Network Working Group   T. Burbridge
Internet-Draft   P. Eardley
Intended status: Informational   British Telecom
Expires: April 24, 2014   M. Bagnulo
                                      Universidad Carlos III de Madrid
                                                       J. Schoenwaelder
                                                       Jacobs University
                                                       October 21, 2013

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

Abstract

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

Requirements Language

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

Status of this Memo

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

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

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

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

Copyright Notice

Table of Contents

1. Introduction

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

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

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

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

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

2. LMAP Information Model

2.1. Information Structure

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

   For clarity the information model is divided into six sections:

   1. Pre-Configuration Information. Information pre-configured on the
       Measurement Agent prior to any communication with other
       components of the LMAP architecture, specifically detailing how
       to register with a Controller

   2. Configuration Information. Information delivered to the MA on
       registration with a Controller or updated during a later
       communication, in particular detailing how to retrieve
       measurement and reporting instruction information from a
       Controller along with information specifically about the MA

   3. Instruction Information. Information that is received by the MA
       from the Controller pertaining to the measurement and reporting
       configuration. This includes measurement configuration, report
       channel configuration, measurement schedules and measurement
       suppression information

   4. Logging Information. Information transmitted from the MA to the
       Controller detailing the results of any configuration operations
       along with error and status information from the operation of the
       MA

   5. Status Information. Information on the general status and
       capabilities of the MA. For example, the set of measurements
       that are supported on the device

   6. Reporting Information. Information transmitted from the MA to
       the Collector including measurement results and the context in
       which they were conducted

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

to the Collector.

2.2. Pre-Configuration Information

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

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

Detail of the information model elements:

1. MA MAC: MAC Address

2. Configuration Channel: Channel

3. MA Certificate: Certificate (optional)

4. MA ID: random UUID (optional)

5. MA password: string (optional)

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

2.3. Configuration Information

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

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

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

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

Detail of the additional information model elements:

1. Measurement Agent ID: UUID

2. Measurement Group ID (optional): String

3. Report MA ID flag (optional): Boolean

4. Instruction Channel: Channel (DISCUSSION: shouldn't we split this
    into 4 different channels i.e. the Measurement Task Configuration
    channel, the Report Channel channel, the Measurement Schedules
    channel and the Measurement Suppression channel?)

5. Status Channel: Channel

6. Logging Channel: Channel

2.4. Instruction Information

The Instruction information model has four sub-elements:

1. Measurement Task Configurations: Set

2. Report Channels: Set

3. Measurement Schedules: Set

4. Measurement Suppression: Object

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

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

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

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

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

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

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

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

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

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

Definition of the information model elements:

1. Measurement Task Configurations: Set

    1. Measurement Task Configuration: Object

        1. Task Name (used for referral from the Measurement
            Schedules): String

        2. Registry Entry: URN

        3. Options: Set (optional)

            1. Interface name (reference by name to one of the
                Interfaces defined in the Status information): String

          4.   Measurement Cycle ID: String (optional)

    2. Report Channels: Set

       1. Report Channel: Channel

    3. Measurement Schedules: Set

       1. Measurement Schedule: Object

          1. Schedule Name: String

          2. Measurement Task Configuration Names (reference by Name
              to one of the measurement tasks defined in the
              Measurement Task Configuration set): List

             1. Task Name: String

          3. Report Channel Names (reference by Name to one of the
              measurement tasks defined in the Measurement Task
              Configuration set): Set

             1. Channel Name: String

          4. Measurement Timing: Timing

    4. Measurement Suppression: Object (optional)

       1. Start: datetime

       2. End: datetime

       3. Set of Measurement Task Configuration Names (optional -
          default all)

          1. Task Name: String

2.5. Logging Information

   The MA will report back success/failure and status information to the
   Controller. These messages will fall into a number of different
   categories:

   1. Success/failure messages in response to information updates from
      the Controller. For example:

      * "Report Channel 'hourly db" configured"

* "Measurement Schedule does not conform to schema, Row 211"

2. Status updates from the operation of the MA. For example:

* "out of memory: cannot record result"

* "Collector 'collector.example.com' not responding"

Each log message will have the following Information model elements:

1. Log Time: datetime

2. Log Event: Object

2.6. Status Information

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

MA Status information model elements:

