

Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: July 19, 2013

M. Bagnulo
UC3M
T. Burbridge
BT
S. Crawford
SamKnows
P. Eardley
BT
A. Morton
AT&T Labs
January 15, 2013

A registry for commonly used metrics. Independent registries
draft-bagnulo-ippm-new-registry-independent-00

Abstract

This document creates a registry for commonly used metrics, defines the rules for assignments in the new registry and performs initial allocations. This document proposes one particular registry structure with independent registries for each of the fields involved. A companion document draft-bagnulo-ippm-new-registry explores an alternative structure with a single registry with multiple sub-registries.

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 July 19, 2013.

Copyright Notice

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

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

Table of Contents

1. Introduction	3
2. The commonly used metrics registry	4
2.1. The metrics registry	4
2.2. The Scheduling registry	5
2.3. The Environment registry	5
2.4. The Output type registry	5
3. Initial assignment for the Scheduling registry	6
3.1. Common parameter definitions	6
3.2. Poisson scheduling	6
3.3. Periodic scheduling	7
3.4. Singleton scheduling	7
4. Initial assignments for the Output Type registry	7
4.1. Raw	7
4.2. Xth percentile interval	8
4.3. Xth percentile mean	8
5. Initial assignments for the Environment registry	8
5.1. Undefined	8
5.2. No cross traffic	9
6. Initial assignments for the Metric registry	10
6.1. Comment	10
6.2. UDP related metrics	10
6.2.1. UDP latency metric	11
6.2.2. UDP packet-loss metric	11
7. ICMP related metrics	12
7.1. ICMP packet-loss metric	12
8. DNS related metrics	12
8.1. DNS latency metric	13
9. Some examples of measurement plans	14
10. Security considerations	15
11. IANA Considerations	15
12. Acknowledgments	15
13. References	15
13.1. Normative References	15
13.2. Informative References	16
Authors' Addresses	16

1. Introduction

This document creates a registry for commonly used metrics. In order to do that, it creates a number of namespaces whose values will be recorded by the registry and will uniquely and precisely identify metrics.

The motivation for having such registry is to allow a controller to request a measurement agent to execute a measurement using a specific metric. Such request can be performed using any control protocol that refers to the value assigned to the specific metric in the registry. Similarly, the measurement agent can report the results of the measurement and by referring to the metric value it can unequivocally identify the metric that the results correspond to.

There was a previous attempt to define a metric registry RFC 4148 [RFC4148]. However, it was obsoleted by RFC 6248 [RFC6248] because it was "found to be insufficiently detailed to uniquely identify IPPM metrics... [there was too much] variability possible when characterizing a metric exactly" which led to the RFC4148 registry having "very few users, if any".

Our approach learns from this, by tightly defining each entry in the registry with only a few parameters open for each. The idea is that the entries in the registry represent different measurement tests, whilst the parameters set things like source and destination addresses that don't change the fundamental nature of the test. The downside of this approach is that it could result in an explosion in the number of entries in the registry. We believe that less is more in this context - it is better to have a reduced set of useful metrics rather than a large set of metrics with questionable usefulness. Therefore this document defines that the registry only includes commonly used metrics that are well defined; hence we require both specification required AND expert review policies for the assignment of values in the registry.

There are a couple of side benefits of having such registry. First the registry could serve as an inventory of useful and used metrics, that are normally supported by different implementations of measurement agents. Second, the results of the metrics would be comparable even if they are performed by different implementations and in different networks, as the metric is properly defined.

This version of the document defines a set of independent registries, that limits the explosion of registry entries by allowing arbitrary combinations of entries in the different entries. The downside is that the list of useful metrics is less defined, as any combination would be defined. Which approach is better is up for discussion.

The registry forms part of a Measurement Plan {do you prefer the term 'Characterization Plan', 'control framework' or 'test schedule'?}. It describes various factors that need to be set by the party controlling the measurements, for example: specific values for the parameters associated with the selected registry entry (for instance, source and destination addresses); and how often the measurement is made. The Measurement Plan might look something like: "Dear measurement agent: Please start test DNS(example.com) and RTT(server.com,150) every day at 2000 GMT. Run the DNS test 5 times and the RTT test 50 times. Do that when the network is idle. Generate both raw results and 99th percentile mean. Send measurement results to collector.com in IPFIX format". The Measurement Plan depends on the requirements of the controlling party. For instance the broadband consumer might want a one-off measurement made immediately to one specific server; a regulator might want the same measurement made once a day until further notice to the 'top 10' servers; whilst an operator might want a varying series of tests (some of which will be beyond those defined in the registry) as determined from time to time by their operational support system. While the registries defined in this document help to define the Measurement Plan its full specification falls outside the scope of this document.

2. The commonly used metrics registry

In this section we define the registry for commonly used metrics. It is composed by the following sub-registries:

- o Scheduling registry
- o Environment registry
- o Output-type registry
- o Metric registry

The rationale for the registry structure is to allow flexibility but yet precise definition of metrics. The metric registry defines the metric itself while the other registries define additional aspects that are needed for the measurement plan and that are needed to fully specify a measurement request from a controller to a measurement agent.

2.1. The metrics registry

Each entry for the metrics registry contain the following information:

- o Value: A text string that uniquely identifies the metric
- o Reference: The specification where the metric is defined

The policy for the assignments in the metric registry is both

specification required AND expert review. This means that in order to create an entry for the metric value a specification defining the metric is required and when that happens, the request for allocation will be reviewed by an expert.

The specification must define the input parameters for the metric as well as the output of the metric. The metric must be well defined, in the sense that two independent implementations must produce uniform and comparable results.

The expert review must make sure that the proposed metric is operationally useful. This means that the metric has proven to be useful in operational/real scenarios.

2.2. The Scheduling registry

Each entry for the scheduling registry contain the following information:

- o Value: The name of the scheduling
- o Reference: the specification where the scheduling is defined

The scheduling defines the scheduling strategy for the metric. Simplest is Singleton scheduling, where an atomic measurement is made. Other strategies make a series of atomic measurements in a "sample" or "stream", with the schedule defining the timing between each distinct measurement. Each atomic measurement could consist of sending a single packet (such as a DNS request) or sending several packets (for example a webpage). A scheduling strategy requires input parameter(s). Assignment in this registry follows the specification required policy.

2.3. The Environment registry

Each entry for the environment registry contain the following information:

- o Value: The name of the environment
- o Reference: the specification where the environment is defined

The environment defines the conditions where the metric is expected to be used. It does not define the metric itself, but the context where the metric is executed. Assignment in this registry follows the specification required policy.

2.4. The Output type registry

Each entry for the output type registry contain the following information:

- o Value: The name of the output type
- o Reference: the specification where the output type is defined

The output type define the type of output that the metric produces. It can be the raw results or it can be some form of statistic. Assignment in this registry follows the specification required policy. The specification of the output type must define the format of the output.

3. Initial assignment for the Scheduling registry

3.1. Common parameter definitions

Although each IPPM RFC defines individual parameters and uses them consistently, the parameter names are not completely consistent across the RFC set. For example, the variable "dT" is used in several different ways. This memo uses one set of parameter names, and the reader is cautioned to map the names according to their definitions.

We define some parameters that are used by several types of scheduling:

- o T0: time to begin a test
 - o Tf: time to end a test
- T0 and Tf are both in seconds and use the date (yyyy-mm-dd) and NTP 64 bit timestamp. T0 includes any control handshaking before the test stream or singleton. Tf is the time the last test data is sent.

As a result, we have:

- o Time when test devices may close the test socket: Tf + Waiting Time (the time to wait before declaring a packet lost is fixed for each metric)
- o Total duration of the test: Tf - T0 + Waiting Time

3.2. Poisson scheduling

The values for this entry are as follows:

- o Value: Poisson
- o Reference: draft-bagnulo-ippm-new-registry

The Poisson scheduling is defined in section 11.1.1 of RFC 2330 [RFC2330] and needs input parameters:

- o T0 and Tf: defined above
- o lambda: the parameter defining the Poisson distribution. Lambda is the mean number of distinct measurements per second in the sample.

3.3. Periodic scheduling

The values for this entry are as follows:

- o Value: Periodic
- o Reference: draft-bagnulo-ippm-new-registry

The Periodic sampling is defined in RFC 3432 [RFC3432]. The additional input parameters for the metric required by Periodic scheduling are:

- o T0 and Tf: defined above
 - * Note that with Periodic sampling, T0 MUST NOT be strictly periodic with other tests of the same type. RFC 3432 [RFC3432] requires randomized start times and describes one way to accomplish this. Also, the duration of the test MUST be limited.
- o incT: the time in seconds between one distinct event and the next, where events typically result in repeating singleton measurements of various types (illustrated below).
 - * for a periodic stream this is the time between packets in the sample, first bit to first bit
 - * for measurements on a process this is the time between the first packets of the process, for example first bit to first bit of the SYN in a TCP 3-way handshake

3.4. Singleton scheduling

The values for this entry are as follows:

- o Value: singleton
- o Reference: draft-bagnulo-ippm-new-registry

The singleton scheduling covers the case when an atomic metric is performed as per RFC 2330 [RFC2330]. The additional input parameter for the metric required by Singleton scheduling is:

- o T0: defined above

4. Initial assignments for the Output Type registry

4.1. Raw

The values for this entry are as follows:

- o Value: Raw
- o Reference: draft-bagnulo-ippm-new-registry

The results of the metric are delivered in the exact way they are produced by the measurements without any further processing or filtering.

4.2. Xth percentile interval

The values for this entry are as follows:

- o Value: Xth-percentile
- o Reference: draft-bagnulo-ippm-new-registry

The additional input parameter for the metric is:

- o X: the percentile (e.g, if the X input parameter is 99, then the output will be the 99th percentile interval.)

The output when using this Output type will be a couple of values, expressed in the same units as the raw output, that is the Xth percentile interval, as defined in section 1.3 of RFC 2330 [RFC2330].

4.3. Xth percentile mean

The values for this entry are as follows:

- o Value: Xth-percentile-mean
- o Reference: draft-bagnulo-ippm-new-registry

The additional input parameter for the metric is

- o X: the percentile (e.g, if the X input parameter is 99, then the output will be the 99th percentile mean.)

The output when using this Output type will be a single value, expressed in the same units as the raw output, that is the mean of the sample only considering the values contained in the Xth percentile interval, as defined in RFC 2330 [RFC2330].

5. Initial assignments for the Environment registry

5.1. Undefined

The values for this entry are as follows:

- o Value: Undefined
- o Reference: draft-bagnulo-ippm-new-registry

The undefined environment is the case where no additional environment settings are defined to perform the metric.

5.2. No cross traffic

The values for this entry are as follows:

- o Value: No-cross-traffic
- o Reference: draft-bagnulo-ippm-new-registry

It is often important that there is no other traffic than the one generated by the measurement itself while doing the measurement. The reasons for this are two-folded, first, it is sometimes important that the traffic created by the measurement doesn't impact the experience of the users of the measured resource. Second it is sometimes important that no other traffic interferes with the measurement. This can be ensured by checking that the level of user traffic is either zero or low enough to be confident that it won't impact or be impacted by the measurement.

The "No cross traffic" condition is satisfied when, during the 5 seconds preceding measurement of the metric:

- o the level of traffic flowing through the interface that will be used to send measurement packets in either direction is less than a threshold value of 1% of the line rate of the aforementioned interface.

The "cross traffic" measurement is made at the interface, associated with the measurement agent, that user traffic flows across. For example, if the probe is attached to the home gateway, then the interface is the service demarcation point where the subscriber connects their private equipment or network to the subscribed service.

Note that the No-cross traffic condition is defined only for the link directly attached to the measurement agent initiating the measurement. There is nothing mentioned about cross traffic on other parts of the path used by measurement packets. In the case the bottleneck of the path is other link than the one directly attached to the device running the measurement agent, it may affect and be affected by the measurement even if the No cross traffic as defined here holds.

DISCUSSION

- o It is not clear we need a registry for this. If the only thing we are going to define is the No cross traffic condition, we can simply set it as an input parameter in each metric.
- o clarify whether traffic for each direction is less than threshold, or the sum

- o current SamKnows probes measure cross-traffic before the measurement of the metric. Another approach would be to measure cross-traffic during the time the metric is measured. Or a hybrid approach. These would either be separate environment entries, or parameterise the existing one.
- o current SamKnows probes define a fixed threshold. it could be a parameter
- o could ignore broadcast traffic (think SamKnows includes)
- o It would be possible to define this a bit more precisely as follows:
 - * The "No cross-traffic" condition is defined for active measurements. The measurement agent runs in a device that has one or more interfaces. In active measurements, the measurement agent sends one or more packets. Lets call if0 the interface through with the packets resulting from the measurement are sent through. The no cross traffic condition is fulfilled when during the 5 seconds prior sending each of the packets of the measurement:
 - + The traffic incoming through if0 that does not belong to the measurement is lower than 1% of the line rate of if0
 - + The traffic coming through the rest of the interfaces towards if0 is less than 1% of the line rate of if0.

6. Initial assignments for the Metric registry

6.1. Comment

Need to work through that we only define T0 and Tf (and not T, dT).

6.2. UDP related metrics

RFC 2681 [RFC2681] defines a Round-trip delay metric and RFC 6673 [RFC6673] defines a Round-trip packet loss metric. We build on these two metrics by specifying several of the open parameters to precisely define several metrics for measuring UDP latency and packet loss. All the UDP related metrics defined in this section use the following:

P-Type:

- o IPv4 header values:
 - * DSCP: set to 0
 - * TTL set to 255
 - * Protocol: Set to 17 (UDP)
- o UDP header values:
 - * Checksum: the checksum must be calculated

- o Payload
 - * Sequence number: 8-byte integer
 - * Timestamp: 8 byte integer. Expressed as 64-bit NTP timestamp as per section 6 of RFC 5905 [RFC5905]
 - * No padding

Timeout: 3 seconds

6.2.1. UDP latency metric

We define the UDP latency metric as follows:

- o Value: UDP_Latency
- o Reference: draft-bagnulo-ippm-new-registry

The methodology for this metric is defined as Type-P-Round-trip-Delay- in RFC 2681 [RFC2681] using the P-Type and Timeout defined above.

