

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

P. Eardley  
BT  
A. Morton  
AT&T Labs  
M. Bagnulo  
UC3M  
T. Burbridge  
BT  
P. Aitken  
A. Akhter  
Cisco Systems  
October 21, 2013

A framework for large-scale measurement platforms (LMAP)  
draft-ietf-lmap-framework-01

## 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. It also defines terminology for LMAP (large-scale measurement platforms). The document is a contribution towards the LMAP working group's milestone.

## 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.

Internet-Draft

LMAP Framework

October 2013

This document is subject to [BCP 78](http://trustee.ietf.org/license-info) 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

<a href="#">1.</a>	Introduction . . . . .	<a href="#">3</a>
<a href="#">2.</a>	Terminology . . . . .	<a href="#">5</a>
<a href="#">3.</a>	Outline of an LMAP-based measurement system . . . . .	<a href="#">7</a>
<a href="#">4.</a>	Constraints . . . . .	<a href="#">10</a>
4.1.	Measurement system is under the direction of a single organisation . . . . .	<a href="#">11</a>
4.2.	Each MA may only have a single Controller at any point in time . . . . .	<a href="#">11</a>
<a href="#">5.</a>	LMAP Protocol Model . . . . .	<a href="#">11</a>
<a href="#">5.1.</a>	Bootstrapping process . . . . .	<a href="#">12</a>
<a href="#">5.2.</a>	Control Protocol . . . . .	<a href="#">14</a>
<a href="#">5.3.</a>	Starting and stopping Measurement Tasks . . . . .	<a href="#">16</a>
<a href="#">5.4.</a>	Report Protocol . . . . .	<a href="#">17</a>
<a href="#">5.5.</a>	Items beyond the scope of the LMAP Protocol Model . . . . .	<a href="#">18</a>
<a href="#">5.5.1.</a>	User-controlled measurement system . . . . .	<a href="#">19</a>
<a href="#">6.</a>	Details of the LMAP framework . . . . .	<a href="#">20</a>
<a href="#">6.1.</a>	Measurement Agent (MA) . . . . .	<a href="#">20</a>
<a href="#">6.1.1.</a>	Measurement Agent embedded in site gateway . . . . .	<a href="#">20</a>
6.1.2.	Measurement Agent embedded behind Site NAT /Firewall	21
<a href="#">6.1.3.</a>	Measurement Agent in-line with site gateway . . . . .	<a href="#">21</a>
<a href="#">6.1.4.</a>	Measurement Agent in multi homed site . . . . .	<a href="#">22</a>
<a href="#">6.2.</a>	Measurement Peer (MP) . . . . .	<a href="#">22</a>
<a href="#">6.3.</a>	Controller . . . . .	<a href="#">23</a>
<a href="#">6.4.</a>	Collector . . . . .	<a href="#">23</a>
<a href="#">7.</a>	Security considerations . . . . .	<a href="#">23</a>
<a href="#">8.</a>	Privacy Considerations for LMAP . . . . .	<a href="#">24</a>
<a href="#">8.1.</a>	Categories of Entities with Information of Interest . . . . .	<a href="#">25</a>
<a href="#">8.2.</a>	Examples of Sensitive Information . . . . .	<a href="#">25</a>
8.3.	Key Distinction Between Active and Passive Measurement Tasks . . . . .	<a href="#">26</a>
<a href="#">8.4.</a>	Communications Model (for Privacy) . . . . .	<a href="#">26</a>

<a href="#">8.4.1.</a>	Controller <-> Measurement Agent . . . . .	<a href="#">27</a>
<a href="#">8.4.2.</a>	Collector <-> Measurement Agent . . . . .	<a href="#">28</a>
<a href="#">8.4.3.</a>	Active Measurement Peer <-> Measurement Agent . . . . .	<a href="#">28</a>
<a href="#">8.4.4.</a>	Passive Measurement Peer <-> Measurement Agent . . . . .	<a href="#">29</a>
<a href="#">8.4.5.</a>	Result Storage and Reporting . . . . .	<a href="#">30</a>

<a href="#">8.5.</a>	Threats . . . . .	<a href="#">30</a>
<a href="#">8.5.1.</a>	Surveillance . . . . .	<a href="#">30</a>
<a href="#">8.5.2.</a>	Stored Data Compromise . . . . .	<a href="#">31</a>
<a href="#">8.5.3.</a>	Correlation and Identification . . . . .	<a href="#">31</a>
<a href="#">8.5.4.</a>	Secondary Use and Disclosure . . . . .	<a href="#">31</a>
<a href="#">8.6.</a>	Mitigations . . . . .	<a href="#">32</a>
<a href="#">8.6.1.</a>	Data Minimization . . . . .	<a href="#">32</a>
<a href="#">8.6.2.</a>	Anonymity . . . . .	<a href="#">33</a>
<a href="#">8.6.3.</a>	Pseudonymity . . . . .	<a href="#">34</a>
<a href="#">8.6.4.</a>	Other Mitigations . . . . .	<a href="#">34</a>
<a href="#">8.7.</a>	The potential role of a Group-ID for privacy . . . . .	<a href="#">34</a>
<a href="#">9.</a>	IANA Considerations . . . . .	<a href="#">36</a>
<a href="#">10.</a>	Acknowledgments . . . . .	<a href="#">36</a>
<a href="#">11.</a>	History . . . . .	<a href="#">36</a>
<a href="#">11.1.</a>	From -00 to -01 . . . . .	<a href="#">37</a>
<a href="#">12.</a>	Informative References . . . . .	<a href="#">37</a>
	Authors' Addresses . . . . .	<a href="#">38</a>

## [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. Several use cases have been proposed for large-scale measurements including:

- o Operators: to help plan their network and identify faults
- o Regulators: to benchmark several network operators and support public policy development

Further details of the use cases can be found at [\[I-D.linsner-lmap-use-cases\]](#). The LMAP framework should be useful for these, as well as other use cases that the LMAP WG doesn't concentrate on, such as to help end users run diagnostic checks like a network speed test.

The LMAP framework has four basic elements: Measurement Agents, Measurement Peers, Controllers and Collectors.

Measurement Agents (MAs) perform network measurements. They are pieces of code that can be executed in specialized hardware (hardware probe) or on a general-purpose device (like a PC or mobile phone).

The Measurement Agents may have multiple interfaces (WiFi, Ethernet, DSL, fibre, etc.) and the measurements may specify any one of these. Measurements may be active (the MA or Measurement Peer (MP) generates test traffic), passive (the MA observes user traffic), or some hybrid form of the two. For active measurement tasks, the MA (or MP) generates test traffic and measures some metric associated with its transfer over the path to (or from) a Measurement Peer. For example, one active measurement task could be to measure the UDP latency between the MA and a given MP. MAs may also conduct passive testing through the observation of traffic. The measurements themselves may be on IPv4, IPv6, and on various services (DNS, HTTP, XMPP, FTP, VoIP, etc.).

The Controller manages one or more MAs by instructing it which measurement tasks it should perform and when. For example it may instruct a MA at a home gateway: "Measure the 'UDP latency' with the Measurement Peer mp.example.org; repeat every hour at xx.05". The Controller also manages a MA by instructing it how to report the measurement results, for example: "Report results once a day in a batch at 4am". We refer to these as the Measurement Schedule and Report Schedule.

The Collector accepts Reports from the MAs with the results from their measurement tasks. Therefore the MA is a device that initiates the measurement tasks, gets instructions from the Controller and reports to the Collector.

There are additional elements that are part of a measurement system, but that are out of the scope for LMAP. We provide a detailed

discussion of all the elements in the rest of the document.

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.

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.

The desirable features for a large-scale measurement systems we are designing for are:

- 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 Large-scale - [[I-D.linsner-lmap-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.](#) Terminology

This section defines terminology for LMAP. Please note that defined terms are capitalized.

**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 be integrated into a measurement system.

**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. It also delivers logging information and capabilities information from the Measurement Agent to the Controller.

**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.

**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.

**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.

Information Model: The protocol-neutral definition of the semantics of the Instructions, the Report, the status of the different elements of the measurement system as well of the events in the system.

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

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 Suppression: a type of Instruction that stops (suppresses) 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 Channel:** a specific Report Schedule and 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.

**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 services, to register a user or a list of users authorized to enjoy these services, and also to set the limits relative to the use that associated users make of these services. (This definition is from [\[Q1741\]](#).)

**Test Traffic:** for Active Measurement Tasks, the traffic generated by the Measurement Agent and/or the Measurement Peer to execute the requested Measurement Task.

### [3.](#) Outline of an LMAP-based measurement system

Figure 1 shows the main components of a measurement system, and the interactions of those components. Some of the components are outside the scope of LMAP. In this section we provide an overview on the whole measurement system, whilst the subsequent sections study the LMAP components in more detail.

The first component is a Measurement Task, which measures some



performance or reliability Metric of interest. An Active Measurement Task involves either a Measurement Agent injecting Test Traffic into the network destined for a Measurement Peer, and/or a MP sending Test Traffic to a MA; one of them measures the some parameter associated with the transfer of the packet(s). A Passive Measurement Task involves only a MA, which simply observes existing traffic – for example, it could simply count bytes or it might calculate the average loss for a particular flow.

It is very useful to standardise Measurement Methods (a Measurement Method is a generalisation of a Measurement Task), so that it is meaningful to compare measurements of the same Metric made at different times and places. It is also useful to define a registry for commonly-used Metrics [[registry](#)] so that a Measurement Method can be referred to simply by its identifier in the registry. The Measurement Methods and registry would hopefully also be referenced by other standards organisations.

In order for a Measurement Agent and a Measurement Peer to execute an Active Measurement Task, they exchange Test Traffic. The protocols used for the Test Traffic is out of the scope of the LMAP WG and falls within the scope of the IETF WGs such as IPPM.

For Measurement Results to be truly comparable, as might be required by a regulator, not only do the same Measurement Methods need to be used but also the set of Measurement Tasks should follow a similar Measurement Schedule and be of similar number. The details of such a characterisation plan are beyond the scope of work in IETF although certainly facilitated by IETF's work.

The next components we consider are the Measurement Agent (MA), Controller and Collector. The main work of the LMAP working group is to define the Control Protocol between the Controller and MA, and the Report Protocol between the MA and Collector. [Section 4](#) onwards considers the LMAP compnents in more detail; here we introduce them.

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 Measurement 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.

Finally we introduce several components that are out of scope of the LMAP WG and will be provided through existing protocols or applications. They affect how the measurement system uses the Measurement Results and how it decides what set of Measurement Tasks to perform.

The MA needs to be bootstrapped with initial details about its Controller, including authentication credentials. The LMAP WG considers the bootstrap process, since it affects the Information Model. However, it does not define a bootstrap protocol, since it is likely to be technology specific and could be defined by the Broadband Forum, DOCSIS or IEEE, depending on the device. 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 (perhaps 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.

Internet-Draft

LMAP Framework

October 2013

A Results Database records all measurements in an equivalent form, for example an SQL database, so that they can be easily accessed by the Data Analysis Tools. The Data Analysis Tools also need to understand the Subscriber's 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 Maintenance) uses the results from the tools.