1.   MA ID: String

2.   MA Device: String

3.   MA hardware: String (optional)

4.   MA firmware: String (optional)

5.   MA software: String (optional)

6.   MA Interfaces: set

1. If name: String

2. If type: String (one of eth, wlan, TBC)

3. If speed: Integer (expressed in Mbps)

4. Link Layer Address: String

5. IP address: Set

1. Protocol: String (one of v4, v6)

          2. Address: String

      6. Gateway: Set (optional)

          1. Protocol: String (one of v4, v6)

          2. Address: String

      7. DNS server: Set (optional)

          1. Protocol: String (one of v4, v6)

          2. Address: String

   7. Last Measurement: datetime

   8. Last Report: datetime

   9. Last Instruction: datetime

   10. Last Configuration: datetime

   11. Supported Measurements: Set

       1. Registry Entry: URN

       2. Version: String (optional)

2.7. Reporting Information

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

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

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

Information model elements:

1. Report Date: datetime

2. Measurement Agent ID: String (optional)

3. Measurement Group ID: String (optional)

4. MA Context: Set (optional)

    1. Context Item: Object

5. Measurement Task: Set

    1. Measurement Task Configuration: Object

    2. Result Headers: List

        1. Column Name: String

    3. Result Data: List

        1. Result Row: Object

            1. Measurement Time: datetime

            2. Cross-traffic: Integer (optional)

            3. Result Columns: List

                1. Column Data

2.8. Channels

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

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

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

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

Channel: Object

1. Channel Name (used for referral from other objects): String

2. Target: URL

3. Certificate: X.509 Certificate

4. Communication Timing: Timing

2.9. Timing Information

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

1. Periodic. Specifies a start, end and interval time in
    milliseconds

2. Calendar: Specifies a calendar based pattern - e.g. 22 minutes
    past each hour of the day on weekdays

3. One Off: A single instance occurring at a specific time

4. Immediate: Should occur as soon as possible

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

2.9.1. Periodic Timing

    Information model elements:

    1. 1. Timing Name: String

    2. 2. Start: datetime (optional)

    3. 3. End: datetime (optional)

    4. 4. Interval: Integer (in milliseconds)

    5. 5. Randomness: Timing Randomness (optional)

2.9.2. Calendar Timing

    Information model elements:

    1. Timing Name: String

    2. Start: datetime (optional)

    3. End: datetime (optional)

    4. Months: Set (optional - default [1-12])

        1. Month: Integer

    5. Weekdays: Set (optional - default [Mon-Sun])

        1. Weekday: String (one off Mon, Tue, Wed, Thu, Fri, Sat Sun)

    6. Days: Set (optional - default [1-31])

        1. Day: Integer

    7. Hours: Set (optional - default [1-24])

        1. Hour:Integer

    8. Minutes: Set (optional - default [1-60])

        1. Minute: Integer

    9. Seconds: Set (optional - default [1-60])

        1. Second: Integer

   10. Randomness: Timing Randomness (optional)

2.9.3. One-Off Timing

   Information model elements:

   1. Time: datetime

   2. Randomness: Timing Randomness (optional)

2.9.4. Immediate Timing

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

2.9.5. Timing Randomness

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

   Information model elements:

   1. Distribution: String

   2. Upper Cut: Integer (optional)

   3. Lower Cut: Integer (optional)

   4. Spread: Integer

3. IANA Considerations

   This document makes no request of IANA .

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

4. Security Considerations

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

5. Acknowledgements

6. Informative References

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

   [I-D.ietf-lmap-framework]
              Eardley, P., Morton, A., Bagnulo, M., Burbridge, T.,
              Aitken, P., and A. Akhter, "A framework for large-scale
              measurement platforms (LMAP)",
              draft-ietf-lmap-framework-00 (work in progress),
              October 2013.

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

   [RFC3444] Pras, A. and J. Schoenwaelder, "On the Difference between
              Information Models and Data Models", RFC 3444,
              January 2003.

Authors' Addresses

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

    Phone:
    Fax:
    Email:
    URI:

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

    Phone:
    Fax:
    Email:
    URI:

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

    Phone:
    Fax:
    Email:
    URI:

    Juergen Schoenwaelder
    Jacobs University

    Phone:
    Fax:
    Email:
    URI: