the ALTO WG.

4. Examples

[NOTE: syntax most certainly wrong!]

4.1. Download speeds

This section shows, as an example, how average download speeds measured in a given time interval can be reported. The aggregation approach in this case is based on ISP and geographical location. Two types of data are reported in this example:

- o data collected from measurements against specific endpoints (e.g. active measurements);
- o data collected from all measurements (e.g. passive measurements).

4.1.1. Network map

```
{
  "meta" : {},
  "data" : {
    "map-vtag" : "1266506139",
    "map" : {
      "ISP1-GEO1" : {
        "ipv4" : [ "10.1.0.0/16", "172.20.0.0/16" ]
      },
      "ISP2-GEO1" : {
        "ipv4" : [ "10.2.0.0/17" ]
      },
      "ISP3-GEO1" : {
        "ipv4" : [ "10.3.0.0/16" ]
      },
      "ISP2-GEO2" : {
        "ipv4" : [ "10.2.128.0/17" ]
      },
      "ISP4-GEO2" : {
        "ipv4" : [ "10.4.0.0/16" ]
      },
      .
      .
      .
      "MSMNT-CL1" : {
        "ipv4" : [ "192.168.0.0/30" ]
      },
      "TOTAL" : {
        "ipv4" : [ "0.0.0.0/0" ]
      }
    }
  }
}
```


4.1.2. Cost map

```
{
  "meta" : {},
  "data" : {
    "cost-mode" : "numerical",
    "cost-type" : "avg-dl-speed",
    "map-vtag" : "1266506139",
    "time-interval" : "2629740",
    "map" : {
      "ISP1-GEO1": { "MSMNT-CL1" : 13.2,
                     "TOTAL" : 10.2},
      "ISP2-GEO1": { "MSMNT-CL1" : 11.4,
                     "TOTAL" : 12.3},
      "ISP3-GEO1": { "MSMNT-CL1" : 13.2,
                     "TOTAL" : 10.2},
      .
      .
      .
    }
  }
}
```

5. Discussion of Useful ALTO Extensions

The base ALTO Protocol as specified in [I-D.ietf-alto-protocol] can in principle be used to enable a more flexible way to provide access to large-scale network measurement results as discussed in the previous sections of this document. However, certain extensions to the base ALTO Protocol that have recently been proposed in the ALTO WG would allow to better enable the use cases discussed in Section 2:

- o Server-initiated Notifications: In [I-D.marocco-alto-ws], it has been proposed to enhance the ALTO protocol such that servers can notify clients about newly available ALTO maps. In the context of this document, this extension would allow applications to be notified when certain new LMAP measurements are available, such as new measurement results on average download speeds. These new results could then be downloaded and used immediately by applications.
- o Incremental Updates: In [I-D.schwan-alto-incr-updates], it has been proposed to enhance the ALTO protocol with incremental updates, such that clients can retrieve partial updates for ALTO maps instead of always downloading a full ALTO map (even when only a small fraction of the ALTO map has changed compared to a previous version). When ALTO is used for querying LMAP results, the corresponding ALTO maps may potentially be quite large (e.g. when a webservice queries for particular, detailed results regarding a whole ISP). In this case, incremental ALTO updates would be a very useful mechanism for applications to retrieve updates of ALTO maps, as a reduced amount of data would be needed for transmitting these maps.

6. Case study: Analyzing a large-scale dataset

Measuring broadband performance is increasingly important as communications continue to move towards the Internet. Internet service providers (ISP), national agencies and other entities gather broadband data and may provide some, or all, of the dataset to the public for analysis. As we argue above, there are two extremes prevalent for presenting large-scale data. One is in the form of charts, figures, or summarized reports amenable for easy and quick consumption. The other extreme includes releasing raw data in the form of large files containing tables formatted as values separated by a delimiter. While the former is indispensable to acquire a summary view of the dataset, it does not suffice for additional analysis beyond what is presented. Conversely, the problem with the latter option (raw files) is that the unsuspecting user perusing them is lost in the deluge of data.

We offer the argument that a reasonable medium between the two extremes may be the ALTO protocol [I-D.ietf-alto-protocol]. A necessary prerequisite for using ALTO is abstracting the network information into a form that is suitable for consumption by the protocol. The implication of using ALTO is that data from any large-scale measurement effort must first be distilled in two maps: a topology map and a cost map. Further analysis and ad-hoc queries can be subsequently performed on the normalized dataset.

In the United States, the Federal Communication Commission (FCC) has embarked on a nationwide performance study of residential wireline broadband service [fcc]. Our aim is to use the raw datasets from this study for analysis and to create a topology map and a cost map from this dataset. ALTO queries aimed at these maps will enable users and interested parties to fulfill the use cases listed in Section 2.

