

Network Working Group
Internet-Draft
Intended status: Informational
Expires: June 9, 2014

P. Eardley
BT
A. Morton
AT&T Labs
M. Bagnulo
UC3M
T. Burbridge
BT
P. Aitken
A. Akhter
Cisco Systems
December 6, 2013

A framework for large-scale measurement platforms (LMAP)
draft-ietf-lmap-framework-02

Abstract

Measuring broadband service on a large scale requires a description of the logical architecture and standardisation 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).

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 June 9, 2014.

Copyright Notice

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

Internet-Draft

LMAP Framework

December 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

1.	Introduction	3
2.	Outline of an LMAP-based measurement system	5
3.	Terminology	8
4.	Constraints	10
4.1.	Measurement system is under the direction of a single organisation	10
4.2.	Each MA may only have a single Controller at any point in time	11
5.	LMAP Protocol Model	11
5.1.	Bootstrapping process	12
5.2.	Control Protocol	14
5.3.	Starting and stopping Measurement Tasks	16
5.4.	Report Protocol	17
5.5.	Items beyond the scope of the LMAP Protocol Model	19
5.5.1.	User-controlled measurement system	20
6.	MA Deployment considerations	20
6.1.	Measurement Agent embedded in site gateway	21
6.2.	Measurement Agent embedded behind Site NAT /Firewall	21
6.3.	Measurement Agent in multi homed site	21
7.	Security considerations	22
8.	Privacy Considerations for LMAP	23
8.1.	Categories of Entities with Information of Interest	23
8.2.	Examples of Sensitive Information	24
8.3.	Key Distinction Between Active and Passive Measurement Tasks	25
8.4.	Privacy analysis of the Communications Models	26
8.4.1.	MA Bootstrapping and Registration	26
8.4.2.	Controller <-> Measurement Agent	27
8.4.3.	Collector <-> Measurement Agent	27
8.4.4.	Active Measurement Peer <-> Measurement Agent	28
8.4.5.	Passive Measurement Peer <-> Measurement Agent	29

8.4.6.	Result Storage and Reporting	29
8.5.	Threats	30
8.5.1.	Surveillance	30
8.5.2.	Stored Data Compromise	30
8.5.3.	Correlation and Identification	31

8.5.4.	Secondary Use and Disclosure	31
8.6.	Mitigations	31
8.6.1.	Data Minimization	32
8.6.2.	Anonymity	32
8.6.3.	Pseudonymity	33
8.6.4.	Other Mitigations	34
9.	IANA Considerations	34
10.	Acknowledgments	35
11.	History	35
11.1.	From -00 to -01	35
11.2.	From -01 to -02	35
12.	Informative References	36
	Authors' Addresses	37

[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 [\[I-D.ietf-lmap-use-cases\]](#). 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 gets instructions from the Controller initiates the measurement tasks, 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.

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.ietf-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.](#) 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 and we introduce the main terms needed for the LMAP framework. The new terms are capitalized. In the next section we provide a terminology section with a compilation of all the LMAP terms and their definition. The subsequent sections study the LMAP components in more detail.

A Measurement Task measures some performance or reliability Metric of interest. An Active Measurement Task involves either a Measurement Agent (MA) injecting Test Traffic into the network destined for a Measurement Peer (MP), 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 [[I-D.bagnulo-ippm-new-registry-independent](#)] 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 Active Measurement Traffic. The protocols used for the Active Measurement Traffic is out of the scope of the LMAP WG and falls within the scope of other 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 components in more detail; here we introduce them.

The Controller manages a MA by instructing it which Measurement Tasks it should perform and when. For example it may instruct a MA at a home gateway: "Run the 'download speed test' with the Measurement Peer 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 Measurement 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 Measurement Tasks, a Controller can initiate a one-off Measurement Task ("Do measurement 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 Measurement Task to diagnose a fault, then it should understand what problem the customer is experiencing and what Measurement Tasks have already been run. The Subscribers' service parameters are already gathered and stored by existing operations systems.

A Results Repository 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.