The input parameters for this metric are:

- o Source IP Address
- o Destination IP Address
- o Source UDP port
- o Destination UDP port
- o Time

The output of this metric is the couple of values formed by the timestamp of the sent packet and the time when the echo was received. They are expressed in seconds and use the date (yyyy-mm-dd) and NTP 64 bit timestamp

6.2.2. UDP packet-loss metric

We define the UDP packet-loss metric as follows:

- o Value: UDP_packet_loss
- o Reference: draft-bagnulo-ippm-new-registry

This metric is defined as Type-P-Round-trip-Loss in RFC 6673 [RFC6673] using the P-Type and Timeout defined above.

The input parameters for this metric are:

- o Source IP Address
- o Destination IP Address
- o Source UDP port
- o Destination UDP port
- o Time T

The output of this metric is a single value 0 (packet was lost) or 1 (packet has arrived before timeout)

7. ICMP related metrics

RFC 6673 [RFC6673] defines a Round-trip packet loss metric. We build on that metrics by specifying several of the open parameters to precisely define a metric for measuring ICMP packet loss. The ICMP related metric defined in this document use the following:

P-Type:

- o IPv4 header values:
 - * DSCP: set to 0
 - * TTL set to 255
 - * Protocol: Set to 1 (ICMP)
- o ICMP header values:
 - * Type: 8 (Echo request)
 - * Code: 0

Observation: reply packets will contain an ICMP type of 0 Echo reply.

Timeout: 3 seconds

7.1. ICMP packet-loss metric

We define the ICMP packet-loss metric as follows:

- o Value: ICMP_Packet_Loss
- o Reference: draft-bagnulo-ippm-new-registry

This metric is defined as Type-P-Round-trip-Loss in RFC 6673 [RFC6673] using the P-Type and Timeout defined above.

The input parameters for this metric are:

- o Source IP Address
- o Destination IP Address
- o Time T

The output of this metric is a single value 0 (packet was lost) or 1 (packet has arrived before timeout)

8. DNS related metrics

RFC 2681 [RFC2681] defines a Round-trip delay metric. We build on that metric by specifying several of the open parameters to precisely define a metric for measuring DNS latency. The metric uses the following parameters:

P-Type:

- o IPv4 header values:
 - * DSCP: set to 0
 - * TTL set to 255
 - * Protocol: Set to 17 (UDP)
- o UDP header values:
 - * Source port: 53
 - * Destination port: 53
 - * Checksum: the checksum must be calculated
- o Payload: The payload contains a DNS message as defined in RFC 1035 [RFC1035] with the following values:
 - * The DNS header section contains:
 - + QR: set to 0 (Query)
 - + OPCODE: set to 0 (standard query)
 - + AA: not set
 - + TC: not set
 - + RD: set to one (recursion desired)
 - + RA: not set
 - + RCODE: not set
 - + QDCOUNT: set to one (only one entry)
 - + ANCOUNT: not set
 - + NSCOUNT: not set
 - + ARCOUNT: not set
 - * The Question section contains:
 - + QNAME: the FQDN provided as input for the test
 - + QTYPE: the query type provided as input for the test
 - + QCLASS: set to IN
 - * The other sections do not contain any Resource Records.

Observation: reply packets will contain a DNS response and may contain RRs.

Timeout: 3 seconds

8.1. DNS latency metric

We define the DNS latency metric as follows:

- o Value: DNS_Latency
- o Reference: draft-bagnulo-ippm-new-registry

The methodology for this metric is defined as Type-P-Round-trip-Delay in RFC 2681 [RFC2681] using the P-Type and Timeout defined above.

The input parameters for this metric are:

- o Source IP Address
- o Destination IP Address (the address of the DNS server to be tested)

- o QTYPE: A RR
- o FQDN: a valid FQDN that will be queried for.
- o Time T

The output of this metric is the timestamp when the packet was sent and the delay that it took to receive a response. Please note that any DNS response is valid, including no records in the answer. (Should we be more explicit about what is the output when there is no reply packet received?)

9. Some examples of measurement plans

A measurement plan will be characterized by the following tuple: (Metric, environment, scheduling, output format). We will next present some measurement plans that are currently used.

A measurement plan for measuring the 99th percentile interval of the UDP latency without cross traffics, using a Poisson stream with rate 1 pkts/sec, stating at time T0 and ending at Tf seconds, between source IP address IPs and source port Ps and destination IP address IPd and destination port Pd would be expressed as:

```
(UDP_Latency(IPs,Ps,IPd,Pd), No-cross-traffic, Poisson(T0,Tf,1),  
Xth-percentile(99))
```

A measurement plan for measuring the UDP packet loss ration without cross traffics, using a Poisson stream with rate 1 pkts/sec, stating at time T0 and ending at Tf seconds, between source IP address IPs and source port Ps and destination IP address IPd and destination port Pd would be expressed as:

```
(UDP_Packet_Loss(IPs,Ps,IPd,Pd), No-cross-traffic,  
Poisson(T0,Tf,1), Xth-percentile-mean(100))
```

A measurement plan for measuring the ICMP packet loss ratio, using a Periodic stream s second between packets, stating at time T0 and ending at Tf seconds, between source IP address IPs and destination IP address IPd would be expressed as:

```
(ICMP_Packet_Loss(IPs,IPd), Undefined, Periodic(T0,Tf,s), Xth-  
percentile-mean(100))
```

A measurement plan for measuring the DNS latency for resolving FQDN foo.com between a resolver in IP address IPs and a server with address IPd at time T would be expressed as:

```
(DNS_Latency(IPs,IPd,foo.com), Undefined, Singleton(T), raw)
```

10. Security considerations

TBD

11. IANA Considerations

TBD

12. Acknowledgments

We would like to thank Henning Schulzrinne for many constructive comments and input on early versions of this document.

13. References

13.1. Normative References

- [RFC2330] Paxson, V., Almes, G., Mahdavi, J., and M. Mathis, "Framework for IP Performance Metrics", RFC 2330, May 1998.
- [RFC3432] Raisanen, V., Grotefeld, G., and A. Morton, "Network performance measurement with periodic streams", RFC 3432, November 2002.
- [RFC2681] Almes, G., Kalidindi, S., and M. Zekauskas, "A Round-trip Delay Metric for IPPM", RFC 2681, September 1999.
- [RFC6673] Morton, A., "Round-Trip Packet Loss Metrics", RFC 6673, August 2012.
- [RFC1035] Mockapetris, P., "Domain names - implementation and specification", STD 13, RFC 1035, November 1987.
- [RFC5905] Mills, D., Martin, J., Burbank, J., and W. Kasch, "Network Time Protocol Version 4: Protocol and Algorithms Specification", RFC 5905, June 2010.
- [RFC2679] Almes, G., Kalidindi, S., and M. Zekauskas, "A One-way Delay Metric for IPPM", RFC 2679, September 1999.
- [RFC2680] Almes, G., Kalidindi, S., and M. Zekauskas, "A One-way Packet Loss Metric for IPPM", RFC 2680, September 1999.
- [RFC3393] Demichelis, C. and P. Chimento, "IP Packet Delay Variation

Metric for IP Performance Metrics (IPPM)", RFC 3393, November 2002.

[RFC5481] Morton, A. and B. Claise, "Packet Delay Variation Applicability Statement", RFC 5481, March 2009.

13.2. Informative References

[RFC4148] Stephan, E., "IP Performance Metrics (IPPM) Metrics Registry", BCP 108, RFC 4148, August 2005.

[RFC6248] Morton, A., "RFC 4148 and the IP Performance Metrics (IPPM) Registry of Metrics Are Obsolete", RFC 6248, April 2011.

Authors' Addresses

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

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

Trevor Burbridge
British Telecom
Adastral Park, Martlesham Heath
Ipswich
ENGLAND

Email: trevor.burbridge@bt.com

Sam Crawford
SamKnows

Email: sam@samknows.com

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

Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: August 25, 2013

M. Bagnulo
UC3M
B. Trammell
ETH Zurich
February 21, 2013

An LMAP application for IPFIX
draft-bagnulo-lmap-ipfix-01

Abstract

This document explores the possibility of using IPFIX to report test results from a Measurement Agent to a Collector, in the context of a large measurement platform.

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 August 25, 2013.

Copyright Notice

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

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

Table of Contents

1. Introduction	3
1.1. A quick introduction to IPFIX	3
1.2. Applying IPFIX to LMAP	4
2. Using IPFIX to report test results	5
3. Example: UDP latency test	7
4. Example: UDP latency test with Options	8
5. What standardization is needed for this?	10
6. Security considerations	10
7. IANA Considerations	11
8. Acknowledgements	11
9. References	11
9.1. Normative References	11
9.2. Informative References	12
Authors' Addresses	12

1. Introduction

A Large-scale Measurement Platform (LMP) is composed by the following fundamental elements: a set of Measurement Agents (MAs), one or more Controllers and one or more Collectors. There may be additional elements in any given such of these platforms, but these three elements are present in all of them. The MAs are pieces of code that run in specialized hardware (hardware probes) or in general purpose devices such as PCs, laptops or mobile phones (software probes). The MA run the tests against other MAs distributed across the Internet. Typically most of the MAs are located in end user networks and a few MAs are located deep into the ISP network, and typically tests are executed from the MAs in the periphery towards MAs located in the core. The Controller is the element that controls the MAs and informs the MAs about what tests to do and when to do them. The protocol between the Controller and the MA is called the Control protocol. After performing the tests, the MAs send the data about the results of the tests performed to the Collector. The protocol used to report test result data from the MA to the Collector is called the Report protocol. In this document we explore the possibility of using IPFIX [I-D.ietf-ipfix-protocol-rfc5101bis] as a Report protocol for large scale measurement platforms.

1.1. A quick introduction to IPFIX

IPFIX [I-D.ietf-ipfix-protocol-rfc5101bis] is a unidirectional, transport-independent export protocol for binary data records, with a focus on network measurement and operations applications. The structure of the data records is described in-band by Templates, which refer to Information Elements (IEs) from a common information model managed by IANA [ipfix-iana]. The basic IEs cover most Layer 3 and Layer 4 measurement needs, and the information model can be extended [I-D.ietf-ipfix-ie-doctors] as well as supplemented by private IEs.

IPFIX organizes data records into Messages. A Message is a sequence of Sets preceded by a Message Header which, among other things, includes an Observation Domain ID (roughly, identifying where the records in the Message were measured) and an Export Time (when the Message was originally sent).

A Set contains Records preceded by a Set Header, which contains a Set ID identifying the type of the records the Set contains. Template Sets, identified by a special Set ID, contain Templates, which are sequences of IE identifiers and lengths; these define the fields of the records they describe. A Template's ID matches the Set ID of the Sets containing records described by the Template.

On-wire data structures in IPFIX are fully discussed in section 3 of [I-D.ietf-ipfix-protocol-rfc5101bis].

Since many records may be described by a single Template, IPFIX's data representation is more efficient than those based on inline record structures (e.g. XML, JSON). Additionally, this arrangement implies that a device that only needs to export one or two fixed-length record types can implement IPFIX with minimal code supporting fixed message and set lengths with fixed-length templates.

IPFIX also supports a feature called Options Templates. An Options Template allows a data record to be scoped to a set of values of particular IEs (called its Scope). For example, a set of test parameters could be scoped to a test identifier IE, and that test identifier exported in a record together with the results. This mechanism allows more efficient data export, as explored in Section 4 below; more information is available in [RFC5473].

1.2. Applying IPFIX to LMAP

In IPFIX terminology [RFC5470], the MA encompasses both the Metering Process (MP) and the Exporting Process (EP), while the Collector is the Collecting Process (CP). IPFIX is used between the EP/MA and the Collector/CP. We propose LMA as an application of IPFIX per [I-D.ietf-ipfix-ie-doctors].

Some considerations about the use of IPFIX for LMP:

- o Separation between Control and Report Protocols: Within a single measurement platform, different protocols can be used for Control and Report, though they must share a common vocabulary representing the measurements to be performed. In particular, if a platform implements IPFIX as a Report protocol, it must implement a different protocol (e.g. NETCONF or other) as a Control protocol.
- o Report protocol diversity: Some platforms may use IPFIX as a Report protocol, while other platforms may decide to use other protocols (e.g. the Broadband forum architecture may decide to use a different one). We believe that it is important to support this protocol diversity. A key element to support such diversity is an independent metric registry (see [I-D.bagnulo-ippm-new-registry-independent]) where values for metric identifiers are recorded independently of the Control and/or Report protocol is used. This affects how we use IPFIX as a Report protocol, as presented in this document.
- o Minimal IPFIX implementation: The unidirectional nature of the protocol and simple wire format make minimal implementations of Exporting Processes possible. These minimal implementations are well suited to small-scale MAs (such as a mobile app or a process

running in a home router). These only need to know about the specific Templates supporting the metric(s) to be reported.

2. Using IPFIX to report test results

In order to use IPFIX to report test results from the MA to the Collector, we need first to understand what information needs to be conveyed. The information transmitted by the MA to the Collector when reporting test(s) results is the following:

- o Information about the MA: in particular a MA identifier
- o Information about the time of the report: when the report was sent (not necessarily when the test was performed)
- o Information describing the test. This includes:
 - * An identifier of the metric used for the test (see the Metric registry of [I-D.bagnulo-ippm-new-registry-independent])
 - * An identifier of the scheduling strategy used to perform the test (see the Scheduling registry of [I-D.bagnulo-ippm-new-registry-independent]) and potential input parameters for the schedule, such as the rate.
 - * An identifier of the output format, (see the Output Type registry of [I-D.bagnulo-ippm-new-registry-independent])
 - * An identifier of the environment, notably, if cross traffic was or not present during the execution of the test. (see the Environment registry of [I-D.bagnulo-ippm-new-registry-independent])
 - * The input parameters for the test, such as source IP address, destination IP address, source and destination ports and so on.
- o Information describing the test results. This widely varies with each test, but can include time each packet was sent and received, number of sent and lost packets or other information.

We next explore how we can encode this information in IPFIX.

In order to convey test information using IPFIX we will naturally use the IPFIX message format and we will define a Template describing the records containing the test result data. We will re-use as many already defined Information Elements (IEs) as possible and we will identify new IEs that are needed.

Part of the information can be conveyed using the fields in the IPFIX header, namely:

- o Information about the MA: In order to convey the MA identifier we can use the Observation Domain field present in the IPFIX header. This would allow to have up to 2^{32} MA, which seems sufficient.
- o Information about the time of the report: The IPFIX header contains an Export Time field that can be used to convey this information.

The information describing the test is included in a Template set that contains multiple IEs for each of the different pieces of information we need to convey. This includes:

- o An identifier of the metric used for the test. In order to convey that we need to define a new IE, let's call it `metricIdentifier`. The values for this element will be the values registered in the Metric registry of [I-D.bagnulo-ippm-new-registry-independent].
- o An identifier of the scheduling strategy used to perform the test. Again, this will be a new IE, called `testSchedule` and its values will be the values defined in the Scheduling registry of [I-D.bagnulo-ippm-new-registry-independent]. The potential input parameters for the schedule, such as the rate, we probably need a new IE for each of these. Usual scheduling distributions only require a rate, so we can define a new IE called `scheduleRate` which value will contain the rate for the requested distribution.
 - * NOTE: The distribution in some cases could be extracted from the results, for example, if the results contain each packet sent, it would be easy to spot a periodic scheduling. Probably not so obvious for the Poisson one. Maybe this would be an optional element to be carried when it is not possible to extract it from the test results.
- o An identifier of the output format. A new IE `outputType` is needed for this and it would take values out of the ones in the Output Type registry of [I-D.bagnulo-ippm-new-registry-independent]. Some of the output formats require an additional input, like the percentile used to trim the outliers when performing means. There are two approaches here. One approach is that the the Output Type registry creates different entries for the different percentiles, which would result in more entries in the Output Type registry (e.g. one entry for the 95th percentile mean and another one for the 90th percentile mean). This may cause an increase number of entries in the Output Type registry, but since there are not too many usual values, it is likely to be manageable. The other approach is to define an additional IE, for instance, the percentile IE that will have the values for the different percentiles used in the output.
- o An identifier of the environment, notably, if cross traffic was or not present during the execution of the test. Again, a new IE is needed for this `testEnvironment`. It will take values of the the Environment registry of [I-D.bagnulo-ippm-new-registry-independent].
- o The input parameters for the test. Most of these can be expressed using existing IEs, such as `sourceIPv4Address`, `destinationIPv4Address`, etc.

Information describing the test results. This widely varies with each test, but can include time each packet was sent and received, number of sent and lost packets or other information. Again most of

these can be expressed using existent IEs, and some new ones can be defined if needed for a particular test.

3. Example: UDP latency test

Let's consider the example of UDP latency. Suppose a MA wants to report the results of a UDP latency test, performed from its own IP address (e.g. 192.0.2.1) to a destination IP address (e.g. 203.0.113.1), using source port 23677 and destination port 34567. The test is performed using a periodic scheduling with a rate of 1 packet per second during 3 seconds and starts at 10:00 CEST. The test was performed without cross-traffic and the output type is raw.

The Template for this would be:

```
metricIdentifier
testSchedule
scheduleRate
outputType
testEnvironment
sourceIPv4Address
destinationIPv4Address
sourceTransportPort
destinationTransportPort
flowStartMilliseconds
flowEndMilliseconds
```

The data set following this template for the example would be:

```
metricIdentifier = UDP_Latency as per
[I-D.bagnulo-ippm-new-registry-independent]
testSchedule = Periodic as per
[I-D.bagnulo-ippm-new-registry-independent]
scheduleRate = 1
outputType = Raw as per
[I-D.bagnulo-ippm-new-registry-independent]
testEnvironment = No-cross-traffic as per
[I-D.bagnulo-ippm-new-registry-independent]
sourceIPv4Address = 192.0.2.1
destinationIPv4Address = 203.0.113.1
sourceTransportPort = 23677
destinationTransportPort = 34567
flowStartMilliseconds = 08:00:00.000 UTC
flowEndMilliseconds = 08:00:00.001 UTC
-----
metricIdentifier = UDP_Latency as per
[I-D.bagnulo-ippm-new-registry-independent]
```



```
testSchedule = Periodic as per
[I-D.bagnulo-ippm-new-registry-independent]
scheduleRate = 1
outputType = Raw as per
[I-D.bagnulo-ippm-new-registry-independent]
testEnvironment = No-cross-traffic as per
[I-D.bagnulo-ippm-new-registry-independent]
sourceIPv4Address = 192.0.2.1
destinationIPv4Address = 203.0.113.1
sourceTransportPort = 23677
destinationTransportPort = 34567
flowStartMilliseconds = 08:00:01.000 UTC
flowEndMilliseconds = 08:00:01.002 UTC
-----
metricIdentifier = UDP_Latency as per
[I-D.bagnulo-ippm-new-registry-independent]
testSchedule = Periodic as per
[I-D.bagnulo-ippm-new-registry-independent]
scheduleRate = 1
outputType = Raw as per
[I-D.bagnulo-ippm-new-registry-independent]
testEnvironment = No-cross-traffic as per
[I-D.bagnulo-ippm-new-registry-independent]
sourceIPv4Address = 192.0.2.1
destinationIPv4Address = 203.0.113.1
sourceTransportPort = 23677
destinationTransportPort = 34567
flowStartMilliseconds = 08:00:02.000 UTC
flowEndMilliseconds = 08:00:02.001 UTC
-----
```

4. Example: UDP latency test with Options

In the previous example, the test description is exported together with the results in the record. If a particular set of test parameters will be repeated often by a given MA, the common properties can be grouped into an Options record, described by an Options Template and identified by a new Information Element, with Data Records referring back to this identifier.

In this case, two templates are used: an Options Template to

The Options Template would be:

```
testParametersId {scope}
metricIdentifier
```

```

testSchedule
scheduleRate
outputType
testEnvironment
sourceIPv4Address
destinationIPv4Address
sourceTransportPort
destinationTransportPort

```

The Template for each Data Record carrying results would be:

```

testParametersId {scope}
flowStartMilliseconds
flowEndMilliseconds

```

The data set carrying the common properties would be:

```

testParametersId = 1
metricIdentifier = UDP_Latency as per
[I-D.bagnulo-ippm-new-registry-independent]
testSchedule = Periodic as per
[I-D.bagnulo-ippm-new-registry-independent]
scheduleRate = 1
outputType = Raw as per
[I-D.bagnulo-ippm-new-registry-independent]
testEnvironment = No-cross-traffic as per
[I-D.bagnulo-ippm-new-registry-independent]
sourceIPv4Address = 192.0.2.1
destinationIPv4Address = 203.0.113.1
sourceTransportPort = 23677
destinationTransportPort = 34567
-----

```

And the data set carrying results would be:

```

testParametersId = 1
flowStartMilliseconds = 08:00:00.000 UTC
flowEndMilliseconds = 08:00:00.001 UTC
-----
testParametersId = 1
flowStartMilliseconds = 08:00:01.000 UTC
flowEndMilliseconds = 08:00:01.002 UTC
-----
testParametersId = 1
flowStartMilliseconds = 08:00:02.000 UTC
flowEndMilliseconds = 08:00:02.001 UTC
-----

```

This approach sacrifices some complexity at the MA (which must assign testParametersIds and use multiple Templates) and the collector (which must track testParametersId of each set of parameters to

reassemble "complete" results) to gain export efficiency. A quantitative measurement of efficiency gains and tradeoffs for a set of specified result records will follow in a future version of this draft.

5. What standardization is needed for this?

So, in order to enable the use of IPFIX for LMP, the following pieces of standardization would be required.

- o The definition of the metric registry. This is not specific for IPFIX as any other Report protocol is likely to require this, but having an independent registry enables multiple report protocols.
- o The definition of new IEs. Some of them are identified above, some other are likely to be needed as well.
- o The definition of the Templates sets for each of the tests to be performed. This is necessary to have a defined Template that different vendors can implement and can use the IPFIX format in the wire, but they don't need to fully implement IPFIX parsing to read arbitrary Template sets, just the ones associated with the relevant metrics.

6. Security considerations

The security requirements for the protocol between the MA and the collector have been identified in [I-D.eardley-lmap-framework] and in [I-D.schulzrinne-lmap-requirements]. The identified requirements are:

- o Mutual authentication and authorization between the MA and the collector. This means that the collector must be able to verify the identity of the MA and to also verify that the MA is authorized to feed data into the collector and that the MA must be able to verify the identity of the collector and recognize it as a valid collector for the data it is reporting.
- o The information flowing between the MA and the collector must be confidential.
- o The integrity of the information flowing from the MA and the collector must be protected.

Not surprisingly these are exactly the same requirements imposed to the design of the IPFIX protocol, in particular for the flow of data between the EP and the CP. As described in the security considerations of IPFIX [I-D.ietf-ipfix-protocol-rfc5101bis], IPFIX address these requirements by imposing the use of TLS or DTLS with mutual authentication through certificates. The authorization relies on having a list of authorized MAs in the collector and a list of collectors in the MAs, identified by information in the Distinguished

Name and/or Common Name of their certificate. Current IPFIX specifications and implementations already support TLS and DTLS and this covers the aforementioned requirements. We are aware that some of the current platforms use ssh as a transport protocol between the MAs and the collector. Using ssh allow avoiding the use of certificates, but may result in a more complex key management (which may not be an issue in certain deployments). We believe it would be possible to define an ssh transport for IPFIX if deemed necessary.

IPFIX recommends the use DNS-IDs in the certificates, which applies to EPs and CPs with relatively static addressing. This is probably not a good fit for MAs, since they are likely to have a dynamic address. In this draft we have proposed to use the Observation domain as identifier for the MAs. While the Observation domain must not be globally unique within IPFIX, it would be possible to make it so in a particular measurement platform. The Observation Domain Identifier could then appear in the Common Name of the certificate in some form. Additionally, access control in very large deployments could rely not on identifying specific MAs, but on ensuring that a peer MA or collector had a certificate signed by one of a set of specified authorized issuers.

7. IANA Considerations

TBD

8. Acknowledgements

We would like to thank Sam Crawford and Al Morton for input on early discussions for this draft.

9. References

9.1. Normative References

- [I-D.ietf-ipfix-protocol-rfc5101bis]
Claise, B. and B. Trammell, "Specification of the IP Flow Information eXport (IPFIX) Protocol for the Exchange of Flow Information", draft-ietf-ipfix-protocol-rfc5101bis-06 (work in progress), February 2013.
- [RFC5470] Sadasivan, G., Brownlee, N., Claise, B., and J. Quittek, "Architecture for IP Flow Information Export", RFC 5470, March 2009.

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

[ipfix-iana]
Internet Assigned Numbers Authority, "IP Flow Information
Export (IPFIX) Entities", IANA IPFIX Registry ,
February 2013.

9.2. Informative References

[RFC5473] Boschi, E., Mark, L., and B. Claise, "Reducing Redundancy
in IP Flow Information Export (IPFIX) and Packet Sampling
(PSAMP) Reports", RFC 5473, March 2009.

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

[I-D.eardley-lmap-framework]
Eardley, P., Burbridge, T., and A. Morton, "A framework
for large-scale measurements",
draft-eardley-lmap-framework-00 (work in progress),
February 2013.

[I-D.schulzrinne-lmap-requirements]
Schulzrinne, H., Johnston, W., and J. Miller, "Large-Scale
Measurement of Broadband Performance: Use Cases,
Architecture and Protocol Requirements",
draft-schulzrinne-lmap-requirements-00 (work in progress),
September 2012.

Authors' Addresses

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

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

Brian Trammell
Swiss Federal Institute of Technology Zurich
Gloriastrasse 35
8092 Zurich
Switzerland

Email: trammell@tik.ee.ethz.ch

Network Working Group
Internet-Draft
Intended status: Informational
Expires: August 22, 2013

M. Boucadair
C. Jacquenet
France Telecom
February 18, 2013

Large scale Measurement of Access network Performance (LMAP):
Requirements and Issues from a Network Provider Perspective
draft-boucadair-lmap-considerations-00

Abstract

This document raises several points related to the ongoing LMAP (Large scale Measurement of Access network Performance) effort. The goal is to contribute to define a scope for LMAP and its expected contribution.

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 August 22, 2013.

Copyright Notice

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

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

Table of Contents

1. Introduction	3
2. Discussion	4
2.1. Service-Specific Measurement	4
2.2. Distorting Measurement Results	5
2.3. On the Impact of Policies	5
2.4. Classes of Service	5
2.5. Pending Questions	6
3. Security Considerations	7
4. IANA Considerations	7
5. Acknowledgments	7
6. Informative References	7
Authors' Addresses	8

1. Introduction

Service Assurance & Fulfilment is a critical component in the service management environment. Within ISP organizations, dedicated organizational and functional structures are implemented to efficiently monitor and assess the overall quality of deployed services and also the service quality as perceived by end-users.

As such, appropriate actions can be taken to solve encountered problems and put any disrupted service back to normal operation. Various tools (e.g., probes, reporting tools, etc.) are deployed to continuously provide feedback on the status of running services and notify managers about operational issues.

For the sake of efficient day-to-day operations, an ISP should implement the Service Fulfilment functions that are responsible for checking if the services delivered to the users are consistent with what has been subscribed and possibly negotiated. These functions may also be used as inputs to Service Assurance related functions.

The ISP should be able to continuously (preferably in real-time) measure and control the level of quality associated to the services delivered to its customers. Indeed, network anomalies such as node outage, link failures, routing disruption and the subsequent overall service performance degradation should be dynamically reported to appropriate management structures (like a Network Operations Center).

Ideally, any issue should be solved (or at least detected and handled as quick as possible) before receiving the complaints from customers. Improvement of current practices should be investigated to enhance the quality of experience as perceived by end-users and also to speed up repair processes whenever a network or service anomaly is detected.

Within this context, the introduction of a high level of automation in the global service delivery and operation chains is promising. This does not mean zero-fault networking: automation is rather meant to optimize communication between the different actors of the service delivery chain and also to guarantee the overall consistency between the different management tools.

Customers should have the ability to check the fulfilment of the Connectivity Provisioning Profile (CPP, [I-D.boucadair-connectivity-provisioning-profile]) they have subscribed to (and possibly negotiated with the service provider). They could thus evaluate how the Service Provider has delivered the service as a function of what has been defined in the service agreement. Customer- or service-specific indicators and related

performance metrics should be accessed by customers so that they can appreciate the level of quality associated to the services they have subscribed to. These data should be updated on a regular basis to adequately reflect the actual status of any service. These indicators (including a combination thereof) should be described and listed in the agreement (see Section 2.13 of [I-D.boucadair-connectivity-provisioning-profile]).

The Large scale Measurement of Access network Performance (LMAP) effort can be defined as a tool supported by the Service Assurance functional block provided to customers to assess whether the services they have subscribed comply with what has been defined in the service level agreement (including the technical parameters exposed in a CPP template, for example).

As discussed in [I-D.boucadair-connectivity-provisioning-profile], performance metrics are not the only relevant indicators to characterize the connectivity service delivered to the customer; other important technical clauses (e.g., reachability scope, traffic conformance, availability, etc.) need also to be taken into account.

Providing customers with tools that can help them better characterize the level of quality associated to the delivery of any service (or a combination thereof) they have subscribed to is likely to enhance their overall quality of experience. As a consequence, such tools would also optimize the overall efficiency of service operation (e.g., by reducing the number of calls placed to online support whenever a problem is pro-actively reported to the customer).

This document discusses several questions to be considered when designing such tools.

This document makes use of the terms defined in [I-D.morton-ippm-lmap-path].

2. Discussion

2.1. Service-Specific Measurement

Various service offerings (e.g., IPTV, VoD, Internet, VoIP, etc.) can be delivered to the same customer. All these services rely upon devices that are involved the forwarding of the corresponding service-specific traffic.

These services are not restricted to the basic IP connectivity service but also include advanced features. The technical clauses that document the IP connectivity service component of these services

may vary one from the other (e.g., a global reachability can be provided for the Internet service while IP connectivity service is restricted to the first SBE/DBE (Session Border Element/Data Border Element) for VoIP services).

Furthermore, some of these services may be delivered over dedicated "virtual" channels (e.g., distinct VCs or addresses can be used for each service).

Assessing whether the delivered service complies with what has been subscribed by the customer or not suggests that measurement actions should be specific to the communication facilities (forwarding paths, virtual channels, tunnels, etc.) used to deliver the service to the customer.

2.2. Distorting Measurement Results

Some services may rely on several components provided by distinct administrative entities. For instance, the DNS service may not be provided by the same operator that provides the IP connectivity. The level of quality associated to the delivery of a service may therefore be affected (e.g., because DNS resolution takes longer than expected) even if traffic performance clauses are honored by the network provider.

The LMAP system should be designed to accommodate such deployment scenario.

2.3. On the Impact of Policies

Issues can be experienced when a customer tries to reach a subset of destinations. These issues may not be necessarily due to performance degradation in the local network but to some policies enforced in the destination networks, at the risk of being unable to deliver the service to some networks (e.g., some government contents cannot be accessed from some networks, because of a security policy enforced by the government).

The measurement system should be designed to accommodate such contexts.

2.4. Classes of Service

Prioritization is used to deliver some services; as such, measurements should be bound to the QoS class used for a given service.

In some networks, DSCP marking inheritance mechanisms are used to

make sure customers cannot injects traffic that belongs to an unauthorized or unsupported class of service. The proposed measurement framework should be designed to handle such designs.

2.5. Pending Questions

Additional considerations should be taken into account as per the following questions:

- Q1: How to determine the measurement scope? How to characterize a measurement scope?
- Q2: Should inter-domain measurement be in the scope?
- Q3: If so, which inter-domain paths should be used to conduct measurement campaigns? Paths used for measurement may not be those used to forward service data.
- Q4: Which metrics to use? How contributing agents negotiate the metric to be used? What measurement methodology (e.g., frequency of measurement requests)? What methodology to aggregate results? What approach to follow if a metric is not returned from a given network segment? How to accommodate the use of metrics that may not be supported by all devices along the whole forwarding path?
- Q5: How measurement and testing methodology are shared between involved parties (e.g., between two service providers)? Should respective responsibilities be negotiated?
- Q6: How to ensure time synchronization?
- Q7: How can a measurement system dynamically discover the measuring entities of a single domain? Across several domains?
- Q8: How to detect a network is LMAP-compliant? How to configure a LMAP client with LMAP server information?
- Q9: How to guarantee the accuracy of collected data?
- Q10: How to control access to measurement results? How to prevent revealing measurement results to external parties?
- Q11: How to map collected data with technical clauses included in a contract/agreement (e.g., CPP)?
- Q12: Flash crowd issues: to what extent measurement traffic can impact the delivered service during a crisis (e.g., an overload situation in some regions of the LMAP domain, where a LMAP domain is an administrative entity that is composed of LMAP-capable nodes operated by a single structure)?
- Q13: How to make sure that the entities involved in measurement do not dramatically affect the accuracy of the measurement (as per Heisenberg principle)? Which procedure to apply to control the reliability of LMAP agents?
- Q14: How to make sure measurement data is not impacted by the home network itself or the machine embedding the measurement agent?

- Q15: How can a network provider instruct a LMAP agent to hold its requests to prevent network congestion situations (e.g., to avoid link overload)?
- Q16: How to make sure measurement data accurately reflect the network performance and not the policies enforced in that network?
- Q17: The LMAP system can be used to assess the level of delivered connectivity service to customers? The system can be embedded in robots enabled in the access segment to emulate the behavior of connected device. How LMAP can accommodate such deployment use case?
- Q18: To what extent conducting a set of measurement actions at T0 will reelect the actual traffic performance to be experienced when invoking the subscribed service?
- Q19: How path diversity impacts measurements?
- Q20: How the system is designed to ensure topology hiding?

3. Security Considerations

TBC.

4. IANA Considerations

This document does not require any action from IANA.

5. Acknowledgments

TBC.

6. Informative References

[I-D.boucadair-connectivity-provisioning-profile]

Boucadair, M., Jacquenet, C., and N. Wang, "IP/MPLS Connectivity Provisioning Profile", draft-boucadair-connectivity-provisioning-profile-02 (work in progress), September 2012.

[I-D.morton-ippm-lmap-path]

Bagnulo, M., Burbridge, T., Crawford, S., Eardley, P., and A. Morton, "A Reference Path and Measurement Points for LMAP", draft-morton-ippm-lmap-path-00 (work in progress), January 2013.

Authors' Addresses

Mohamed Boucadair
France Telecom
Rennes, 35000
France

Email: mohamed.boucadair@orange.com

Christian Jacquenet
France Telecom
Rennes, 35000
France

Email: christian.jacquenet@orange.com

Network Working Group
Internet-Draft
Intended status: Informational
Expires: August 29, 2013

P. Eardley
T. Burbridge
BT
A. Morton
AT&T
February 25, 2013

A framework for large-scale measurements
draft-eardley-lmap-framework-01

Abstract

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

Status of this Memo

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

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

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

This Internet-Draft will expire on August 29, 2013.

Copyright Notice

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

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

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

Table of Contents

1. Introduction	3
2. Outline of framework	4
3. Constraints	6
3.1. Measurement system is under the direction of a single organisation	6
3.2. Each MA may only have a single Controller at any point in time	7
3.3. A Measurement Agent acts autonomously	7
4. Work required in LMAP	8
4.1. Defining the Test and Report Schedules	8
4.2. Defining the Report	9
5. Related work required but out of scope of LMAP	9
5.1. Standard measurement tests	10
5.2. Characterisation plan	10
5.3. Other elements	10
6. IANA Considerations	10
7. Security Considerations	11
8. Acknowledgements	12
9. Informative References	12
Authors' Addresses	13

1. Introduction

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

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

The LMAP framework should be useful for all these.

The key requirement is for large scale. A measurement system might have at least ~100k Measurement Agents.

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

- o Standardised - in terms of the tests that they perform, the components, the data models and protocols for transferring information between the components. For example so that it is meaningful to compare measurements made of the same metric at different times and places. Today's systems are proprietary in some or all of these aspects.
- o Extensible - it should be easy to add or modify tests, for example an improved test methodology or to measure a performance metric not previously considered important (eg bufferbloat).
- o Large-scale - [use-cases] envisages Measurement Agents in every home gateway and edge device such as set-top-boxes and tablet computers. Existing systems have up to a few thousand Measurement Agents (without judging how much further they could scale).
- o Diversity - a measurement system should handle different types of Measurement Agent - for example Measurement Agents may come from different vendors, be in wired and wireless networks and be on devices with IPv4 or IPv6 addresses.

This section forms the problem statement / aim of the proposed LMAP working group, in the context of the constraints and scope given below.

2. Outline of framework

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

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

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

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

Two Measurement Agents (MAs) jointly perform an active measurement test, by generating test traffic and measuring some metric associated with its transfer over the path from one MA to the other; for example the time taken to transfer a 'test file'. Some tests may involve more than two MAs; for example, to measure 'latency under load'. A single MA may also conduct passive testing through the observation of traffic (ie without the involvement of a second MA); for example an end user's mix of applications.

The MA functions are implemented either in specialised hardware or as code on general purpose devices like a PC, tablet or smartphone.

- o Comment: It may be useful to distinguish two types of MA - a 'complete MA' that interacts with the Controller and Collector, and a 'remote MA' that only takes part in active tests (and does not interact with the Controller and Collector). This is for further study.
- o Comment: A 'complete MA' may want to run a test against a normal, non-LMAP device, for example the time for a DNS response or a webpage download from www.example.com.

The Controller manages a MA by instructing it which tests it should perform and when. For example it may instruct a MA at a home gateway: "Run the 'download speed test' with the test server at the

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

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

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

- o Comment: It is possible that communications between two Collectors, two Controllers and a Controller and Collector may be useful in some use cases, perhaps to help a measurement system scale. It is for further study whether such communications should be in or out of scope of LMAP.
- o Comment: The Initialiser, Subscriber Parameter Database, Results Database, Data Analysis Tools and OAM are out of scope of LMAP. They may be provided through existing protocols or applications and are likely to be part of a complete large-scale measurement system.

An Initialiser configures a MA with details about its Controller, including authentication credentials. Possible protocols are SNMP, NETCONF or (for Home Gateways) CPE WAN Management Protocol (CWMP) from the Auto Configuration Server (ACS) (as specified in TR-069).

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

A Results Database records all measurements in an equivalent form, for example an SQL database [schulzrinne], so that they can be easily

accessed by the Data Analysis Tools whilst the LMAP system implementor can choose local solutions for each component.

The Data Analysis Tools receive the results from the Collector or via the Results Database. They might visualise the data or identify which component or link is likely to be the cause of a fault or degradation.

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

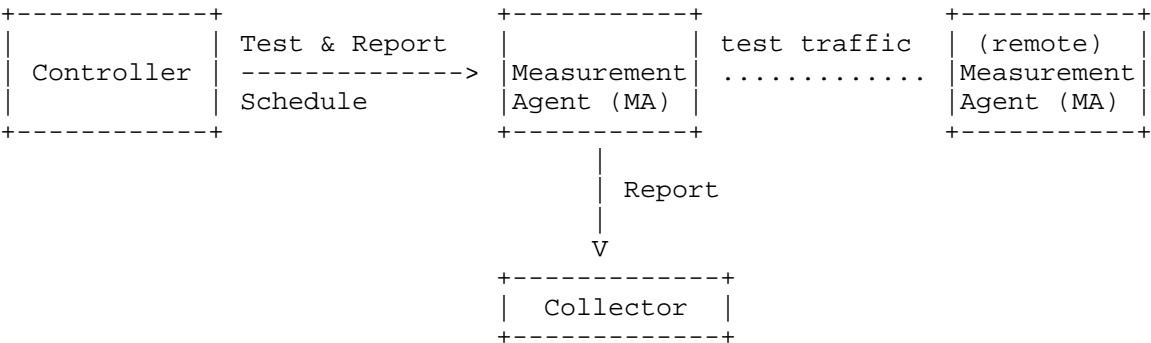


Figure 1: Schematic of main elements of LMAP framework

3. Constraints

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

Explanation: 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 privacy and network security.

Given the novelty of large-scale measurement efforts, the expectation is that inter-organization coordination is an out-of-band consideration. There could be scenarios where measurement data, or a suitably anonymised version of it, is shared between organisations, via their Data Analysis Tools for instance. Consideration is outside the scope of LMAP.

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

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

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

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

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

3.3. A Measurement Agent acts autonomously

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

Firstly, this means that the MA initiates measurement tests. For the typical case where the MA is on a home gateway or edge device, this means that the MA initiates a 'download speed test' by asking a remote MA to send the file. The main rationale is that, for a test

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

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

4. Work required in LMAP

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

4.1. Defining the Test and Report Schedules

The Test and Report Schedules contain the instructions sent by the Controller to the MA. The Schedules could be combined into a single Schedule, this is for further study.

The Test Schedule would include things like:

- o Which tests to operate and (if applicable) to which remote MAs
- o The testing schedule
- o Any test or environmental parameters
- o How to react to the presence of user or other test traffic (if not inherent in the test design)

The Report Schedule would include things like:

- o How often to report results (e.g. time or volume of data)
- o Where to report results
- o What to do if reporting fails

Defining Test and Report Schedule(s) is in scope of LMAP:

- o Information model: the abstract definition of the Schedule
- o Data model: which instantiates the information model in a particular language. It could be done using an existing IETF data modeling language, for example YANG as sketched in [lmap-yang].
- o Protocol: how the Test and Report Schedule(s) are delivered from the Controller to the MA. Possibilities would include NETCONF [RFC6241] as discussed in [lmap-netconf] or a RESTful interface [yang-api].

4.2. Defining the Report

The Report contain the measurement results sent by the Controller to the MA. The Report includes things like:

- o The results of the test (typically at least all the singleton measurements, including the time they were measured)
- o The MA's identifier
- o The test and its parameters (essentially an 'echo' of the Test Schedule with the parameters actually used, which avoids the Collector having to ask the Controller).

Defining the Report is in scope of LMAP:

- o Information model: the abstract definition of the Report
- o Data model: which instantiates the information model in a particular language. It could be done using an existing IETF data modeling language, for example IPFIX, as sketched in [lmap-ipfix] or
- o Protocol: how the Report is delivered from the MA to the Collector. Possibilities would include IPFIX, as sketched in [lmap-ipfix] or a RESTful interface.

5. Related work required but out of scope of LMAP

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

5.1. Standard measurement tests

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

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

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

5.2. Characterisation plan

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

5.3. Other elements

The other elements discussed in Section 2 may also benefit from standardisation: Initialiser, Subscriber Parameter Database, Results Database, Data Analysis Tools and operator's OAM.

The Initialiser-MA protocol is likely to be technology specific and so for different types of device could be defined by the Broadband Forum, DOCSIS or IEEE. The Data Analysis Tools and operator's OAM are also beyond the scope of the IETF. For the Subscriber Parameter Database and Results Database, it is possible that there could be work to define a data model - it is suggested that this is for later study and should be out of the initial scope of IETF work.

6. IANA Considerations

This document makes no request of IANA.

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

7. Security Considerations

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

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

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

Availability should also be considered. While the loss of some MAs may not be considered critical, the unavailability of the Collector could mean that valuable business data or data critical to a regulatory process is lost. Similarly, the unavailability of a Controller could mean that the MAs continue to operate an incorrect test schedule or fail to initiate.

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

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

8. Acknowledgements

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

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

9. Informative References

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

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

Authors' Addresses

Philip Eardley
BT

Trevor Burbridge
BT

Al Morton
AT&T

INTERNET-DRAFT
Intended Status: Informational
Expires: August 29, 2013

Marc Linsner
Cisco Systems
Philip Eardley
Trevor Burbridge
BT
February 25, 2013

Large-Scale Broadband Measurement Use Cases
draft-linsner-lmap-use-cases-02

Abstract

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

Status of this Memo

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

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

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

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

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

Copyright and License Notice

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

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

Table of Contents

1	Introduction	3
1.1	Terminology	3
2	Use Cases	3
2.1	Internet Service Provider (ISP) Use Case	3
2.2	End User Network Diagnostics	4
2.3	Multi-provider Network Measurements	4
2.4	Over the Top Providers	4
2.5	Regulators	5
3	Details of ISP Use Case	6
3.1	Existing Capabilities and Shortcomings	6
3.2	Understanding the quality experienced by customers	6
3.3	Benchmarking and competitor insight	8
3.4	Understanding the impact and operation of new devices and technology	8
3.5	Design and planning	9
3.6	Identifying, isolating and fixing network problems	11
3.7	Comparison with the regulator use case	12
3.8	Conclusions	13
4	Security Considerations	14
5	IANA Considerations	14
6	Contributors	14
7	References	14
7.1	Normative References	14
	Authors' Addresses	15

1 Introduction

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

1.1 Terminology

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

2 Use Cases

2.1 Internet Service Provider (ISP) Use Case

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

- o Identifying, isolating and fixing problems in the network, services or with CPE and end user equipment. Such problems may be common to a point in the network topology (e.g. a single exchange), common to a vendor or equipment type (e.g. line card or home gateway) or unique to a single user line (e.g. copper access). Part of this process may also be helping users understand whether the problem exists in their home network or with an over-the-top service instead of with their BB product.
- o Design and planning. Through identifying the end user experience the ISP can design and plan their network to ensure specified levels of user experience. Services may be moved closer to end users, services upgraded, the impact of QoS assessed or more capacity deployed at certain locations. SLAs may be defined at network or product boundaries.
- o Benchmarking and competitor insight. The operation of sample panels across competitor products can enable an ISP to assess where they play in the market, identify opportunities where other products operate different technology, and assess the performance

of network suppliers that are common to both operators.

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

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

2.2 End User Network Diagnostics

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

2.3 Multi-provider Network Measurements

As an extension of the first use case, multiple network providers and third parties, such as a regulatory body, may collaborate to gather network performance data on a one-time or recurring basis, using a subset of customers of the service providers. The form of collaboration is beyond the scope of this paper, however it should be understood that a data collection platform must serve multiple stakeholder interests.

The main consumer of this use case is someone other than the 'last mile' provider.

2.4 Over the Top Providers

Possibly an extension to the Multi-Provider use case, OTT providers have an interest to ensure Quality of Experience (QOE) associated with content consumption. The uniqueness to this use case compared to those mentioned above is the feature that the measurement client

will run on software based apps and embedded apps such as those found in set-top boxes or disc players. It is assumed that measurement tests run by OTT providers would only include the metrics associated with layer 3 and up.

The main consumer of this use case are content providers.

2.5 Regulators

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

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

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

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

LMAP standards could answer regulators shared needs by providing scalable, cost-effective, scientifically robust solutions to the measurement and collection of broadband performance information.

The main consumer of this use case are regulators

3 Details of ISP Use Case

3.1 Existing Capabilities and Shortcomings

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

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

Faced with a gap in current vendor offerings some ISPs have taken the approach of placing proprietary test capabilities on their home gateway and other consumer device offerings (such as Set Top Boxes). This also means that different device platforms may have different and largely incomparable tests, developed by different company sub-divisions managed by different systems.

3.2 Understanding the quality experienced by customers

Operators want to understand the quality of experience (QoE) of their broadband customers. The understanding can be gained through a "panel", ie a measurement probe is deployed to a few 100 or 1000 of its customers. The panel needs to be a representative sample for each of the operator's technologies (FTTP, FTTC, ADSL...) and broadband options (80Mb/s, 20Mb/s, basic...), ~100 probes for each. The operator would like the end-to-end view of the service, rather than (say) just the access portion. So as well as simple network

statistics like speed and loss rates they want to understand what the service feels like to the customer. This involves relating the pure network parameters to something like a 'mean opinion score' which will be service dependent (for instance web browsing QoE is largely determined by latency above a few Mb/s).

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

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

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

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

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

1. Averaged data (over say 1 month) is generally ok
2. A panel (subset) of only a few customers is OK
3. Both active and passive measurements are possible, though the former seems easier
4. Regularly scheduled tests are fine (providing active tests back off if the customer is using the line). Scheduling can be done some time ahead ('starting tomorrow, run the following test every day').
5. The operator needs to devise metrics and compound measures

that represent the QoE

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

3.3 Benchmarking and competitor insight

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

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

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

1. Averaged data (over say 1 month) is generally ok
2. A panel (subset) of only a few customers is OK
3. Both active and passive measurements are possible, though the former seems easier
4. Regularly scheduled tests are fine (providing active tests back off if the customer is using the line). Scheduling can be done some time ahead ('starting tomorrow, run the following test every day').
5. The performance metrics are whatever the operator wants to benchmark. As well as QoE measures, it may want to measure some network-specific parameters.
6. As well as the performance of the access link, the performance of different network segments, including end-to-end.

3.4 Understanding the impact and operation of new devices and technology

Another type of measurement is to test new capabilities and services

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

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

1. New tests need to be devised that test a prospective capability.
2. Most of the tests are probably simply: "send one packet and record what happens", so an occasional one-off test is sufficient.
3. A panel (subset) of only a few customers is probably OK, to gain an understanding of the impact of a new technology, but it may be necessary to check an individual line where the roll-out is per customer.
4. An active measurement is needed.

3.5 Design and planning

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

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

It may also be possible to run stress tests for risk analysis, for example 'if whizzy new application (or device) becomes popular, which parts of my network would struggle, what would be the impact on other services and how many customers would be affected'.

Another example is that the operator may want to monitor performance

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

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

Through a better understanding of its own network and its suppliers, the operator should be able to focus investment more effectively - in the right place at the right time with the right technology. What-if tests could help quantify the advantage that a new technology brings and support the business case for larger roll-out.

The characteristics of large scale measurements emerging from these examples:

1. A key challenge is how to integrate results from measurements into existing network planning and management tools
2. New tests may need to be devised for the what-if and risk analysis scenarios.
3. Capacity constraints first reveal themselves during atypical events (early warning). So averaging of measurements should be over a much shorter time than the sub use case discussed above.
4. A panel (subset) of only a few customers is OK for most of the examples, but it should probably be larger than the QoE use case #1 and the operator may also want to regularly change who is in the subset, in order to sample the revealing outliers.
5. Measurements over a segment of the network ("end-to-middle") are needed, in order to refine understanding, as well as end-to-end measurements.
6. The primary interest is in measuring specific network performance parameters rather than QoE.
7. Regularly scheduled tests are fine
8. Active measurements are needed; passive ones probably aren't

3.6 Identifying, isolating and fixing network problems

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

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

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

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

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

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

and get the result whilst the customer is still on-line session.

The characteristics of large scale measurements emerging from these examples:

1. A key challenge is how to integrate results from measurements into the operator's existing Test and Diagnostics system.
2. Results from the tests shouldn't be averaged
3. Tests are generally run on an ad hoc basis, ie specific requests for immediate action
4. "End-to-middle" measurements, ie across a specific network segment, are very relevant
5. The primary interest is in measuring specific network performance parameters and not QoE
6. New tests are needed for example to check the home network (ie the connection from the home hub to the set top boxes or to a tablets on wifi)
7. Active measurements are critical. Passive ones may be useful to help understand exactly what the customer is experiencing.

3.7 Comparison with the regulator use case

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

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

- o Performance metrics: A regulator and operator are generally interested in the same performance metrics. Both would like standardised metrics, though this is more important for regulators.
- o Sampling: The regulator wants an average across a representative sample of broadband customers (per operator, per

type of BB contract). The operator also wants to measure individual lines with a problem.

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

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

- o Pre-processing: A regulator would like standard ways of processing the collected data, to remove outlier measurements and aggregate results, because this can significantly affect the final "averaged" result. Pre-processing is not important for an operator.

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

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

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

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

3.8 Conclusions

There is a clear need from an ISP point of view to deploy a single

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

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

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

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

4 Security Considerations

TBD

5 IANA Considerations

TBD

6 Contributors

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

James Miller jamesmilleresquire@gmail.com

7 References

7.1 Normative References

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

[LMAP-REQ] Schulzrinne, H., "Large-Scale Measurement of Broadband

Performance: Use Cases, Architecture and Protocol Requirements", draft-schulzrinne-lmap-requirements, September, 2012

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

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

Authors' Addresses

Marc Linsner
Marco Island, FL
USA

Email: mlinsner@cisco.com

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

Email: philip.eardley@bt.com

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

Email: trevor.burbridge@bt.com

Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: August 28, 2013

M. Bagnulo
UC3M
T. Burbridge
BT
S. Crawford
SamKnows
P. Eardley
BT
A. Morton
AT&T Labs
February 24, 2013

A Reference Path and Measurement Points for LMAP
draft-morton-ippm-lmap-path-01

Abstract

This document defines a reference path for Large-scale Measurement of Broadband Access Performance (LMAP) and measurement points for commonly used performance metrics.

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 August 28, 2013.

Copyright Notice

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

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

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

Table of Contents

1. Introduction	3
2. Purpose and Scope	3
3. Terms and Definitions	4
3.1. Reference Path	4
4. Reference Path	4
5. Measurement Points	5
6. Translation Between Ref. Path and Tech. X	7
7. Security considerations	8
8. IANA Considerations	8
9. Acknowledgements	8
10. References	8
10.1. Normative References	8
10.2. Informative References	9
Authors' Addresses	10

1. Introduction

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

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

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

The motivation for this memo is to provide an unambiguous framework to describe measurement coverage, or scope of the reference path. This is an essential part of the metadata to describe measurement results. Measurements conducted over different path scopes are not a valid basis for performance comparisons.

2. Purpose and Scope

The scope of this memo is to define a reference path for LMAP activities with sufficient level of detail to determine the location of different measurement points without ambiguity.

The bridge between the reference path and specific network technologies (with differing underlying architectures) is within the scope of this effort. Both wired and wireless technologies are in-scope.

The purpose is to create an efficient way to describe the location of the measurement point(s) used to conduct a particular measurement so that the measurement result will adequately described in this regard. This should serve many measurement uses, including diagnostic (where the same metric may be measured over many different path scopes) and

comparative (where the same metric may be measured on different network infrastructures).

3. Terms and Definitions

3.1. Reference Path

A reference path is a serial combination of routers, switches, links, radios, and processing elements that comprise all the network elements traversed by each packet between the source and destination hosts. The reference path is intended to be equally applicable to all networking technologies, therefore the components are generically defined, but their functions should have a clear counterpart or be obviously omitted in any network technology.

4. Reference Path

This section defines a reference path for Internet Access.