6.1. Challenges in data analysis

The FCC Measuring Broadband America (MBA) study consisted of 7,782 volunteers spread across the United States with adequate geographic diversity. Volunteers opted in for the study, however, each of the volunteers remained anonymous. An opaque integral number (unit_id) represented a subscriber in the raw dataset. This unit_id remains constant during the duration of the study in the dataset and uniquely identifies a volunteer subscriber, even if the subscriber switches the ISP. More detail about the methodology used is described in [fcc].

The dataset consisted of 12 tables, each table corresponding to the data drawn from a certain performance test. For the analysis we

present in this document we focus on the "curr_dns" table, which contains the time taken for the ISP's recursive DNS resolver to return a DNS A RR for a popular website domain name. This test was ran approximately every hour in a 24-hour period, and produced about 75-78 million records per month. This resulted in a typical file size in the range of 6-7 GBytes per month. We note that the "curr_dns" table is one of the smaller tables in the dataset.

The first challenge, therefore, was to arrive at computing resources comparable in scale with respect to the dataset consisting of millions of records spread across gigabyte-sized files. To analyze the volume of data we used a canonical Map-Reduce computational paradigm on a Hadoop cluster (more details on the methodology are outlined in Section 6.2).

A second, more pressing challenge, was to identify the geographic location of the unit_ids generating the data. In order to derive a topological map and impose costs on the links, it is important to know the physical locations of the unit_ids that contributed the measurements. However, in the MBA dataset, the population is anonymized and the individual subscriber reporting the measurement data is simply referred to by an opaque integral number. Therefore, an important task was to use the information in the public tables to reveal a coarse location of the subscriber.

We outline the methodology we used to do so in the next section. We stress that this methodology does not identify the specific location of a subscriber, who still remains anonymous. Instead, it simply locates the subscriber in a larger metropolitan region. This level of granularity suffices for our work.

6.2. Geo-locating the units

To geo-locate the units, we simply note that broadband subscriber devices are likely to be configured using DHCP by their ISP. Besides imparting an IP address to the subscriber device, DHCP also populates the DNS name servers the subscriber devices uses for DNS queries. In most installations, these DNS name servers are located in close physical proximity of the subscriber device. The FCC technical appendix states that the DNS resolution tests were targeted directly at the ISP's recursive resolvers to circumvent caching and users configuring the subscriber device to circumvent the ISP's DNS resolvers. Therefore, a reasonable approximation of a subscribers geo-location could be the geographic location of the DNS name server serving the subscriber. We use this very heuristic to geo-locate a subscriber.

Thus our first, and very simple filter consisted of obtaining a

mapping from a unit_id (representing a subscriber) to one or more DNS name servers that the unit_id is sending DNS requests to. It turned out that while this was a necessary condition for advancing, it was not a sufficient one. The raw data would need to be further processed to reduce inconsistencies and remove outliers. A number of interesting artifacts were uncovered during further processing of the data. These artifacts informed the selection of the unit_ids for further analysis.

The artifacts are documented below.

- o A handful of unit_ids were geo-located in areas outside the contiguous United States, such as Ukraine, Poland or the United Kingdom. We theorize that the subscribers corresponding to the unit_ids geo-located outside the contiguous United States had simply configured their devices to use alternate DNS servers, probably located outside the United States. We removed these records before conducting our analysis.
- o We also observed a reasonable number of non-ISP DNS resolvers, especially Google's 8.8.8.8 and 8.8.4.4 and OpenDNS 208.67.222.222 and 208.67.220.220. These 4 public DNS servers are geo-located in California. We removed these records to ensure that the specific location that these resolvers represented was not oversampled.
- o We noticed that a large number of unit_ids were being geo-located in Potwin, Kansas. Intrigued as to why there appeared to be a large population of Internet users being located in a small rural community in Kansas, we investigated further. It appears that Potwin, Kansas is the geographical center of the United States and a number of ISPs have chosen to establish data centers in or around the Potwin area. These ISPs generally locate their primary or secondary DNS name servers in Potwin-area data centers, thus accounting for the popularity of Potwin as an Internet destination. We continue to further investigate on minimizing the impact of such natural aggregation points that, if not accounted for, will skew our results in an unwarranted direction.
- o We observed some unit_ids changing ISPs during the observation period. This is a normal occurrence and to the extent that the unit_id is geo-located in the same geographical area after the change in ISP, we do not exclude such unit_ids from further analysis.

Subsequent filters extracted the stable unit_ids from our dataset. In order to determine which unit_id are stable, i.e., remain constant with respect to their geographic location over the observation period from January to December 2012, we extracted for each unit_id the IP

address of each DNS name server it consulted. This is obtained by applying the map reduce paradigm on the DNS dataset. We extracted for each `unit_id` the triggered DNS servers and obtained the individual DNS servers accessed by a `unit_id`. This was repeated for each month of the observation period. The resulting sets were cleaned up of private IP addresses and other artifacts discussed above. The cleaned set consisted of about 8000 distinct `unit_id`.