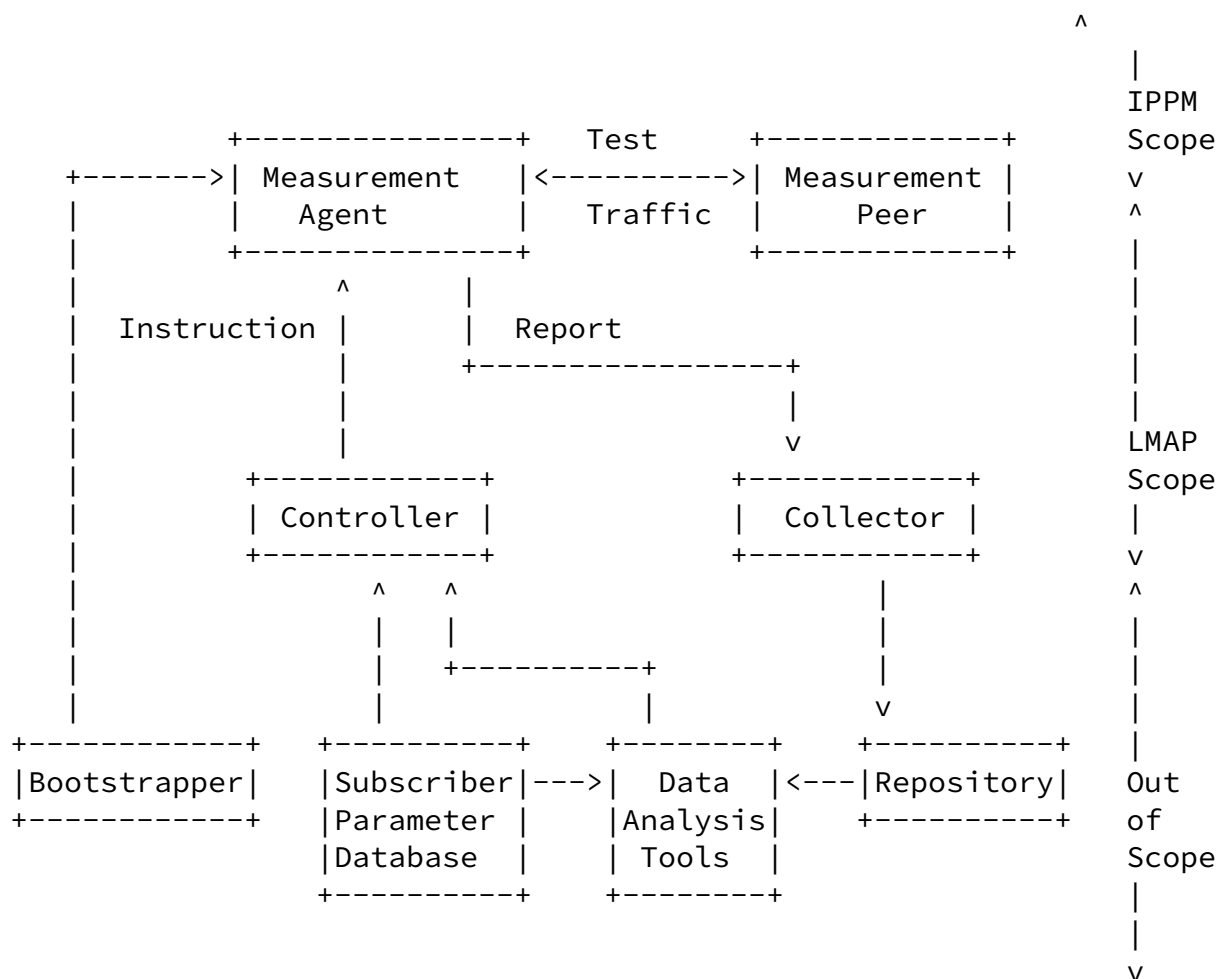


Figure 1: Schematic of main elements of an LMAP-based

(showing the elements in and out of the scope of the LMAP WG)

3. 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.

Capabilities Information: The list of the Measurement Methods that the MA can perform, plus information about the device hosting the MA (for example its interface type and speed and its IP address).

Channel: a schedule, a target and the associated security information for that target. In the case of a Report Channel it is a specific Report Schedule, a Collector and its associated security information.

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 Failure 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.

Failure Information: Information about the MA's failure to action or execute an Instruction, whether concerning Measurement Tasks or Reporting.

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.

Internet-Draft

LMAP Framework

December 2013

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). 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.

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, 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.

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

[4.](#) Constraints

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, 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.

Eardley, et al.

Expires June 9, 2014

[Page 10]

Internet-Draft

LMAP Framework

December 2013

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:

- 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 Active Measurement 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

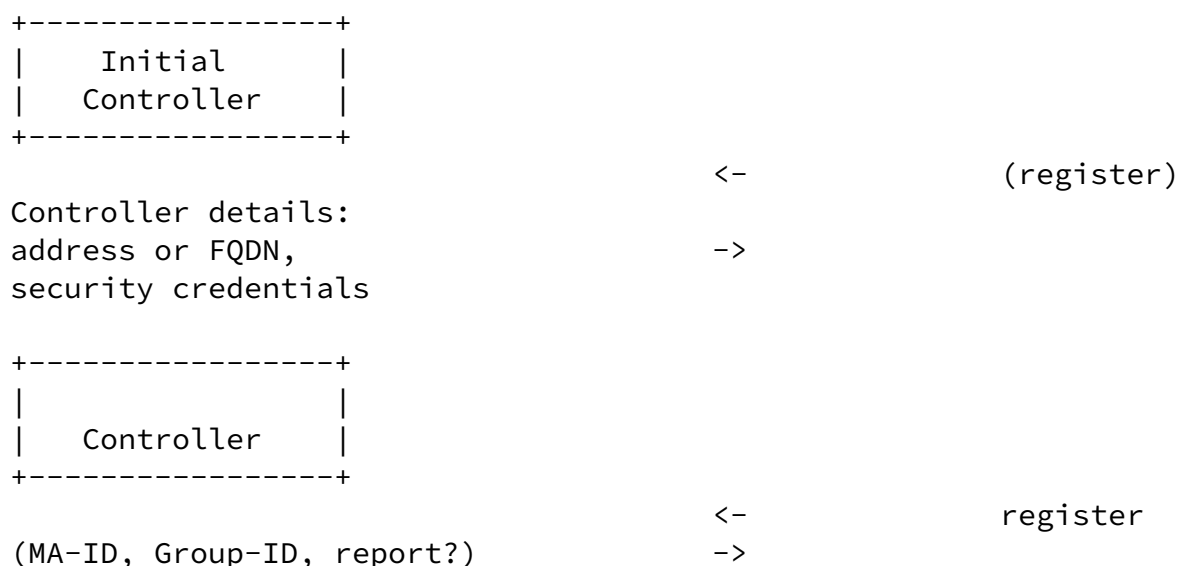
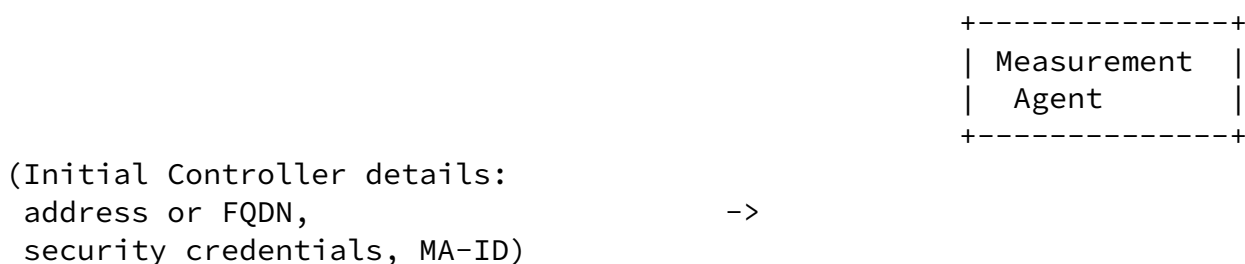
The Protocol Model is closely related to the Information Model [[I-D.burbridge-lmap-information-model](#)], which is the abstract definition of the information carried by the protocol model. The purpose of both is to provide a protocol and device independent view, which can be implemented via specific 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.

The diagrams show the flow of LMAP information, however there may need to be other protocol interactions. For example, typically the

MA is behind a NAT, so it needs to initiate communications in order that the Controller can communicate with it. The communications channel also needs to be secured before it is used. Another example is that the Collector may want to 'pull' Measurement Results from a MA.

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).



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; if the details of the initial Controller are statically configured; if the measurement system has specific Controllers for different devices types; or perhaps as a way of handling failure of the Controller.

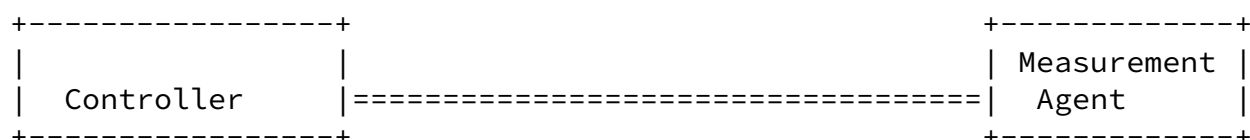
If the MA has not learnt its identifier (MA-ID) while bootstrapping, it will do so when the MA registers with the Controller; it may also be told a Group-ID and whether to include the MA-ID as well as the Group-ID in its Reports. A Group-ID would be shared by several MAs and could be useful for privacy reasons (for instance to hinder tracking of a mobile MA device). The MA may also tell the Controller its Capabilities (such as the Measurement Methods it can perform) (see next sub-section).

If the device with the MA re-boots, then the MA need to re-register, so that it can receive a new Instruction. To avoid a "mass calling event" after a widespread power restoration affecting many MAs, it is sensible for an MA to pause for a random delay (perhaps in the range of one minute) before re-registering.

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.

[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.



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

The Instruction contains:

- o what Measurement Tasks 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 measurement task if the end user is active')
- o when to do them: the Measurement Schedule details the timings of regular measurement tasks, one-off measurement tasks
- 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.

A new Instruction replaces (rather than adds to) those elements that it includes. For example, if the new Instruction includes (only) a Measurement Schedule, then that replaces the old Measurement Schedule but does not alter the configuration of the Measurement Tasks and Report Channels.

If the Instruction includes several Measurement Tasks, these could be scheduled to run at different times or possibly at the same time - some Tasks may be compatible, in that they do not affect each other's Results, whilst with others great care would need to be taken.

A Measurement Task may create more than one Measurement Result. For example, one Result could be reported periodically, whilst another could be an alarm that is reported immediately a the measured value of a Metric goes below a threshold.

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).

However, the Control protocol includes a Capabilities detection feature, through which 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 create a lot of traffic. See the

next section for more information on stopping Measurement Tasks. Note that if a Controller wants to permanently stop a Measurement Task, it should send a new Measurement Schedule, as suppression is intended to temporarily stop Tasks. The Controller can send an "un-suppress" message to indicate that the temporary problem is solved and Active Measurement Tasks can begin again.

The figure shows that the various messages are acknowledged, which means that they have been delivered successfully.

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. These could concern either the Measurement Tasks and Schedules, or the Reporting. In both cases there are two broad categories of failure. Firstly, the MA cannot action the Instruction (for example, it doesn't include a parameter that is mandatory for the requested Measurement Method; or it is missing details of the target Collector). Secondly, the MA cannot execute the Measurement Task or deliver the Report (for example, the MA unexpectedly has no spare CPU cycles; or the Collector is not responding). 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.

Finally, note that the MA doesn't do a 'safety check' with the Controller (that it should still continue with the requested Measurement Tasks) - nor does it inform the Controller about Measurement Tasks starting and stopping. 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

December 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. As an example, 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.

The Controller may want a MA to run the same Measurement Task indefinitely (for example, "run the 'upload speed' Measurement Task once an hour until further notice"). To avoid the MA generating traffic forever after a Controller has permanently failed, it is suggested that the Measurement Schedule includes a time limit ("run the 'upload speed' Measurement Task once an hour for the next 30 days") and that the Measurement Schedule is updated regularly (say, every 10 days).

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.

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

Eardley, et al.