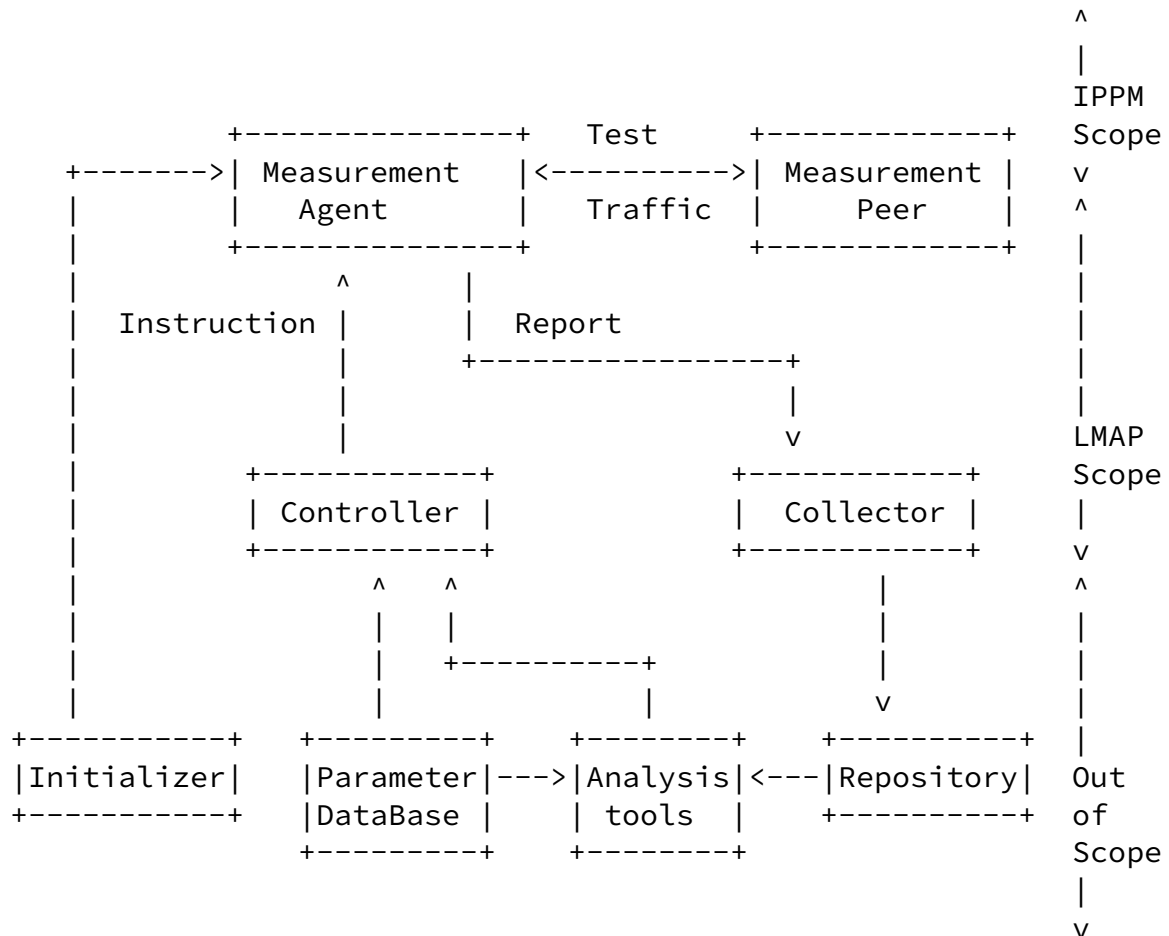


Figure 1: Schematic of main elements of an LMAP-based measurement system  
(showing the elements in and out of the scope of the LMAP WG)

#### [4.](#) Constraints

Eardley, et al.

Expires April 24, 2014

[Page 10]

---

Internet-Draft

LMAP Framework

October 2013

The LMAP framework makes some important assumptions, which constrain the scope of the work to be done.

##### [4.1.](#) 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.

##### [4.2.](#) Each MA may only have a single Controller at any point in time

A MA is instructed by one Controller and is in one measurement system. The constraint avoids different Controllers giving a MA conflicting instructions and so means that the MA does not have to manage contention between multiple Measurement (or Report) Schedules. This simplifies the design of MAs (critical for a large-scale infrastructure) and allows a Measurement Schedule to be tested on specific types of MA before deployment to ensure that the end 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).

## 5. LMAP Protocol Model

A protocol model presents ([RFC4101](#)) "an architectural model for how the protocol operates ... a short description of the system in overview form, ... [which] needs to answer three basic questions:

1. What problem is the protocol trying to achieve?
2. What messages are being transmitted and what do they mean?
3. What are the important, but unobvious, features of the protocol?"

An LMAP system goes through the following phases:

Eardley, et al.

Expires April 24, 2014

[Page 11]

---

Internet-Draft

LMAP Framework

October 2013

- o a bootstrapping process before the MA can take part in the three items below
- o a Control Protocol, which delivers an Instruction from a Controller and a MA. The Instruction details what Measurement Tasks the MA should perform and when, and how it should report the Measurement Results
- o the actual Measurement Tasks are performed. An Active Measurement Task involves sending test traffic between the Measurement Agent and a Measurement Peer, whilst a Passive Measurement Task involves (only) the Measurement Agent observing existing user traffic. The LMAP WG does not define Measurement Methods, however the IPPM WG does.
- o a Report Protocol, which delivers a Report from the MA to a Collector. The Report contains the Measurement Results.

In the diagrams the following convention is used:

- o (optional): indicated by round brackets
- o [potentially repeated]: indicated by square brackets

The Protocol Model is closely related to the Information Model, which

is the abstract definition of the information carried by the protocol model. The purpose of both is to provide a protocol and device independent view, which can be implemented via specific protocols. The LMAP WG will define a specific Control Protocol and Report Protocol, but other Protocols could be defined by other standards bodies or be proprietary. However it is important that they all implement the same Information and Protocol Model, in order to ease the definition, operation and interoperability of large-scale measurement systems.

### [5.1.](#) Bootstrapping process

The primary purpose of bootstrapping is to enable the MA and Controller to be integrated into a measurement system. In order to do that, the MA needs to retrieve information about itself (like its identity in the measurement system), about the Controller and the Collector(s) as well as security information (such as certificates and credentials).

```
+-----+
| Measurement |
|   Agent     |
+-----+
```

(Initial Controller details:  
address or FQDN,  
security credentials)

->

```
+-----+
|   Initial   |
| Controller  |
+-----+
```

<-

(register)

Controller details:  
address or FQDN,  
security credentials

->

```
+-----+
| Controller  |
+-----+
```

<-

register

Typically the MA is behind a NAT, so needs to initiate communications, in order that the Controller can communicate with it. The normal NAT interactions are not shown in the figure.

The MA knows how to contact a Controller through some device /access specific mechanism. For example, this could be in the firmware, downloaded, manually configured or via a protocol like TR-069. The Controller could either be the one that will send it Instructions (see next sub-section) or else an initial Controller. The role of an initial Controller is simply to inform the MA how to contact its actual Controller; this could be useful, for example, for load balancing or if the details of the initial Controller are statically configured or if the measurement system has specific Controllers for different devices types. When the MA registers with the Controller it learns its MA identifier; it may also be told a Group-ID and whether to include the MA-ID as well as the Group-ID in its Reports. A Group-ID would be shared by several MAs and could be useful for privacy reasons (for instance to hinder tracking of a mobile MA device). The MA may also tell the Controller the list of Measurement Methods that its capable of (see next sub-section).

Whilst the LMAP WG considers the bootstrapping process, it is out of scope to define a bootstrap mechanism, as it depends on the type of device and access.

Open issue: what happens if a Controller fails, how is the MA is homed onto a new one?

## [5.2.](#) Control Protocol

The primary purpose of the Control Protocol is to allow the Controller to configure a Measurement Agent with Measurement Instructions, which it then acts on autonomously.

```
+-----+
|               |
```

```
+-----+
| Measurement |
```

Controller	Agent
------------	-------

Instruction:

[(Measurement Task (parameters)), (Measurement Schedule), (Report Channel(s))]	->	
	<-	ACK
(Capability request)	->	
	<-	List of Measurement Methods
ACK	->	
Suppress	->	
	<-	Failure report: (reason)
ACK	->	

The Instruction contains:

- o what measurements to do: the Measurement Methods 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')
- o when to do them: the Measurement Schedule details the timings of regular tests, one-off tests
- o how to report the Measurement Results: via Reporting Channel(s), each of which defines a target Collector and Report Schedule

An Instruction could contain one or more of the above elements, since the Controller may want the MA to perform several different Measurement Tasks (measure UDP latency and download speed), at several frequencies (a regular test every hour and a one-off test



immediately), and report to several Collectors. The different elements can be updated independently at different times and regularities, for example it is likely that the Measurement Schedule will be updated more often than the other elements.

In general we expect that the Controller knows what Measurement Methods the MA supports, such that the Controller can correctly instruct the MA. Note that the Control Protocol does not allow negotiation (which would add complexity to the MA, Controller and Control Protocol for little benefit).

The MA can send to the Controller the complete list of Measurement Methods that it is capable of. Note that it is not intended to indicate dynamic capabilities like the MA's currently unused CPU, memory or battery life. The list of Measurement Methods could be useful in several circumstances: when the MA first communicates with a Controller; when the MA becomes capable of a new Measurement Method; when requested by the Controller (for example, if the Controller forgets what the MA can do or otherwise wants to resynchronize what it knows about the MA).

The Controller has the ability to send a "suppress" message to MAs. This could be useful if there is some unexpected network issue and so the measurement system wants to eliminate inessential traffic. As a result, temporarily the MA does not start new Active Measurement Tasks, and it may also stop in-progress Measurement Tasks, especially ones that are long-running &/or creates a lot of traffic. See the next section for more information on stopping Measurement Tasks.

The figure shows that the various messages are acknowledged, which means that they have been delivered successfully. However, the "suppress" message is not acknowledged, since it is likely to be broadcast to several /many MAs at a time when the measurement system wants to eliminate inessential traffic. Note also that the MA does not inform the Controller about Measurement Tasks starting and stopping.

There is no need for the MA to confirm to the Controller that it has understood and acted on the Instruction, since the Controller knows the capabilities of the MA. However, the Control Protocol must support robust error reporting by the MA, to provide the Controller with sufficiently detailed reasons for any failures. There are two broad categories of failure: the MA cannot action the Instruction (for example, it doesn't include a parameter that is mandatory for

the requested Measurement Method); or the Measurement Task could not be executed (for example, the MA unexpectedly has no spare CPU cycles). Note that it is not considered a failure if a Measurement Task (correctly) doesn't start - for example if the MA detects cross-traffic; instead this is reported to the Collector in the normal manner (see Section below).

Comment: the detailed list of reasons below would be more appropriate in the Information Model i-d.

- o no value for a mandatory parameter
- o time of test is in past
- o type wrong, eg string given where expect integer
- o Schedule refers to a Measurement configuration or Report Channel that doesn't exist
- o MA has crashed
- o MA doesn't (any longer) understand requested Method
- o MA has run out of CPU, memory, battery power
- o Collector has disappeared
- o MP has disappeared

Finally, note that the MA doesn't do a 'safety check' with the Controller (that it should still continue with the requested Measurement Tasks) - it simply carries out the Measurement Tasks as instructed, unless it gets an updated Instruction.

The LMAP WG will define a Control Protocol and its associated Data Model that implements the Protocol & Information Model. 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-YANG.

### [5.3.](#) Starting and stopping Measurement Tasks

The LMAP WG is neutral to what the actual Measurement Task is. The WG does not define a generic start and stop process, since the correct approach depend on the particular Measurement Task; the details are defined as part of each Measurement Method, and hence potentially by the IPPM WG.

Internet-Draft

LMAP Framework

October 2013

Once the MA gets its Measurement and Report Schedules from its Controller then it acts autonomously, in terms of operation of the Measurement Tasks and reporting of the result. One implication is that the MA initiates Measurement Tasks. Therefore for the common case where the MA is on a home gateway, the MA initiates a 'download speed test' by asking a Measurement Peer to send the file.

Many Active Measurement Tasks begin with a pre-check before the test traffic is sent. Action could include:

- o the MA checking that there is no cross-traffic (ie that the user isn't already sending traffic);
- o the MA checking with the Measurement Peer that it can handle a new Measurement Task (in case the MP is already handling many Measurement Tasks with other MAs);
- o the first part of the Measurement Task consisting of traffic that probes the path to make sure it isn't overloaded.

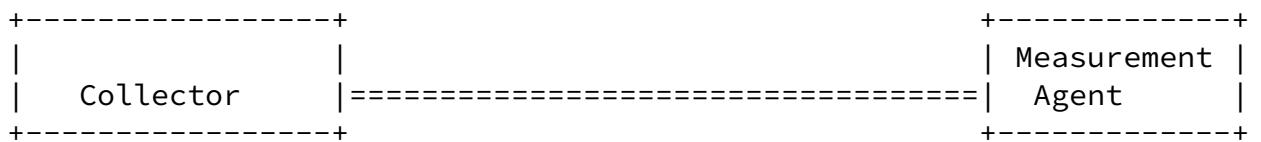
It is possible that similar checks continue during the Measurement Task, especially one that is long-running &/or creates a lot of Test Traffic, which may be abandoned whilst in-progress. A Measurement Task could also be abandoned in response to a "suppress" message (see previous section). Action could include:

- o For 'upload' tests, the MA not sending traffic
- o For 'download' tests, the MA closing the TCP connection or sending a TWAMP Stop control message.

Comment: presumably Passive Measurement Tasks don't do pre-checking or stopping?

#### [5.4.](#) Report Protocol

The primary purpose of the Report Protocol is to allow a Measurement Agent to report its Measurement Results to a Collector, and the context in which they were obtained.



<- Report:  
[MA-ID &/or Group-ID,  
Measurement Results,

Eardley, et al.

Expires April 24, 2014

[Page 17]

Internet-Draft

LMAP Framework

October 2013

ACK -> Measurement Task]

The MA acts autonomously in terms of reporting; it simply sends Reports as defined by the Controller's Instruction.

The Report contains:

- o the MA's identifier, or perhaps a Group-ID to anonymise results
- o the actual Measurement Results, including the time they were measured
- o the details of the Measurement Task (to avoid the Collector having to ask the Controller for this information later)

Depending on the requirements of the measurement system, the MA might label, or perhaps not include, Measurement Results impacted by for instance cross-traffic or the MP being busy. If applicable the Measurement Report includes the start and end of suppression.

The MA may report the results to more than one Collector, if the Instruction says so. It could report a different subset of Results to different Collectors.

The LMAP WG will define a Report Protocol and its associated Data Model that implements the Protocol & Information Model. 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 IPFIX.

## [5.5.](#) Items beyond the scope of the LMAP Protocol Model

There are several potential interactions between LMAP elements that are out of scope of definition by the LMAP WG:

1. It does not define a coordination process between MAs. Whilst a measurement system may define coordinated Measurement Schedules across its various MAs, there is no direct coordination between MAs.

2. It does not define interactions between the Collector and Controller. It is quite likely that there will be such interactions, probably intermediated by the data analysis tools. For example if there is an "interesting" Measurement Result then the measurement system may want to trigger extra Measurement Tasks that explore the potential cause in more detail.
3. It does not define coordination between different measurement systems. For example, it does not define the interaction of a MA in one measurement system with a Controller or Collector in a different measurement system. Whilst it is likely that the Control and Report protocols could be re-used or adapted for this scenario, any form of coordination between different organisations 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 the LMAP WG. Note that a single MA is instructed by a single Controller and is only in one measurement system.
  - \* An interesting scenario is where a home contains two independent MAs, for example one controlled by a regulator and one controlled by an ISP. Then the test traffic of one MA is treated by the other MA just like any other user traffic.
4. It does not specifically define a user-initiated measurement system, see sub-section.