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

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

GRA = Globally Routable Address, GW = Gateway

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

- o Subsc. (Subscriber) device - This is a host that normally originates and terminates communications conducted over the IP packet transfer service.
- o Private Net #x - This is a network of devices owned and operated by the Internet Access Service Subscriber. In some configurations, one or more private networks and the device that provides the Access Service Demarcation point are collapsed in a single device (and ownership may shift to the service provider), and this should be noted as part of the path description.
- o Access (Service) Demarcation point - this varies by technology but is usually defined as the Ethernet interface on a residential gateway or modem where the scope of access packet transfer service begins and ends. In the case of a WiFi Service, this would be an Air Interface within the intended service boundary (e.g., walls of

the coffee shop). The Demarcation point may be within an integrated endpoint using an Air Interface (e.g., LTE UE). Ownership may not affect the demarcation point; a Subscriber may own all equipment on their premises, but it is likely that the service provider will certify such equipment for connection to their access network, or a third-party will certify standards compliance.

- o Intra IP Access - This is the first point in the access architecture beyond the Access Service Demarc. where a globally routable IP address is exposed and used for routing. In architectures that use tunneling, this point may be equivalent to the GRA GW. This point could also collapse to the device providing the Access Service Demarc., in principle. Only one Intra IP Access point is shown, but they can be identified in any access or transit network.
- o GRA GW - the point of interconnection between the access administrative domain and the rest of the Internet, where routing will depend on the GRAs in the IP header.
- o Transit GRA GW - Networks that intervene between the Subscriber's Access network and the Destination Host's network are designated "transit" and involve two GRA GW.

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

In the case that a private address space is used throughout an access architecture, then the Access Service Demarc. and the Intra IP Access points must use the same address space and be separated by the shared and dedicated access link infrastructure, such that a test between these points produces a useful assessment of access performance.

5. Measurement Points

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

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

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

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

...	Transit	--	GRA	--	Service	--	Private	--	Private	--	Destination
	GRA GW		GW		Demarc.		Net #n		Net #n+1		Host
	mpX90		mp890		mp800						mp900

GRA = Globally Routable Address, GW = Gateway

The numbering for measurement points (mpNNN) allows for considerable local use of unallocated numbers.

Notes:

- o Some use the terminology "on-net" and "off-net" when referring to Internet Service Provider (ISP) measurement coverage. With respect to the reference path, tests between mp100 and mp190 are "on-net".
- o Widely deployed broadband access measurements have used pass-through devices[SK] (at the subscriber's location) directly connected to the service demarcation point: this would be located at mp100.
- o The networking technology used at all measurement points must be indicated, especially the interface standard and configured speed.
- o If it can be shown that a link connecting to a measurement point has reliably deterministic or negligible performance, then the remote end of the connecting link is an equivalent point for some methods of measurement (To Be Specified Elsewhere). In any case, the presence of such a link must be reported.
- o Many access network architectures have a traffic aggregation point (e.g., CMTS or DSLAM) between mp100 and mp150. We designate this point mp120, but it won't currently fit in the figure.
- o A Carrier Grade NAT (CGN) deployed in the Subscriber's access network would be positioned between mp100 and mp190, and the egress side of the CGN will typically be designated mp150.
- o In the case that a private address space is used in an access architecture, then mp100 may need to use the same address space as its remote measurement point counterpart, so that a test between these points produces a useful assessment of network performance. Tests between mp000 and mp100 could use private address space, and when the egress side of a CGN is at mp150, then the private address side of the CGN could be designated mp149 for tests with mp100.

- o Measurement points at Transit GRA GWs are numbered mpX00 and mpX90, where X is the lowest positive integer not already used in the path.

6. Translation Between Ref. Path and Tech. X

This section and those that follow are intended to provide a more exact mapping between particular network technologies and the reference path.

We provide an example for 3G Cellular access below.

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

	_____UE_____		_____RAN+Core_____		_____GGSN_____	
--	--------------	--	--------------------	--	----------------	--

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

We next provide a few examples of DSL access. Consider first the case where:

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

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

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

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

	---UE---		-----CPE/NAT-----		-----BRAS-----		-----Access Network-----	
--	----------	--	-------------------	--	----------------	--	--------------------------	--

GRA = Globally Routable Address, GW = Gateway

Consider next the case where:

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

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

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

Subsc. device	Private Net #1	Private Net #2	Access Demarc.	Intra IP Access	GRA GW	Transit GRA GW
mp000			mp100	mp150	mp190	mp200
--UE--	-----CPE/NAT-----		-----	-CGN-	-----	
			---Access Network---			

GRA = Globally Routable Address, GW = Gateway

7. Security considerations

Specification of a Reference Path and identification of measurement points on the path represent agreements among interested parties, and they present no threat to the readers of this memo or to the Internet itself.

8. IANA Considerations

TBD

9. Acknowledgements

Thanks to Matt Mathis for review and comments.

10. References

10.1. Normative References

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

- [RFC3432] Raisanen, V., Grotefeld, G., and A. Morton, "Network performance measurement with periodic streams", RFC 3432, November 2002.
- [RFC2681] Almes, G., Kalidindi, S., and M. Zekauskas, "A Round-trip Delay Metric for IPPM", RFC 2681, September 1999.
- [RFC6673] Morton, A., "Round-Trip Packet Loss Metrics", RFC 6673, August 2012.
- [RFC1035] Mockapetris, P., "Domain names - implementation and specification", STD 13, RFC 1035, November 1987.
- [RFC5905] Mills, D., Martin, J., Burbank, J., and W. Kasch, "Network Time Protocol Version 4: Protocol and Algorithms Specification", RFC 5905, June 2010.
- [RFC2679] Almes, G., Kalidindi, S., and M. Zekauskas, "A One-way Delay Metric for IPPM", RFC 2679, September 1999.
- [RFC2680] Almes, G., Kalidindi, S., and M. Zekauskas, "A One-way Packet Loss Metric for IPPM", RFC 2680, September 1999.
- [RFC3393] Demichelis, C. and P. Chimento, "IP Packet Delay Variation Metric for IP Performance Metrics (IPPM)", RFC 3393, November 2002.
- [RFC5481] Morton, A. and B. Claise, "Packet Delay Variation Applicability Statement", RFC 5481, March 2009.
- [RFC5835] Morton, A. and S. Van den Berghe, "Framework for Metric Composition", RFC 5835, April 2010.

10.2. Informative References

- [RFC4148] Stephan, E., "IP Performance Metrics (IPPM) Metrics Registry", BCP 108, RFC 4148, August 2005.
- [RFC6248] Morton, A., "RFC 4148 and the IP Performance Metrics (IPPM) Registry of Metrics Are Obsolete", RFC 6248, April 2011.
- [SK] Crawford, Sam., "Test Methodology White Paper", SamKnows Whitebox Briefing
Note <http://www.samknows.com/broadband/index.php>,
July 2011.

Authors' Addresses

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

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

Trevor Burbridge
British Telecom
Adastral Park, Martlesham Heath
IPswitch
ENGLAND

Email: trevor.burbridge@bt.com

Sam Crawford
SamKnows

Email: sam@samknows.com

Phil Eardley
British Telecom
Adastral Park, Martlesham Heath
IPswitch
ENGLAND

Email: philip.eardley@bt.com

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

Email: acmorton@att.com

Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: August 19, 2013

J. Schoenwaelder
Jacobs University Bremen
February 15, 2013

Considerations on using NETCONF with LMAP Measurement Agents
draft-schoenw-lmap-netconf-00.txt

Abstract

This document discusses how the NETCONF protocol can be used to configure LMAP measurement agents.

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 August 19, 2013.

Copyright Notice

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

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

Table of Contents

1. Introduction	3
2. Properties of Measurement Agents	3
3. Properties of the NETCONF Protocol	3
4. Discussion	4
4.1. Connection Initiation	4
4.2. Client and Server Role	5
4.3. Identification of Configuration Versions	5
4.4. Pushing of Measurement Results	5
4.5. NETCONF versus YANG-API	6
5. Security Considerations	6
6. IANA Considerations	7
7. Acknowledgements	7
8. Informative References	7
Author's Address	8

1. Introduction

This document discusses how the NETCONF protocol [RFC6241] can be used to configure Large-Scale Measurement of Broadband Performance (LMAP) measurement agents (MAs), sometimes also called measurement clients [I-D.schulzrinne-lmap-requirements].

MAs may be deployed as separate hardware devices or as functions embedded in consumer electronic devices and home routers or as pure software solutions that can be installed on off-the-shelf computing equipment. Measurement agents receive instructions from a controller when and how to conduct what measurements (the measurement schedule) and how and when to report measurement results to a data collector (the report schedule). Further information about the interaction between MAs and controllers and collectors can be found in [I-D.schulzrinne-lmap-requirements].

2. Properties of Measurement Agents

Measurement Agents (MAs) have a number of important properties:

1. MAs are often deployed behind Network Address Translators (NATs). This might even be true if MAs are part of a device on the demarcation line between a service provider and a home network due to the usage of Carrier Grade NATs in the service provider network.
2. MAs may run on devices that are not always powered up and online.
3. A single controller may be responsible for a large number of MAs.
4. A large fraction of the MAs may be inactive (i.e., they do not perform any measurements) at any given point in time. Inactive MAs may need to be enabled on demand for example to troubleshoot specific problems (e.g., as part of customer helpdesk services) or to balance measurement traffic load.

3. Properties of the NETCONF Protocol

The Network Configuration Protocol (NETCONF) [RFC6241] provides mechanisms to install, manipulate, and delete the configuration of network devices. It uses an Extensible Markup Language (XML)-based data encoding for the configuration data as well as the protocol messages. The NETCONF protocol operations are realized as remote procedure calls (RPCs).

The NETCONF protocol runs by default over the Secure Shell protocol (SSH) [RFC6242] but it can also be used over Transport Layer Security (TLS) [RFC5539] with pairwise authentication using X.509 certificates.

NETCONF has been originally designed to be used on network devices such as backbone routers. A device supporting NETCONF has an embedded NETCONF server. Configuration management applications use embedded NETCONF clients to connect to NETCONF servers and then issue RPC calls to manipulate the configuration state of the devices.

4. Discussion

This section discusses certain technical challenges related to the usage of NETCONF to configure MAs.

4.1. Connection Initiation

Due to the nature of LMAP MAs (likely located behind NATs), it is crucial that MAs initiate connections to a controller. This is currently not supported in NETCONF. There were previous attempts to provide a so called call-home mechanism for NETCONF, see for example [I-D.kwatsen-reverse-ssh]. The current state of the art, however, is that there is no standardized call home mechanism for NETCONF over SSH.

For the NETCONF over TLS transport [RFC5539], which relies on mutual authentication using X.509 certificates, it seems easier to support call home. In fact, the NETCONF over TLS transport specification is currently being updated and hence there is an opportunity to get call home support into this update on relative short notice. The work essentially requires to detail aspects such as port numbers used to connect from a NETCONF server (acting as a TCP client) to a NETCONF client (acting as a TCP server). In addition, a YANG data model would be desirable that can be used to configure the call home policy (when to call home) and the addresses to connect to and perhaps the certificate to use.

Decoupling the NETCONF server / client role from the TCP server / client role turns out to be straight-forward since every NETCONF session starts with an exchange of <hello> messages. The <hello> message sent by the NETCONF server includes a session-id while the NETCONF client does not send a session-id. As such, both endpoints can easily identify and verify who is acting as NETCONF client and NETCONF server. [RFC6241] already handles the possible error cases (i.e., a NETCONF server connecting to a NETCONF server or a NETCONF client connecting to a NETCONF client).

4.2. Client and Server Role

Some large scale measurement deployments use proprietary protocols where the server role is on the controller. In a nutshell, the MA connects to the controller running a server and checks if there is a configuration update to load. If so, the MA fetches the necessary new configuration information and then applies it locally.

The NETCONF protocol, however, assumes that the NETCONF server role is taken by the device that is configured. This would be in the LMAP use case the MA while the NETCONF client would be running on the controller.

Even though this may seem like a major difference in the way the interaction works, it appears that NETCONF can provide the functionality needed. A MA initiating a transport connection and subsequently taking the NETCONF server role enables the controller (acting as a NETCONF client) after the <hello> exchange to take the initiative to determine whether any configuration changes need to be applied to the device. If so, standard <edit-config> operations can be used to modify the device's configuration.

4.3. Identification of Configuration Versions

As mentioned above, the controller (running a NETCONF client) must determine whether a device's configuration needs updates. While this could be achieved by retrieving the configuration using <get-config> and comparing the result with the expected configuration, this approach is not very efficient. It will be much more effective if the NETCONF server would indicate the version of the configuration it is currently using. The version can either be identified by a version number or a time-stamp of the last configuration change or simply an opaque tag that is handed out and interpreted only by the controller. While the configuration version might simply be modeled as a regular data object that the NETCONF client retrieves in the usual way, it might be useful to consider optimizations, e.g., carrying the configuration version as part of a new capability in the <hello> exchange.

4.4. Pushing of Measurement Results

NETCONF has not been designed as a data push protocol. While a NETCONF extension [RFC5277] provides support for event notifications, this mechanism requires in its simplest form that a NETCONF client first subscribes to an event stream and that the session used to carry event notification stays open. This is not scalable in the LMAP scenario.

One possible way to work around this limitation within the framework of the current NETCONF protocol is to make use of the event notification replay feature: A MA is locally collecting measurement results. After connecting to a collector (acting as a NETCONF client), the collector subscribes to an event stream with a request to replay the measurement results collected since the last time data has been fetched from the MA. An alternative, of course, would be to model test results as part of an LMAP data model and to use NETCONF <get> operations to retrieve the data.

That said, if close to soft real-time pushing of measurement results from the MA to the collector is required, then NETCONF likely is not the right choice.

4.5. NETCONF versus YANG-API

NETCONF provides a feature rich solution for network configuration management, including support for concurrent access to a NETCONF server by multiple NETCONF clients, different configuration datastores, explicit validation of configurations, and a confirmed-commit procedure to support configuration change transactions spanning multiple devices. A recent proposal called YANG-API [I-D.bierman-netconf-yang-api] aims at providing a simplified interface that follows RESTful principles and is compatible with a resource-oriented device abstraction.