Expires June 9, 2014

[Page 17]

Internet-Draft

LMAP Framework

December 2013

```
| Collector |=====| Agent |
+-----+                                     +-----+
```

```

                                     <-      Report:
                                     [MA-ID &/or Group-ID,
                                     Measurement Results,
                                     Measurement Task]
ACK                                ->
```

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)

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

Optionally, a Report is not sent when there are no Measurement Results.

In the initial LMAP Information Model and Report Protocol, for simplicity we assume that all Measurement Results are reported as-is, but allow extensibility so that a measurement system (or perhaps a

second phase of LMAP) could allow a MA to pre-process Measurement Results before it reports them. Potential examples of pre-processing by the MA are:

- o labelling, or perhaps not including, Measurement Results impacted by for instance cross-traffic or the MP being busy
- o detailing the start and end of suppression
- o filtering outlier Results
- o calculating some statistic like average (beyond that defined by the Measurement Task itself)

The measurement system may define what happens if a Collector unexpectedly does not hear from a MA. Possible solutions could include the ability for a Collector to 'pull' Measurement Results from a MA, or (after an out-of-scope indication from the Collector to the Controller) for the Controller to send a fresh Report Schedule to the MA. The measurement system also needs to consider carefully how to interpret missing Results; for example, if the missing Results are ignored and the lack of a Report is caused by its broadband being broken, then the estimate of overall performance, averaged across all MAs, would be too optimistic.

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.](#) MA Deployment considerations

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 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.](#) 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.

[6.2.](#) 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.

[6.3.](#) 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.). [\[RFC6419\]](#) 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.

[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. The measurement system must secure the various components of the system from unauthorised access or corruption.

We assume that each Measurement Agent will receive measurement tasks 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 fool a MA into injecting falsified data into the measurement platform or to corrupt the results of a real MA. The results must also be held and processed securely after collection and analysis.

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

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.

Internet-Draft

LMAP Framework

December 2013

- 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.

Although privacy is a protection extended to individuals, we include discussion of ISPs and other LMAP system operators in this section. These organizations have sensitive information involved in the LMAP system and revealed by measurements, and many of the same mitigations are applicable. Further, the ISPs store information on their subscribers beyond that used in the LMAP system (e.g., billing information), and there should be a benefit in considering all the needs and potential solutions coherently.

[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 Sub-IP layer addresses and names (e.g., MAC address, BS id, SSID)
- 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

Examples of Internet Service Provider Sensitive Information:

- o Measurement Device Identification (Equipment ID and IP address)

Eardley, et al.

Expires June 9, 2014

[Page 24]

Internet-Draft

LMAP Framework

December 2013

- 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. The LMAP system would not typically expose all of the information above, but could expose a combination of items which could be correlated with other pieces collected by an attacker (as discussed in the section on Threats below).

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

There are many possible definitions for the two main categories of measurement types, active and passive. For the purposes of this memo, we describe Passive and Active Measurements 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). We note that some authorities make a distinction on time of storage, and information that is kept only temporarily to perform a communications function is not subject to regulation (e.g., Active Queue Management, Deep Packet Inspection). Passive measurements could reveal all websites a subscriber visits and the applications and/or services they use.

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 about the Subscriber present and stored in the measurement system compared to the passive case, as follows:

- o IP address in use (and possibly Sub-IP addresses and names)
- o Status as a study volunteer 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 (e.g., measurement results).

Both Active and Passive measurements potentially expose the description of Internet Access service and specific service parameters, such as subscribed rate and type of access.

[8.4.](#) Privacy analysis of the Communications Models

This section examines each of the protocol exchanges described at a high level in [Section 5](#) and some example measurement tasks, and identifies specific sensitive information which must be secured during communication for each case. With the protocol-related sensitive information identified, we have can better consider the threats described in the following section.

From the privacy perspective, all entities participating in LMAP protocols 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.

[8.4.1.](#) MA Bootstrapping and Registration

[Section 5.1](#) provides the communication model for the Bootstrapping process.

Although the specification of mechanisms for Bootstrapping the MA are beyond the LMAP scope, designers should recognize that the Bootstrapping process is extremely powerful and could cause an MA to

join a new or different LMAP system with Control/Collection entities, or simply install new methods of measurement (e.g., a passive DNS Query collector). A Bootstrap attack could result in a breach of the LMAP system with significant sensitive information exposure depending on the capabilities of the MA, so sufficient security protections are warranted.

The Bootstrapping (or Registration) process provides sensitive information about the LMAP system and the organization that operates it, such as

- o Initial Controller IP address or FQDN
- o Assigned Controller IP address or FQDN
- o Security certificates and credentials

During the Bootstrap process (or Registration process that follows), the MA receives its MA-ID which is a persistent pseudonym for the subscriber in the case that the MA is located at a service

demarcation point. Thus, the MA-ID is considered sensitive information, because it could provide the link between subscriber identification and measurements or observations on traffic.

Also, the Bootstrap or Registration process could assign a Group-ID to the MA. The specific definition of information represented in a Group-ID is to be determined, but several examples are envisaged including use as a pseudonym for a set of subscribers, a class of service, an access technology, or other important categories. Assignment of a Group-ID enables anonymization sets to be formed on the basis of service type/grade/rates. Thus, the mapping between Group-ID and MA-ID is considered sensitive information.

8.4.2. Controller <-> Measurement Agent

The high-level communication model for interactions between the LMAP Controller and Measurement Agent is illustrated in [Section 5.2](#). 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 (MA-ID, Group-ID) are exchanged with a capability request, then measurement-related information of interest such as the parameters, schedule, metrics, and IP addresses of measurement devices. Thus, the measurement Instruction contains sensitive information which must be secured. For example, the fact that an ISP is running additional measurements beyond the set reported externally is sensitive information, as are the additional measurements themselves. The schedule/timing of specific measurements is also sensitive, because an attacker intending to bias the results without being detected can use this information to great advantage.

An organization operating the Controller having no service relationship with a user who hosts the measurement agent *could* gain real-name mapping to public IP address through user participation in an LMAP system (this applies to the Measurement Collection protocol, as well).

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

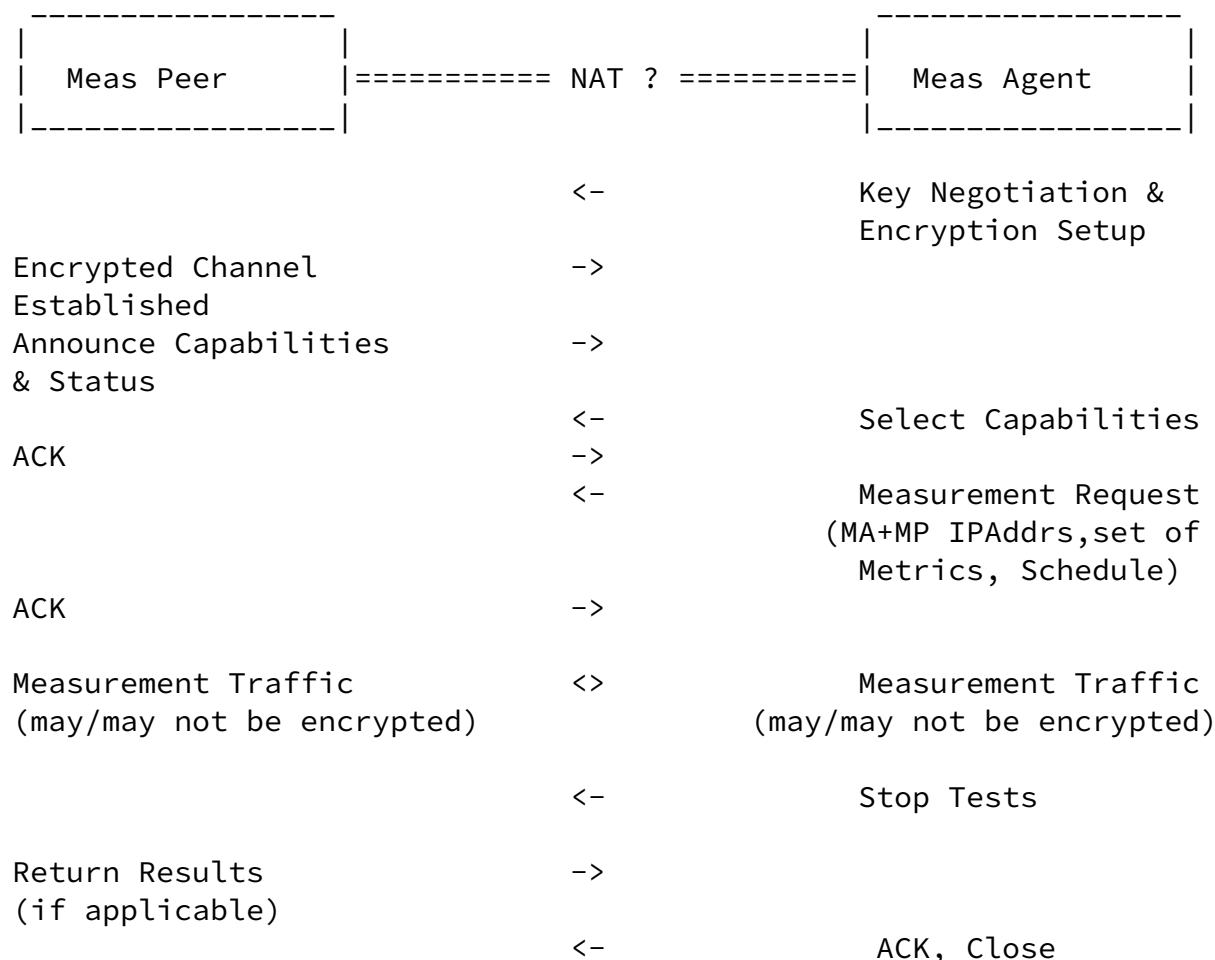
The high-level communication model for interactions between the LMAP Measurement Agent and Collector is illustrated in [Section 5.4](#). The primary purpose of this exchange is to authenticate and collect results from a Measurement Agent, which it has measured autonomously and stored.

Beyond the Controller-MA exchange, the new and highly-sensitive information exposed in the Collector-MA exchange is the measurement results. Organizations collecting LMAP measurements have the responsibility for Data Control. Thus, the results and other information communicated in the Collector protocol must be secured.

[8.4.4.](#) 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.

We note the potential for additional observers in the figures below by indicating the possible presence of a NAT, which has additional significance to the protocols and direction of initiation.



This exchange primarily exposes the IP addresses of measurement devices and the inference of measurement participation from such

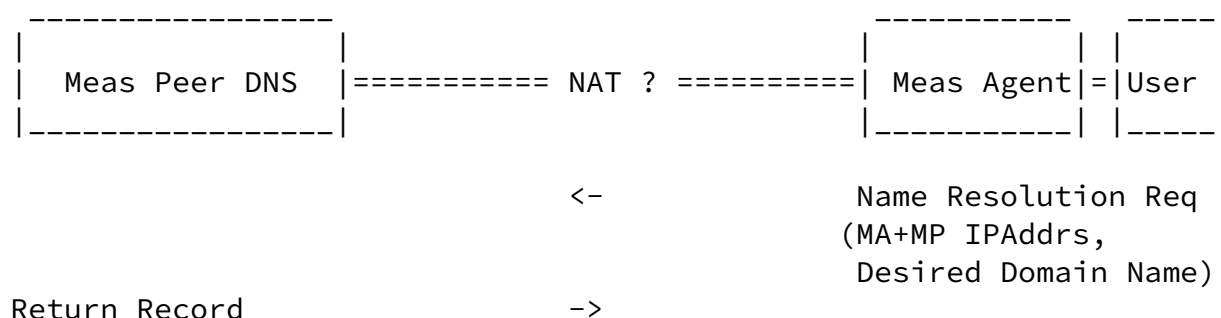
traffic. There may be sensitive information on key points in a service provider's network included. There may also be access to measurement-related information of interest such as the metrics, schedule, and intermediate results carried in the measurement packets (usually a set of timestamps).

If the measurement traffic is unencrypted, as found in many systems

today, then both timing and limited results are open to on-path observers, and this should be avoided when the degradation of secure measurement is minimal.

[8.4.5.](#) 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.