In order to determine the stability of each `unit_id` we proceeded to sum up the occurrences of IP addresses over the whole observation period separated in monthly files. If the IP address of a DNS server occurred 12 times this meant that the `unit_id` always accessed the same DNS server and therefore remained stable over the observation period. The obtained stable `unit_ids`, around 1500, will be used for further analysis. Assuming a 99% confidence level and ± 3 point margin of error, we will require a sample of 1494 `unit_ids`. With our stable `unit_id` set of 1500 `unit_ids`, we are now positioned to perform further analysis on the dataset to create the full topology and cost maps.

Table 1 presents a sample of the geographic location data that we have uncovered for `unit_ids`. A complete list of identified units superimposed on the geographical map of the United States is available at <http://cdb.io/13UOHgD>.

Unit ID	City, State	Latitude/Longitude
872	Morganville, NJ	40.35950089,-74.26280212
885	Madison, WI	43.07310104,-89.40119934
898	Foley, AL	30.40660095,-87.68360138
7969	Manteca, CA	37.79740143,-121.2160034
8024	Quincy, MA	42.25289917,-71.00229645

Sample unit identification tuples

Table 1

7. Security considerations

There are no security artifacts invalidated due to our analysis in Section 6. All of our analysis was performed on publicly available data. However, we do note that some privacy may have been lost based on our analysis. In the raw dataset, the unit identifiers are opaque strings with no immediate correlation with a geographic location. After our analysis, while the unit identifiers still remain opaque, they are nonetheless correlated to a specific, though coarse, geographic location.

8. IANA considerations

This document does not contain any IANA considerations.

9. Conclusion

This document argues that, compared to existing solutions, there may be a need for a more flexible way to provide access to large-scale network measurement results. Further, the document argues that the ALTO protocol is a good candidate to enable querying for specific, possibly aggregated, measurement results in a flexible way. Examples of how such a flexible query mechanism for large-scale measurement results could look like based on ALTO are given.

With respect to the case study in Section 6, identification of the geographic location of the unit_ids generating the performance data is essential in order to continue the work. We have presented a methodology and some early results in identifying a geographic location. This location, although coarse, suffices for our future work that will consist of further data mining and analysis to create appropriate ALTO network and cost maps.

10. References

10.1. Normative References

- [RFC5693] Seedorf, J. and E. Burger, "Application-Layer Traffic Optimization (ALTO) Problem Statement", RFC 5693, October 2009.

10.2. Informative References

- [I-D.ietf-alto-protocol]
Alimi, R., Penno, R., and Y. Yang, "ALTO Protocol", draft-ietf-alto-protocol-20 (work in progress), October 2013.
- [I-D.marocco-alto-ws]
Marocco, E. and J. Seedorf, "WebSocket-based server-to-client notifications for the Application-Layer Traffic Optimization (ALTO) Protocol", draft-marocco-alto-ws-01 (work in progress), July 2012.
- [I-D.schulzrinne-lmap-requirements]
Schulzrinne, H., Johnston, W., and J. Miller, "Large-Scale Measurement of Broadband Performance: Use Cases, Architecture and Protocol Requirements", draft-schulzrinne-lmap-requirements-00 (work in progress), September 2012.
- [I-D.schwan-alto-incr-updates]
Schwan, N. and B. Roome, "ALTO Incremental Updates", draft-schwan-alto-incr-updates-02 (work in progress), July 2012.
- [fcc] United States Federal Communications Commission, "Measuring Broadband America", Accessed July 12, 2013, <http://www.fcc.gov/measuring-broadband-america>.

Appendix A. Acknowledgment

Jan Seedorf is partially supported by the mPlane project (mPlane: an Intelligent Measurement Plane for Future Network and Application Management), a research project supported by the European Commission under its 7th Framework Program (contract no. 318627). The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of the mPlane project or the European Commission.

Authors' Addresses

Jan Seedorf
NEC
Kurfuerstenanlage 36
Heidelberg 69115
Germany

Phone: +49 6221 4342 221
Fax: +49 6221 4342 155
Email: seedorf@neclab.eu

David Goergen
University of Luxembourg

Email: david.goergen@uni.lu

Radu State
University of Luxembourg

Email: radu.state@uni.lu

Vijay K. Gurbani
Bell Labs, Alcatel-Lucent

Email: vkg@bell-labs.com

Enrico Marocco
Telecom Italia
Via G. Reiss Romoli, 274
Turin 10148
Italy

Email: enrico.marocco@telecomitalia.it