While implementations of YANG-API are in progress, it seems too early to decide whether the benefits of RESTful YANG-API are significant enough to consider it as a possible alternative for LMAP. In particular, it might take a few years for YANG-API to become a stable specification.

5. Security Considerations

The NETCONF protocol [RFC6241] can run over several different transports. Since the protocol manipulates sensitive configuration information, NETCONF requires that all transports provide authentication, data integrity, confidentiality, and replay protection.

There are currently two transport for NETCONF on the standards track. The NETCONF over SSH transport [RFC6242] provides authentication and data encryption services. The NETCONF over SSH specification further requires that the identity of the SSH server must be verified and authenticated by the SSH client according to local policy before password-based authentication data or any configuration or state data is sent to or received from the SSH server. Similarly, the identity

of the SSH client must also be verified and authenticated by the SSH server according to local policy to ensure that the incoming SSH client request is legitimate before any configuration or state data is sent to or received from the SSH client. Neither side should establish a NETCONF over SSH connection with an unknown, unexpected, or incorrect identity on the opposite side.

The NETCONF over TLS transport [RFC5539], currently being revised in [I-D.ietf-netconf-rfc5539bis], provides authentication and data encryption services. In particular, [RFC5539] assumes that both peers authenticate each other using X.509 certificates while [I-D.ietf-netconf-rfc5539bis] adds the possibility to use pre-shared keys.

The NETCONF access control model [RFC6536] provides an authorization model for NETCONF. It allows to configure access control rules that can be used to restrict NETCONF protocol access for particular users to a pre-configured subset of all available NETCONF protocol operations and content. The NETCONF access control model should be required for LMAP implementations that potentially allow access from multiple controllers.

6. IANA Considerations

TBD

7. Acknowledgements

TBD

8. Informative References

[I-D.bierman-netconf-yang-api]

Bierman, A. and M. Bjorklund, "YANG-API Protocol",
draft-bierman-netconf-yang-api-01 (work in progress),
November 2012.

[I-D.ietf-netconf-rfc5539bis]

Badra, M., Luchuk, A., and J. Schoenwaelder, "NETCONF Over
Transport Layer Security (TLS)",
draft-ietf-netconf-rfc5539bis-01 (work in progress),
October 2012.

[I-D.kwatsen-reverse-ssh]

Watsen, K., "Reverse Secure Shell (Reverse SSH)",

draft-kwatsen-reverse-ssh-01 (work in progress),
June 2011.

- [I-D.schulzrinne-lmap-requirements]
Schulzrinne, H., Johnston, W., and J. Miller, "Large-Scale
Measurement of Broadband Performance: Use Cases,
Architecture and Protocol Requirements",
draft-schulzrinne-lmap-requirements-00 (work in progress),
September 2012.
- [RFC5277] Chisholm, S. and H. Trevino, "NETCONF Event
Notifications", RFC 5277, July 2008.
- [RFC5539] Badra, M., "NETCONF over Transport Layer Security (TLS)",
RFC 5539, May 2009.
- [RFC6241] Enns, R., Bjorklund, M., Schoenwaelder, J., and A.
Bierman, "Network Configuration Protocol (NETCONF)",
RFC 6241, June 2011.
- [RFC6242] Wasserman, M., "Using the NETCONF Protocol over Secure
Shell (SSH)", RFC 6242, June 2011.
- [RFC6536] Bierman, A. and M. Bjorklund, "Network Configuration
Protocol (NETCONF) Access Control Model", RFC 6536,
March 2012.

Author's Address

Juergen Schoenwaelder
Jacobs University Bremen

Email: j.schoenwaelder@jacobs-university.de

Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: August 19, 2013

J. Schoenwaelder
Jacobs University Bremen
February 15, 2013

A YANG Data Model for LMAP Measurement Agents
draft-schoenw-lmap-yang-00.txt

Abstract

This document sketches a data model for configuring and scheduling tests for large scale broadband access network measurements. The data model is defined using the YANG data modeling language.

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 August 19, 2013.

Copyright Notice

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

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

Table of Contents

1. Introduction	3
2. Data Model Overview	3
3. YANG Module	4
4. Security Considerations	9
5. IANA Considerations	9
6. Acknowledgements	9
7. References	9
7.1. Normative References	9
7.2. Informative References	10
Author's Address	10

1. Introduction

This document sketches a data model for configuring and scheduling tests for large scale broadband access network measurements. The data model is defined using the YANG [RFC6020] data modeling 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 [RFC2119].

2. Data Model Overview

The data model has the following structure, where square brackets are used to enclose a list's keys, "?" means that the leaf is optional, and "*" denotes a leaf-list:

```

module: acme-lmap
  +--rw lmap
    +--rw tests
      |   +--rw test [name]
      |   |   +--rw name          string
      |   |   +--rw description?  string
      |   |   +--rw program       string
      |   |   +--rw option [name]
      |   |   |   +--rw name      string
      |   |   |   +--rw value?    string
      |   |   +--rw argument*     string
    +--rw schedules
      +--rw schedule [name]
        +--rw name          string
        +--rw test?         leafref
        +--rw enabled?      boolean
        +--rw (schedule-type)?
          |   +--:(periodic)
          |   |   +--rw interval?      uint32
          |   +--:(calendar)
          |   |   +--rw weekday?       weekday-set
          |   |   +--rw month-set?     months-set
          |   |   +--rw day*           int8
          |   |   +--rw hour*          int8
          |   |   +--rw minute*        int8
          |   +--:(one-shot)
        +--ro failures?     yang:counter32
        +--ro last-failure?  string
        +--ro last-failed?  yang:date-and-time

```

3. YANG Module

The data model uses types imported from [RFC6021].

```
module acme-lmap {  
  
  namespace "http://www.example.org/yang/acme-lmap";  
  
  prefix "lmap";  
  
  import ietf-yang-types {  
    prefix yang;  
  }  
  
  organization  
    "Large-Scale Measurement of Broadband Performance Project";  
  
  contact  
    "Web:      <http://www.example.org/>  
  
    Editor:    Juergen Schoenwaelder  
              <j.schoenwaelder@jacobs-university.de>";  
  
  description  
    "This module provides a data model for configuring tests and  
    test schedules running on LMAP measurement agents.  
  
    Note that the whole data model can be generalized to invoke  
    other actions that are not measurement tests.";  
  
  revision "2013-02-01" {  
    description  
      "Initial version";  
    reference  
      "TBD";  
  }  
  
  typedef weekday-set {  
    type bits {  
      bit sunday;  
      bit monday;  
      bit tuesday;  
      bit wednesday;  
      bit thursday;  
      bit friday;  
      bit saturday;  
    }  
    description
```

```
    "A type modeling sets of weekdays in the Greco-Roman
    tradition.";
}

typedef months-set {
    type bits {
        bit january;
        bit february;
        bit march;
        bit april;
        bit may;
        bit june;
        bit july;
        bit august;
        bit september;
        bit october;
        bit november;
        bit december;
    }
    description
        "A type modeling sets of months in the Julian and Gregorian
        tradition.";
}

container lmap {

    container tests {
        config true;

        list test {
            key name;

            description
                "The list of tests configured on the lmap probe.";

            leaf name {
                type string;
                description
                    "The unique name of a configured test.";
            }

            leaf description {
                type string;
                description
                    "A short description of the configured test.";
            }

            leaf program {
```

```
    type string;
    mandatory true;
    description
      "The (local) program to invoke in order to execute
       the test.";
  }

  list option {
    key "name";
    ordered-by user;

    description
      "The options passed to the program in order to carry out
       the test. This is a list of key / value pairs and may be
       used to model command line options.";

    leaf name {
      type string;
      description
        "The name of the options.";
    }

    leaf value {
      type string;
      description
        "The value of the options.";
    }
  }

  leaf-list argument {
    type string;
    description
      "The list of arguments passed to the test.";
  }
}

container schedules {
  config true;

  list schedule {
    key name;

    leaf name {
      type string;
      description
        "The locally-unique, administratively assigned name for
         this scheduled test.";
    }
  }
}
```

```
}

leaf test {
  type leafref {
    path "/lmap/tests/test/name";
  }
  description
    "The test invoked by this schedule";
}

leaf enabled {
  type boolean;
  description
    "Indicates whether the test is enabled or disabled.";
}

choice schedule-type {

  case periodic {
    leaf interval {
      type uint32;
      units "seconds";
      description
        "The number of seconds between two action invocations of
        a periodic test. Implementations must guarantee
        that test invocations will not occur before at least
        the specified number of seconds have passed.

        The scheduler must ignore all periodic schedules that
        have a interval value of 0. A periodic schedule
        with a scheduling interval of 0 seconds will therefore
        never invoke a test.

        Implementations may be forced to delay invocations in
        the face of local constraints. A scheduled test should
        therefore not rely on the accuracy provided by the
        scheduler implementation.";
    }
  }

  case calendar {
    leaf weekday {
      type weekday-set;
      description
        "The set of weekdays on which this test should be
        executed. And empty set means any weekday.";
    }
  }
}
```

```
leaf month-set {
    type months-set;
    description
        "The set of months on which this test should be
        executed. And empty set means any month.";
}

leaf-list day {                // make this a set?
    type int8 {
        range "-31..31";
    }
    description
        "The set of days in a months on which this test should
        be executed. Negative days are counted backwards from
        the end of the month.";
}

leaf-list hour {              // make this a set?
    type int8 {
        range "0..23";
    }
    description
        "The set of hours in a day on which this test should
        be executed.";
}

leaf-list minute {           // make this a set?
    type int8 {
        range "0..59";
    }
    description
        "The set of minutes in an hour on which this test
        should be executed.";
}

case one-shot {              // separate or add disable counter
}                             // to the calendar schedule?

leaf failures {
    config false;
    type yang:counter32;
    description
        "The number of failures that occurred while invoking
        scheduled tests.";
}
```

```
    leaf last-failure {
      config false;
      type string;
      description
        "The most recent error that occurred during the invocation of
        a scheduled test.";
    }

    leaf last-failed {
      config false;
      type yang:date-and-time;
      description
        "The date and time when the most recent failure occurred.";
    }
  }
}
}
```

4. Security Considerations

TBD

5. IANA Considerations

TBD

6. Acknowledgements

Some ideas were lifted from [RFC2591].

7. References

7.1. Normative References

- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.
- [RFC6020] Bjorklund, M., "YANG - A Data Modeling Language for the Network Configuration Protocol (NETCONF)", RFC 6020, October 2010.

[RFC6021] Schoenwaelder, J., "Common YANG Data Types", RFC 6021, October 2010.

7.2. Informative References

[RFC2591] Levi, D. and J. Schoenwaelder, "Definitions of Managed Objects for Scheduling Management Operations", RFC 2591, May 1999.

Author's Address

Juergen Schoenwaelder
Jacobs University Bremen

Email: j.schoenwaelder@jacobs-university.de

LMAP
Internet-Draft
Intended status: Informational
Expires: March 25, 2013

H. Schulzrinne
W. Johnston
J. Miller
FCC
September 21, 2012

Large-Scale Measurement of Broadband Performance: Use Cases,
Architecture and Protocol Requirements
draft-schulzrinne-lmap-requirements-00

Abstract

Measuring broadband performance on a large scale is important for network diagnostics by providers and users, as well for as public policy. To conduct such measurements, user networks gather data, either on their own initiative or instructed by a measurement controller, and then upload the measurement results to a designated measurement server. This document describes a logical architecture and summarizes key requirements for protocols to connect the components. The system is designed to support residential and small-enterprise networks, using either wired or wireless networks. The architecture supports an extensible set of active and passive measurements, but the details of the metrics themselves are beyond the scope of this document.

Status of this Memo

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

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

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

This Internet-Draft will expire on March 25, 2013.

Copyright Notice

Copyright (c) 2012 IETF Trust and the persons identified as the document authors. All rights reserved.

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

Table of Contents

1. Introduction	3
2. Use Cases	6
3. Architecture Overview	8
3.1. Measurement client	8
3.2. Measurement server	8
3.3. Measurement controller	8
3.4. Data collector	8
3.5. Network parameter server	9
4. Protocols	10
5. Initiation of Measurements	12
6. Requirements	13
7. Security Considerations	16
8. IANA Considerations	17
9. Acknowledgements	18
10. References	19
10.1. Normative References	19
10.2. Informative References	19
Authors' Addresses	20

1. Introduction

Measuring actual network performance is crucial to managing consumer and enterprise networks, but, when performed at scale, it also allows third parties to gain insight into the actual performance of such networks, facilitating consumer choice and allowing to evaluate the state of broadband performance in a country, among other public policy goals. A number of network performance metrics have been defined, such as [2], but there is no overall architecture and set of protocols that facilitates gathering such measurements in a coordinated way, at scales drawing on thousands or millions of nodes.

Large-scale measurement efforts (e.g., [3]) use proprietary, custom-designed mechanisms to coordinate the measurement clients. They require that the organization running the measurements deploy thousands of dedicated hardware components or rely on end-system software modules that are subject to exogeneous factors, such as home networks, that may distort the results. Thus, this document proposes an overall architecture, with emphasis on the functional and security requirements for the protocols connecting the elements of the architecture, that will make it possible to build measurement capabilities into home and enterprise edge routers, personal computers, mobile devices and other edge devices.

Any usage and implementation will likely impose a number of additional operational requirements and a statistical sampling methodology. For example, the Measurement Broadband America project [3] within the US Federal Communications Commission (FCC) has established specific operational guidelines on data validity and commits to specific requirements for open access to measurement data, software tools and documentation of measurement methodology and statistical approaches. While crucial for deployment, these are beyond the scope of this protocol requirements document. Also, as is customary for IETF-managed protocols, this document does not mandate a specific hardware or operating system platform for implementation.

We suggest that the IETF IP Performance Metrics (IPPM) working group take on defining any additional performance metrics as needed. Such an effort should be undertaken as a collaborative effort with the Broadband Forum (BBF) [4]; other SDOs may also take on aspects of this problem area.