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 sensitive information of interest (listed above) can be collected and stored in the passive monitoring scenario. Thus, the LMAP Collection of passive measurements provides the additional sensitive information exposure to a Collection-path observer, and this information must be secured.

[8.4.6.](#) 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.

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. For LMAP systems this includes deleting or modifying collected measurement records, as well as data theft.

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.

Internet-Draft

LMAP Framework

December 2013

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 itself.

[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 and measured ISP's permission should be obtained before measurement. Although user traffic is only indirectly involved, active measurement results provide limited information about the subscriber/ISP and may constitute secondary use. Use of the measurements in unauthorized marketing campaigns would qualify as 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. The informed consent of the Subscriber (and, if different, the end user) is needed, including the specific purpose of the measurements. The approval process could involve showing the Subscriber their measured information and results before instituting periodic collection, or before all instances of collection, with the option to cancel collection temporarily or permanently.

It should also be clear who is legally responsible for data protection (privacy); in some jurisdictions this role is called the 'data controller'. It is good practice to time limit the storage of personal information.

Although the details of verification would be impenetrable to most subscribers, the MA could be architected as an "app" with open

source-code, pre-download and embedded terms of use and agreement on measurements, and protection from code modifications usually provided by the app-stores. Further, the app itself could provide data reduction and temporary storage mitigations as appropriate and certified through code review.

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. The Data Controller has the (legal) responsibility to maintain data protections described in the Subscriber's agreement and agreements with other organizations.

Another standard method for de-personalising data is to blur it by adding synthetic data, data-swapping, or perturbing the values in ways that can be reversed or corrected.

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 Burbidge 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

11.2. From -01 to -02

- o change to INFORMATIONAL track (previous version had typo'd Standards track)
- o new definitions for Capabilities Information and Failure Information
- o clarify that diagrams show LMAP-level information flows. Underlying protocol could do other interactions, eg to get through NAT or for Collector to pull a Report
- o add hint that after a re-boot should pause random time before re-register (to avoid mass calling event)
- o delete the open issue "what happens if a Controller fails" (normal methods can handle)
- o add some extra words about multiple Tasks in one Schedule

- o clarify that new Schedule replaces (rather than adds to) and old one. similarly for new configuration of Measurement Tasks or Report Channels.
- o clarify suppression is temporary stop; send a new Schedule to permanently stop Tasks
- o alter suppression so it is ACKed

- o add un-suppress message
- o expand the text on error reporting, to mention Reporting failures (as well as failures to action or execute Measurement Task & Schedule)
- o add some text about how to have Tasks running indefinitely
- o add that optionally a Report is not sent when there are no Measurement Results
- o add that a Measurement Task may create more than one Measurement Result
- o clarify /amend /expand that Reports include the "raw" Measurement Results - any pre-processing is left for lmap2.0
- o add some cautionary words about what if the Collector unexpectedly doesn't hear from a MA
- o add some extra words about the potential impact of Measurement Tasks
- o clarified various aspects of the privacy section
- o updated references
- o minor tweaks

12. Informative References

- [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.

Linsner, M., Eardley, P., and T. Burbridge, "Large-Scale Broadband Measurement Use Cases", [draft-ietf-lmap-use-cases-00](#) (work in progress), October 2013.

[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-homenet-arch]

Chown, T., Arkko, J., Brandt, A., Troan, O., and J. Weil, "IPv6 Home Networking Architecture Principles", [draft-ietf-homenet-arch-11](#) (work in progress), October 2013.

[RFC6419] Wasserman, M. and P. Seite, "Current Practices for Multiple-Interface Hosts", [RFC 6419](#), November 2011.

[RFC6887] Wing, D., Cheshire, S., Boucadair, M., Penno, R., and P. Selkirk, "Port Control Protocol (PCP)", [RFC 6887](#), April 2013.

[RFC5533] Nordmark, E. and M. Bagnulo, "Shim6: Level 3 Multihoming Shim Protocol for IPv6", [RFC 5533](#), June 2009.

[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-01](#) (work in progress), October 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.

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

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

Email: paitken@cisco.com

Internet-Draft

LMAP Framework

December 2013

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

Email: aakhter@cisco.com

Eardley, et al.

Expires June 9, 2014

[Page 39]