### [5.5.1.](#) User-controlled measurement system

The WG concentrates on the cases where an ISP or a regulator runs the measurement system. However, we expect that LMAP functionality will also be used in the context of an end user-controlled measurement system. There are at least two ways this could happen (they have various pros and cons):

1. a user could somehow request the ISP- (or regulator-) run measurement system to test his/her line. The ISP (or regulator) Controller would then send an Instruction to the MA in the usual LMAP way. Note that a user can't directly initiate a Measurement Task on an ISP- (or regulator-) controlled MA.
2. a user could deploy their own measurement system, with their own MA, Controller and Collector. For example, the user could download all three functions onto the same user-owned end device; then the LMAP Control and Report protocols do not need to be used, but using LMAP's Information Model would still be beneficial. The MP could be in the home gateway or outside the home network; in the latter case the MP is highly likely to be

run by a different organisation, which raises extra privacy considerations.

In both cases there will be some way for the user to initiate the Measurement Task(s). The mechanism is out-of-scope of the LMAP WG, but could include the user clicking a button on a GUI or sending a text message. Presumably the user will also be able to see the Measurement Results, perhaps summarised on a webpage. It is suggested that these interfaces conform to the LMAP guidance on the privacy of the Measurement Results and Subscriber information.

## [6.](#) Details of the LMAP framework

This section contains a more detailed discussion of the four components of the LMAP framework.

### [6.1.](#) Measurement Agent (MA)

The Measurement Agent is the component that is responsible for

executing the Measurement Tasks. 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 measurement could make use of one or multiple Measurement Agents in a single measurement 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 Tasks) 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.

#### [6.1.1.](#) 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 Measurements 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.

#### 6.1.2. Measurement Agent embedded 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.

#### 6.1.3. 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. It allows the Measurement Agent to back off a potentially disruptive Active Measurements Task to avoid impacting the user. A Passive Measurements Task allows the Measurement Agent to gather data without



the overhead of Test Traffic (of interest to both the site user and network operator) as well as potentially provide a greater number of samples. 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).

#### [6.1.4.](#) 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 the MA itself. For example, if the MA 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 MA is to replicate the end user's network experience, it is important to understand the current practices.

#### [6.2.](#) Measurement Peer (MP)

A Measurement Peer is the other side of an Active Measurements Task – the target of Test Traffic from a Measurement Agent. The Measurement Peer 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), even another Measurement Agent in another home could perform as a Measurement Peer for a given Measurement Task. Particularly useful could be a MP that is well placed bandwidth-wise and can handle thousand of sessions of Test Traffic.

### [6.3.](#) Controller

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

For scaling purposes there may be 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.

### [6.4.](#) Collector

A Collector is responsible for receiving the Measurement Results from the Measurement Agent at the end of a Measurement Task. 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, perhaps regionally located. A large scale test making use of multiple Collectors would need to aggregate/consolidate their results for the complete picture.

## [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

Internet-Draft

LMAP Framework

October 2013

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 fool a MA into injecting 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 do not operate a correct Measurement Schedule.

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.

## [8.](#) Privacy Considerations for LMAP

Comment: It may be better to create a separate draft about 'LMAP threats and considerations' containing this section and perhaps the security section.

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, i.e. that privacy features are well-defined.

This section provides a set of privacy considerations for LMAP. This section benefits greatly from the timely publication of [[RFC6973](#)]. There are dependencies on the integrity of the LMAP security mechanisms, described in the Security Considerations section above.

We begin with a set of assumptions related to protecting the sensitive information of individuals and organizations participating in LMAP-orchestrated measurement and data collection.

#### [8.1.](#) Categories of Entities with Information of Interest

LMAP protocols need to protect the sensitive information of the following entities, including individuals and organizations who participate in measurement and collection of results.

- o Individual Internet Users: Persons who utilize Internet access services for communications tasks, according to the terms of service of a service agreement. Such persons may be a Service Subscriber, or have been given permission by the subscriber to use the service.
- o Internet Service Providers: Organizations who offer Internet access service subscriptions, and thus have access to sensitive information of Individuals who choose to use the service. These organizations desire to protect their subscribers and their own sensitive information which may be stored in the process of measurement and result collection.
- o Other LMAP system Operators: Organizations who operate measurement systems or participate in measurements in some way.

#### [8.2.](#) Examples of Sensitive Information

This section gives examples of sensitive information which may be measured or stored in a measurement system, and which is to be kept private by default in the LMAP core protocols.

Examples of Subscriber or authorized Internet User Sensitive Information:

- o IP address in use
- o Personal Identification (Real Name)
- o Location (street address, city)
- o Subscribed Service Parameters
- o Contents of Traffic (Activity, DNS queries, Destinations, Equipment types, Account info for other services, etc.)
- o Status as a study volunteer and Schedule of (Active) Measurement Tasks

Eardley, et al.

Expires April 24, 2014

[Page 25]

---

Internet-Draft

LMAP Framework

October 2013

Examples of Internet Service Provider Sensitive Information:

- o Measurement Device Identification (Equipment ID and IP address)
- o Measurement Instructions (choice of measurements)
- o Measurement Results (some may be shared, others may be private)
- o Measurement Schedule (exact times)
- o Network Topology (Locations, Connectivity, Redundancy)
- o Subscriber billing information, and any of the above Subscriber Information known to the provider.
- o Authentication credentials (e.g., certificates)

Other organizations will have some combination of the lists above.

### [8.3.](#) Key Distinction Between Active and Passive Measurement Tasks

For the purposes of this memo, we define Passive and Active Measurements Tasks as follows:

Passive: measurements conducted on Internet User traffic, such that sensitive information is present and stored in the measurement system (however briefly this storage may be).

Active: measurements conducted on traffic which serves only the purpose of measurement. Even if a user host generates active measurement traffic, there is significantly limited sensitive information present and stored in the measurement system compared to the passive case, as follows:

- o IP address in use
- o Status as a study volunteer and schedule of active tests

On the other hand, the sensitive information for an Internet Service Provider is the same whether active or passive measurements are used.

#### [8.4.](#) Communications Model (for Privacy)

This section briefly presents a set of communication models for LMAP. We assume that the Measurement Agent is located behind a NAT/Firewall, so it performs the role of Initiator for all communications.

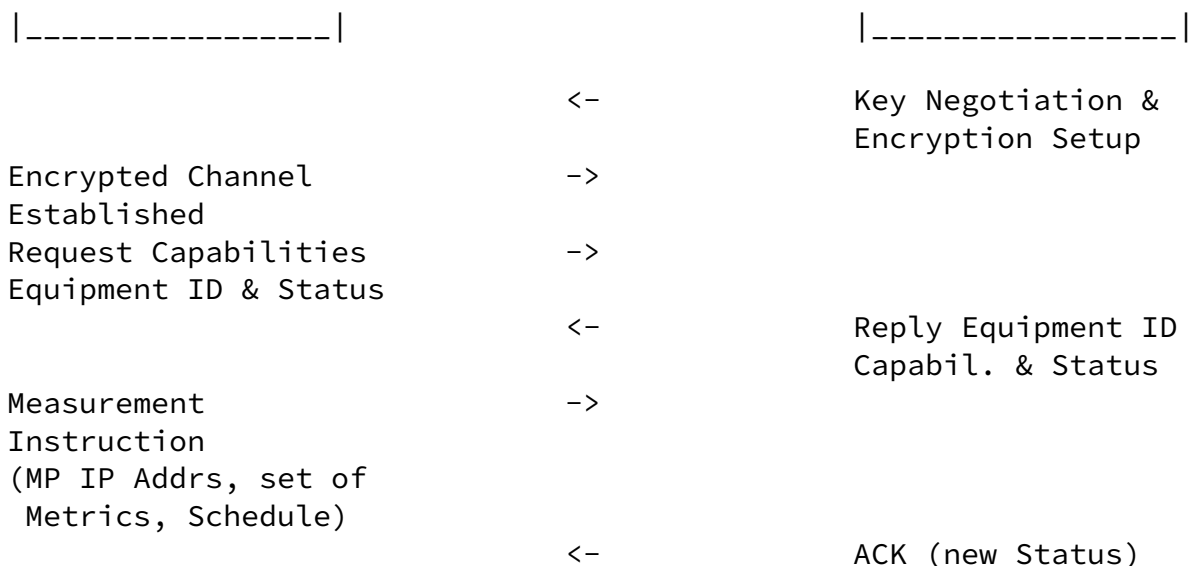
From a privacy perspective, all LMAP entities can be considered "observers" according to the definition in [[RFC6973](#)]. Their stored information potentially poses a threat to privacy, especially if one or more of these functional entities has been compromised.

Likewise, all devices on the paths used for control, reporting, and measurement are also observers. We note this in the figures below by identifying the possible presence of a NAT, which has additional significance to the protocols and direction of initiation.

##### [8.4.1.](#) Controller <-> Measurement Agent

The high-level communication model for interactions between the LMAP Controller and Measurement Agent is illustrated below. The primary purpose of this exchange is to authenticate and task a Measurement Agent with Measurement Instructions, which the Measurement Agent then acts on autonomously.



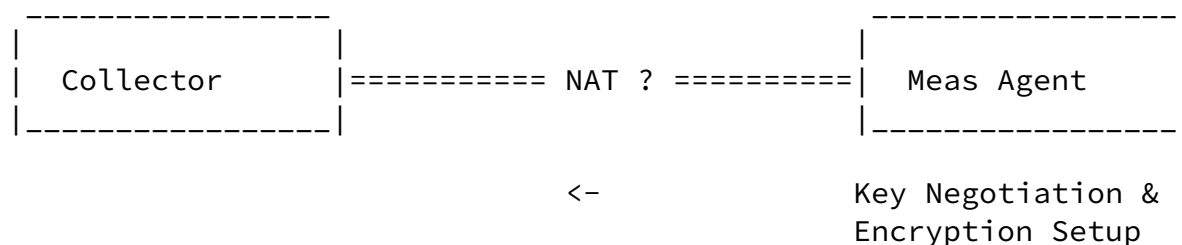


Primarily IP addresses and pseudonyms are exchanged first, then measurement-related information of interest such as the metrics, schedule, and IP addresses of measurement devices.

An organization operating the controller having no service relationship with the user who hosts the measurement agent *\*could\** gain real-name mapping to public IP address through user participation in an LMAP system.

#### [8.4.2.](#) Collector <-> Measurement Agent

The high-level communication model for interactions between the LMAP Measurement Agent and Collector is illustrated below. The primary purpose of this exchange is to authenticate and collect results from a Measurement Agent, which it has measured autonomously and stored.



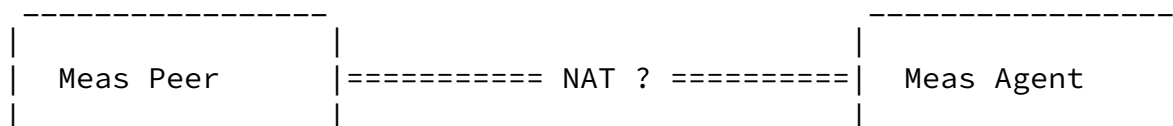
Encrypted Channel	->	
Established		
Request Capabilities?	->	
Equipment ID & Status		
	<-	Reply Equipment ID
		Capabil. & Status
	<-	Measurement Results
		(MP IP Addrs, set of
		Metrics, Values)
ACK	->	

Primarily IP addresses and pseudonyms are exchanged first, then measurement-related information of interest such as the metrics, schedule, results, and IP addresses of measurement devices.

An organization operating the collector having no service relationship with the user who hosts the measurement agent *\*could\** gain real-name mapping to public IP address through user participation in an LMAP system.

#### [8.4.3.](#) Active Measurement Peer <-> Measurement Agent

Although the specification of the mechanisms for measurement is beyond the LMAP scope, the high-level communications model below illustrates measurement information and results flowing between active measurement devices as a potential privacy issue. The primary purpose of this exchange is to execute measurements and store the results.



	<-	Key Negotiation &
		Encryption Setup
Encrypted Channel	->	
Established		
Announce Capabilities	->	
& Status		
	<-	Select Capabilities



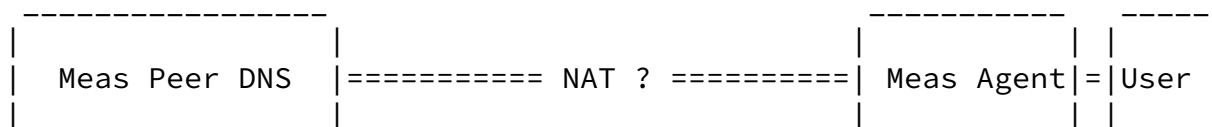
ACK	->	
	<-	Measurement Request (MA+MP IPAddr, set of Metrics, Schedule)
ACK	->	
Measurement Traffic (may/may not be encrypted)	<>	Measurement Traffic (may/may not be encrypted)
	<-	Stop Tests
Return Results (if applicable)	->	
	<-	ACK, Close

This exchange primarily exposes the IP addresses of measurement devices and the inference of measurement participation from such traffic. There may be information on key points in a service provider's network. There may also be access to measurement-related information of interest such as the metrics, schedule, and results.

If the measurement traffic is unencrypted, as found in many systems today, then both timing and limited results are open to observers.

#### [8.4.4.](#) Passive Measurement Peer <-> Measurement Agent

Although the specification of the mechanisms for measurement is beyond the LMAP scope, the high-level communications model below illustrates passive monitoring and measurement of information flowing between production network devices as a potential privacy issue. The primary purpose of this model is to illustrate collection of user information of interest with the Measurement Agent performing the monitoring and storage of the results. This particular exchange is for DNS Response Time, which most frequently uses UDP transport.



	<-	Name Resolution Req (MA+MP IPAddr, Desired Domain Name)
Return Record	->	

This exchange primarily exposes the IP addresses of measurement devices and the intent to communicate with, or access the services of "Domain Name". There may be information on key points in a service provider's network, such as the address of one of its DNS servers. The Measurement Agent may be embedded in the User host, or it may be located in another device capable of observing user traffic.

In principle, any of the Internet User information of interest (listed above) can be collected and stored in the passive monitoring scenario.

#### [8.4.5.](#) Result Storage and Reporting

Although the mechanisms for communicating results (beyond the initial Collector) are beyond the LMAP scope, there are potential privacy issues related to a single organization's storage and reporting of measurement results. Both storage and reporting functions can help to preserve privacy by implementing the mitigations described below.

### [8.5.](#) Threats

This section indicates how each of the threats described in [\[RFC6973\]](#) apply to the LMAP entities and their communication and storage of "information of interest".

#### [8.5.1.](#) Surveillance

[Section 5.1.1 of \[RFC6973\]](#) describes Surveillance as the "observation or monitoring of and individual's communications or activities."

All of passive measurement is surveillance, with inherent risks.

Active measurement methods which avoid periods of user transmission indirectly produce a record of times when a subscriber or authorized user has utilized their Internet access service.

Active measurements may also utilize and store a subscriber's currently assigned IP address when conducting measurements that are relevant to a specific subscriber. Since the measurements are time-stamped, the measurement results could provide a record of IP address assignments over time.

Internet-Draft

LMAP Framework

October 2013

Either of the above pieces of information could be useful in correlation and identification, described below.

#### [8.5.2.](#) Stored Data Compromise

[Section 5.1.2 of \[RFC6973\]](#) describes Stored Data Compromise as resulting from inadequate measures to secure stored data from unauthorized or inappropriate access.

The primary LMAP entity subject to compromise is the results storage which serves the Collector function (also applicable to temporary storage on the Collector itself). Extensive security and privacy threat mitigations are warranted for the storage system. Although the scope of its measurement and storage is smaller than the collector's, an individual Measurement Agent stores sensitive information temporarily, and also needs protections.

The LMAP Controller may have direct access to storage of Service Parameters, Subscriber information (location, billing, etc.), and other information which the controlling organization considers private, and needs protection in this case.

The communications between the local storage of the Collector and other storage facilities (possibly permanent mass storage), is beyond the scope of the LMAP work at this time, though this communications channel will certainly need protection as well as the mass storage.

#### [8.5.3.](#) Correlation and Identification

Sections [5.2.1](#) and [5.2.2](#) of [\[RFC6973\]](#) describes Correlation as combining various pieces of information to obtain desired characteristics of an individual, and Identification as using this process to infer identity.

The main risk is that the LMAP system could un-wittingly provide a key piece of the correlation chain, starting with an unknown Subscriber's IP address and another piece of information (e.g., Subscriber X utilized Internet access from 2000 to 2310 UTC, because the active measurements were deferred, or sent a name resolution for [www.example.com](#) at 2300 UTC).

#### [8.5.4.](#) Secondary Use and Disclosure

Sections [5.2.3](#) and [5.2.4](#) of [[RFC6973](#)] describes Secondary Use as unauthorized utilization of an individual's information for a purpose the individual did not intend, and Disclosure is when such information is revealed causing other's notions of the individual to change, or confidentiality to be violated.

The collection and reporting of passive traffic measurements is a form of secondary use, and subscribers' permission should be obtained before measurement. Although user traffic is only indirectly involved, active measurement results provide limited information about the subscriber and may constitute secondary use.

## [8.6](#). Mitigations

This section examines the mitigations listed in [section 6 of \[\[RFC6973\]\(#\)\]](#) and their applicability to LMAP systems. Note that each section in [[RFC6973](#)] identifies the threat categories that each technique mitigates.

### [8.6.1](#). Data Minimization

[Section 6.1 of \[\[RFC6973\]\(#\)\]](#) encourages collecting and storing the minimal information needed to perform a task.

There are two levels of information needed for LMAP results to be useful for a specific task: Network Operator and User troubleshooting, and General results reporting.

The minimal supporting information for general results is conducive to protection of sensitive information, as long as the results can be aggregated into large categories (e.g., the month of March, all subscribers West of the Mississippi River). In this case, all individual identifications (including IP address of the MA) can be excluded, and only the results applicable to the desired measurement path are provided.. However, this implies a filtering process to reduce the information fields allocated to this task, because greater detail was needed to conduct the measurements in the first place.

For a Network Operator and User troubleshooting a performance issue or failure, potentially all the network information (e.g., IP addresses, equipment IDs, location), measurement schedule, service configuration, measurement results and other information may assist

in the process. This includes the information needed to conduct the measurements, and represents a need where the maximum relevant information is desirable, therefore the greatest protections should be applied.

We note that a user may give temporary permission for passive measurements to enable detailed troubleshooting, but withhold permission for passive measurements in general. Here the greatest breadth of sensitive information is potentially exposed, and the maximum privacy protection must be provided.

For MAs with access to the sensitive information of users (e.g., within a home or a personal host/handset), it is desirable for the results collection to minimize the data reported, but also to balance this desire with the needs of troubleshooting when a service subscription exists between the user and organization operating the measurements.

For passive measurements where the MA reports flow information to the Collector, the Collector may perform pre-storage minimization and other mitigations (below) to help preserve privacy.

#### [8.6.2.](#) Anonymity

[Section 6.1.1 of \[RFC6973\]](#) describes a way in which anonymity is achieved: "there must exist a set of individuals that appear to have the same attributes as the individual", defined as an "anonymity set".

Experimental Methods for anonymization of user identifiable data applicable to passive measurement have been identified in [\[RFC6235\]](#). However, the findings of several of the same authors is that "there is increasing evidence that anonymization applied to network trace or flow data on its own is insufficient for many data protection applications as in [\[Bur10\]](#)."

Essentially, the details of passive flow measurements can only be accessed by closed organizations, and unknown injection attacks are always less expensive than the protections from them. However, some forms of summarized passive measurement may protect the user's

sensitive information sufficiently well, and so each metric must be evaluated in the light of privacy.

The methods in [\[RFC6235\]](#) could be applied more successfully in active measurement, where there are protections from injection attack. The successful attack would require breaking the integrity protection of the LMAP reporting protocol and injecting measurement results (known fingerprint, see [section 3.2 of \[RFC6973\]](#)) for inclusion with the shared and anonymized results, then fingerprinting those records to ascertain the anonymization process.

Beside anonymization of measured results for a specific user or provider, the value of sensitive information can be further diluted by summarizing the results over many individuals or areas served by the provider. There is an opportunity enabled by forming anonymity sets [\[RFC6973\]](#) based on the reference path measurement points in [\[I-D.ietf-ippm-lmap-path\]](#). 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 anonymization

applies to the Internet Service Provider, where their Internet gateway would be referred to as "mp190".

#### [8.6.3.](#) Pseudonymity

[Section 6.1.2 of \[RFC6973\]](#) indicates that pseudonyms, or nicknames, are a possible mitigation to revealing one's true identity, since there is no requirement to use real names in almost all protocols.

A pseudonym for a measurement device's IP address could be an LMAP-unique equipment ID. However, this would likely be a permanent handle for the device, and long-term use weakens a pseudonym's power to obscure identity.

#### [8.6.4.](#) Other Mitigations

Sections [6.2](#) and [6.3](#) of [\[RFC6973\]](#) describe User Participation and Security, respectively.

Where LMAP measurements involve devices on the Subscriber's premises or Subscriber-owned equipment, it is essential to secure the Subscriber's permission with regard to the specific information that

will be collected.

LMAP protocols, devices, and the information they store clearly need to be secure from unauthorized access. This is the hand-off between privacy and security considerations, found elsewhere in this memo.

#### [8.7.](#) The potential role of a Group-ID for privacy

A group identifier may be useful to help maintain privacy. Several MAs would share the same Group-ID. This has been suggested where the endusers are sensitive about privacy, for example mobile users do not want their location tracked. Some possibilities are discussed below.

A Group-ID might be used when Results are forwarded by a Collector to a third party. The measurement system operates using MA-IDs, however if Results are sent to a third party then Results from several MAs are aggregated together, in order to prevent the third party tracking them to an individual MA or enduser.

A Group-ID could be used for Reporting. The Controller's Instruction still refers to an MA using its MA-ID, but Results are reported to the Collector including a Group-ID but not an MA-ID. This might be useful where the measurement system is not run by the ISP (in the mobile example, the user clearly wants the operator track their location). The Group-ID needs to be sensible, for example MAs with the same broadband product (it is not sensible to aggregate Results from MAs on 2Mb/s and 300Mb/s lines). Note that:

- o A malicious MA could bias overall results by reporting more or less often than it is supposed to. The use of a Group-ID makes this harder for a Collector to detect.
- o An attacker is more likely to be able to break the MA-Collector communications, since it can only be secured at the group level,

for instance with a shared password. The attacker could then report false Results. Securing at the individual MA level intrinsically reveals the MA's identity

- o A malicious Collector can probably use other information to deaggregate the Results per MA, for example by tracking its IP address or analysing some per-MA 'fingerprint' information associated with the MA-Collector transmission protocol
- o A conspiratorial Controller could create a per-MA fingerprint (for example a unique set of parameters for the Measurement Tasks or simply a regular time at which the MA reports), which the Collector uses to identify the MA
- o A well-behaved Collector ensures that it only stores the Group-ID and throws away per-MA information. Then it cannot subsequently disaggregate Results per MA - such a breakdown might be requested by a government agency, an attacker or even by the measurement system itself (say after a change of company policy). In this case, the MA-Collector communication can be secured per MA, providing authentication is changed regularly and/or cannot be linked to the repository with the Measurement Results. In principle the scenario doesn't need a Group-ID to be defined for the Report Protocol - since the Collector can implement the Group-ID locally, after Results are reported.

A Group-ID could be used for Control as well as Reporting. The same Instruction is broadcast to all MAs, which check that they have a matching group-id before carrying out the Instruction. Notes:

- o The first three bullets above still apply

- o In addition, the Controller-MA communication is now also less secure
- o All the MAs with the same Group-ID probably need to be able to run exactly the same set of Measurement Methods.
- o At least at first glance, failure handling is harder. It is much less useful for the MA to inform the Controller that it cannot



understand or execute an Instruction - the Controller simply knows that one or more MAs, with a particular Group-ID, cannot understand or execute the Instruction. There also seems no point an MA reporting the Measurement Methods that it understands (which is intended to help a Controller that has forgotten an MA's capabilities, perhaps after a crash)

Conclusion - this topic needs more discussion. The use of per-MA authentication for security seems in tension with the use of Group-IDs for privacy.

## [9.](#) IANA Considerations

There are no IANA considerations in this memo.

## [10.](#) Acknowledgments

This document is a merger of three individual drafts: [draft-eardley-lmap-terminology-02](#), [draft-akhter-lmap-framework-00](#), and [draft-eardley-lmap-framework-02](#).

Thanks to numerous people for much discussion, directly and on the LMAP list. This document tries to capture the current conclusions. Thanks to Juergen Schoenwaelder for his detailed review of the terminology.

Philip Eardley, Trevor Burbridge and Marcelo Bagnulo 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.

## [11.](#) History

First WG version, copy of [draft-folks-lmap-framework-00](#).

### [11.1.](#) From -00 to -01

- o new sub-section of possible use of Group-IDs for privacy
- o tweak to definition of Control protocol
- o fix typo in figure in S5.4

## 12. Informative References

[I-D.linsner-lmap-use-cases]

Linsner, M., Eardley, P., and T. Burbridge, "Large-Scale Broadband Measurement Use Cases", [draft-linsner-lmap-use-cases-04](#) (work in progress), October 2013.

[lmap-yang]

, "A YANG Data Model for LMAP Measurement Agents", ,  
<<http://tools.ietf.org/html/draft-schoenw-lmap-yang>>.

[lmap-netconf]

, "Considerations on using NETCONF with LMAP Measurement Agents", ,  
<<http://tools.ietf.org/html/draft-schoenw-lmap-netconf>>.

[lmap-ipfix]

, "An LMAP application for IPFIX", ,  
<<http://tools.ietf.org/html/draft-bagnulo-lmap-ipfix>>.

[registry]

, , , , , "A registry for commonly used metrics. Independent registries", , <<http://tools.ietf.org/html/draft-bagnulo-ippm-new-registry-independent>>.

[RFC6241]

, , , , "Network Configuration Protocol (NETCONF)", ,  
<<http://tools.ietf.org/html/rfc6241>>.

[yang-api]

, "YANG-API Protocol", ,  
<<http://tools.ietf.org/html/rfc6241>>.

[schulzrinne]

, , , "Large-Scale Measurement of Broadband Performance: Use Cases, Architecture and Protocol Requirements", ,  
<<http://tools.ietf.org/html/draft-schulzrinne-lmap-requirements>>.

[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>>.
- [Bur10] Burkhardt, M., Schatzmann, D., Trammell, B., and E. Boschi, "The Role of Network Trace Anonymization Under Attack", January 2010.
- [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.
- [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.
- [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.
- [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.
- [RFC6235] Boschi, E. and B. Trammell, "IP Flow Anonymization Support", [RFC 6235](#), May 2011.
- [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.

Internet-Draft

LMAP Framework

October 2013

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

Paul Aitken

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

Email: [paitken@cisco.com](mailto:paitken@cisco.com)

Eardley, et al.

Expires April 24, 2014

[Page 39]

---

Internet-Draft

LMAP Framework

October 2013

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

Email: [aakhter@cisco.com](mailto:aakhter@cisco.com)