In some applications, such as data gathering by local regulatory entities, extensive logging at various levels, from packet arrival times to events, will be used to assure all parties of the validity of the data gathered. However, logging is beyond the scope of this document.

Both active and passive measurement techniques have been widely accepted in practice. In active measurements, the end system emits traffic and observes a performance metric, or has another end point do so. Examples of active measurements include round-trip delay [2], one-way delay [5] and throughput [6] metrics, service availability, as well as a range of measurements that try to emulate application behavior, such as VoIP, HTTP retrievals or media streaming. Passive measurements observe existing user traffic flows. We note that there is some overlap between NetFlow [7] measurements and passive measurements described here. The delineation between the two and possible re-use of functionality are left to further discussion.

For both active and passive measurements, a measurement client sends or observes traffic, respectively. For active measurements, the measurement client may need a measurement server as well as a recipient of the measurement traffic. (In some cases, such as measurements modeling user access to network services, such as web page retrieval performance, the measurement traffic is exchanged with a production server, such as a web server, but this requires careful design to avoid overloading that server with measurement traffic.) Since we are interested in large-scale measurements, we assume that a measurement controller provides the measurement client with information on what to measure and when to perform the measurements. Finally, in some cases, a measurement data collector gathers data, typically samples rather than aggregate data, collected by the measurement clients for later analysis. The data models and file formats for supporting the exchange of the test parameters as well as test results require standardization.

As noted above, it appears likely that metrics will evolve and new ones will be added over time. Components of the platform may be designed and operated by different, independent entities, or, at minimum, data gathered by the platform may be used by different parties for different purposes. For example, a regulator or ISP might contract with third parties to manage various components of a measurement effort, and all data communications must securely support the delegation and authentication of rights and responsibilities to perform any operational parameter supported by the measurement architecture. Thus, it will be important to agree to on a set of metrics and associated metric-specific protocol parameters. For example, the TCP throughput metric defined in [6] depends on the TCP congestion avoidance algorithm. Each measurement run generates one or more data samples, e.g., a set of throughput values. The controller needs to convey those parameters to the measurement client and the data collector needs to be able to determine unambiguously which parameters were used for a specific set of data samples.

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 [1]. Although RFC 2119 was written with protocols in mind, the key words are used in this document to indicate the strength of a requirement.

2. Use Cases

Large-scale, automated measurements are helpful in a number of use cases. We illustrate the scope with three examples:

Provider network measurements: Internet service providers have an interest in knowing how well their networks are performing, as viewed from their customers' perspective. Such performance information allows them to identify bottlenecks and observe the impact of changes in user behavior, e.g., the emergence of new network applications or time-of-day patterns. Here, the provider is not interested in the performance of an individual edge network or device, but rather wants to get a statistically-valid sample of performance across their network. Service providers may be interested in both the end device performance, i.e., the performance as seen by edge devices in home and enterprise networks, as well as the edge performance, i.e., as seen by the network device directly attached to their network, such as a cable modem, DSL modem or enterprise edge router. To reduce the network load, providers are unlikely to gather measurements from all clients all the time, but rather sample randomly across both time and their user population. The measurement controller directs the measurement client what measurements are to be performed, what measurement servers to use, when to measure and at which data collector it should deposit the measurement data.

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

Multi-provider network measurements: As an extension of the first use case, multiple network providers and third parties, such as a regulatory body, may collaborate to gather network performance data on a one-time or recurring basis, using a subset of customers of the service providers. The form of collaboration is beyond the scope of this paper, however it should be understood that a data collection platform must serve multiple stakeholder interests.

In the description above, the network provider can either be a

commercial or not-for-profit entity distinct from the network edge users, or it can be the information technology department in a local area network. Particularly for the user diagnostics use case, it may be helpful for the measurement client to obtain parameters of their connectivity, such as the nominal uplink and downlink speed. In other cases, only the entity performing the data analysis may need to know the nominal performance parameters.

3. Architecture Overview

We define a measurement platform to consist of one or more measurement clients, measurement controllers and data collection servers. Based on the use cases above, we summarize their functions below.

3.1. Measurement client

The measurement client is the reference point for measurements. For active measurements, it sends measurement traffic to the measurement server or other network elements. For passive measurements, it observes network performance metrics. Client measurement functionality must be implementable in a variety of user contexts and provide for communications within different network segments, such as the access link between a broadband subscribers modem and an ISP network, as well as consumer electronic device communicating to measurement server features in a wireless LAN device.

3.2. Measurement server

The measurement server is only needed for active measurements that require two network nodes. The measurement server typically operates as a traffic source or sink. To allow scaling, different clients within a measurement platform may use different measurement servers. Clients may also select, for example, the closest measurement server if the influence of wide-area connectivity on measurement results is to be minimized.

3.3. Measurement controller

The measurement controller provides the measurement client with instructions on when and how to conduct what measurements, i.e., the measurement schedule. For example, it might instruct the client to conduct a particular kind of throughput measurement every ten minutes, and to deposit the throughput samples into a particular data collector. Measurement controllers may be capable of accepting inputs from other controllers, scaling up the scope of the measurement system. As one example, an ISP operating a testing platform for its own network may accept test requests from an external controller as part of a nationwide testing program that it is participating in.

3.4. Data collector

The data collector collects time-stamped measurement samples from measurement clients. It generally makes these measurement samples available only to authorized users. The data collector may store

measurement samples in a database or as files and may make them available via download or SQL query. Access control, internal data storage and access methods to data are beyond the scope of this document.

We logically separate the data collector from the measurement server for both functional and performance reasons. In general, data collected should not be transferred to the collector while a measurement is in progress. Also, a measurement client on a mobile host may decide to delay transferring measurement data until a low-cost or high-speed connection to the server becomes available.

3.5. Network parameter server

In some of the use cases, it is necessary for the analysis to compare the measured against the nominal network performance, or correlate measured parameters with the type and key parameters of the user's network connection. For example, for evaluating network delay measurements, it is helpful to know what kind of access technology (e.g., FTTP, DSL, cable, cellular data or satellite) and nominal speed the network connection offers.

4. Protocols

With the description of the elements above and the relationships between them, a set of protocols needs to be defined. The key functions of the protocols are described briefly below.

Measurement client to measurement server: Each metric will have its own set of measurement protocols, and these are beyond the scope of this document. For example, a VoIP metric may use a defined set of UDP packets to estimate performance.

Measurement client to measurement controller: The measurement client queries the measurement controller to obtain an updated measurement schedule. The measurement schedule returned by the controller indicates the type of measurements the measurement client should perform, the measurement servers and on what schedule to conduct the measurements. For example, it might indicate to run a VoIP emulation test every day for ten minutes to a specific server, spanning a one-week measurement campaign. The collector also indicates one or more addresses of data collectors to the client.

Measurement controller to measurement controller: A measurement controller can request that another controller undertake a specific testing program and could indicate specific tests, schedules and sample parameters appropriate to the intended objectives. Other data could include the identity and identity verification of the requester, a specific test identifier, e.g. Nationwide Test XX, and information necessary for the data collector so that data is accessible to authorized parties.

Measurement client to data collector: The measurement client will typically perform one or more measurements, and then, during the pause between measurements, transmit the collected samples to the data collector. The samples must be tagged with identifying information, such as when they were collected, edge device information (e.g., the mobile device or cable modem) and which measurement host was used. For mobile measurements, the sample data is likely to contain location data, possibly of reduced spatial resolution to protect user privacy.

Measurement client to network parameter server: The measurement client may query the network parameter server, typically located in the service providers network, for information about its nominal service parameters, based on its network address, link layer address, or hardware identifiers such as the IMEI for mobile nodes. The data returned may include information such as nominal uplink and downlink speeds, data quotas and physical and data link

layer technology. (Data quotas may be important for deciding which data-intensive measurements a client wishes to run.)

While basic network connection information is unlikely to change rapidly, it may change at unpredictable instants. For example, a network provider may upgrade the connection speed of subsets of their customers, customers may change their subscription or provider may adjust the monthly data transfer quota.

We assume that the measurement server, controller and data collector cooperate in configuring appropriate parameters. For example, the controller needs to be able to determine which measurement servers and data collectors are currently available and the client is authorized to use. Discovery of suitable data collectors is considered beyond the scope of this effort.

5. Initiation of Measurements

Either the client or the measurement controller could in principle initiate measurements. For periodic measurements or one-off user-triggered diagnostics, it is sufficient for the end system to contact the controller, e.g., periodically every week. Client-initiated measurements have a number of advantages. In particular, they make it less likely that measurement hosts can be abused to generate denial-of-service traffic. They also avoid problems allowing inbound requests through network address translators (NATs) and firewalls.

However, there may be cases where the network provider wishes to initiate a one-time measurement or change the measurement parameters before the client next contacts the controller. For such cases, a publish-subscribe mechanism may be considered, where the measurement client subscribes to measurement schedule updates with the measurement controller.

6. Requirements

We distinguish requirements for the different component by a prefix: Requirements labeled A-* describe the overall platform architecture, M-* indicate requirements primarily affecting the measurement client, C-* those for the controller, D-* for the data collector and N-* for the functions necessary to obtain network parameter. In many cases, a single requirement governs more than one entity or protocol, so the labeling should be considered rough.

A-1: The architecture **MUST** allow for one-time measurements initiated by end users, sampled measurements initiated by network providers and measurements by one or more third parties.

A-2: Measurement clients and servers **MUST** support an extensible set of performance metrics.

A-3: Measurement clients, measurement servers and data collectors **MAY** be operated by different administrative entities, including entities other than the Internet service provider.

A-4: Measurement clients **MUST** be able perform both active and passive measurements.

A-6: All entities **MUST** be able to authenticate the entities they communicate with.

A-7: Each measurement sample **MUST** be unambiguously associated with the measurement parameters, either by reference or by value.

A-8: To ensure availability and scaling, implementations **MUST** be able to implement multiple measurement controllers, measurement servers and data collectors with appropriate load balancing and failover.

M-1: The architecture **MUST** allow a single measurement client to participate in one or more independent measurement platforms.

M-2: A measurement client **SHOULD** be able to automatically switch from a non-responsive to an alternate measurement server.

M-3: A measurement client **MUST** be able to register with the data collection platform automatically, announcing its availability and relevant system parameters. (For example, a cable or DSL modem may indicate its make and model number.)

- M-4: A measurement client MUST be able to declare what kind of measurements it can perform, e.g., by enumerating a set of measurement identifiers.
- C-1: The measurement system MUST support measurements that are scheduled according to a pre-defined calendar.
- C-2: The measurement controller MUST be able to specify the interval on how often it wishes to be contacted for updated measurement schedules.
- C-3: A measurement client SHOULD be able to automatically discover controllers provided by their Internet service provider.
- C-4: A measurement client MUST be able to authenticate and authorize the measurement controller.
- C-5: The data exchange between the client and controller MUST allow for optional encryption and integrity protection.
- D-1: The protocol messages for measurement samples MUST allow new measurement types and parameters.
- D-2: It MUST be possible to protect the integrity and confidentiality of the measurement data exchanged between the measurement client and the data collector.
- D-3: The data exchange protocol between measurement server and data collector SHOULD allow the definition of common data elements, e.g., for network addresses and timestamps.
- D-4: The measurement client SHOULD be able to automatically fail over to alternate data collectors.
- D-5: Clients MUST be able to either send data immediate or delay sending measurement data to the collector, e.g., to use a low-traffic period or a low-cost network.
- D-6: Clients MUST be able to interleave data samples from different measurement metrics to the data collector.
- D-7: The data collector SHOULD be able to ascertain whether the measurement client clock is at least approximately synchronized to its own.

- D-8: The data exchange between measurement client and data collector MUST be subject to flow and congestion control.
- D-9: The measurement client MUST be able to ascertain that it is initiating a session with the desired data collector rather than an impostor.
- N-1: Measurement clients SHOULD be able to obtain nominal network service parameters in a machine-readable format, such as advertised speed and typical latency. (This may not be necessary in all measurement use cases.)
- N-2: The set of network parameters MUST be extensible in a backward-compatible manner.
- N-3: The measurement client SHOULD be able to determine the network parameter server without manual configuration.
- N-4: The protocol between measurement client and network parameter server SHOULD support a variety of client identifiers, such as network addresses, link-layer addresses, AAA identifiers or hardware identifiers.
- N-5: The data exchanged between the network parameter server and the measurement client SHOULD ensure its confidentiality and integrity.
- N-6: The protocol SHOULD support suitable authentication functionality to restrict access to network parameters to authorized nodes. Authorized nodes may include third parties, such as data collectors.
- N-7: The entity querying the network parameter server MUST be able to assure itself that it is communicating with an authentic server.
- N-8: Clients of the network parameter server SHOULD be able to be automatically informed of changes in parameters.

7. Security Considerations

The large-scale measurement architecture has to prevent third parties' use of the measurement clients in bot-nets or for other nefarious or malicious purposes. A malicious third party could cause a measurement client to initiate probe traffic to victim hosts rather than measurement servers. We rely on user-initiated requests, secured with transport-layer security and server certificates, to ensure that only user-authorized entities issue control commands. Users may also authenticate themselves via local shared secrets. We note that there are similarities in approach with M2M data communications and we suggest that reference of ongoing work on the M2M signaling gateway framework or other models may be useful.

Measurements may also inadvertently expose information that the owner of the measurement client considers privacy-sensitive. Privacy considerations may differ depending on whether the measurement client, measurement server or data collector are operated by the same entity or not, and what trust relationships these entities have with each other. It must be possible to protect the confidentiality of the measurement data exchanged between the measurement client and the data collector. For mobile measurements, location information is likely to be crucial to interpreting measurement results. A measurement client may want to substitute rough location [8] to reduce the ability of a third party to track its movements and whereabouts.

8. IANA Considerations

This document does not request any IANA actions.

9. Acknowledgements

The document is based on discussion within the FCC Measuring Broadband America project.

DISCLAIMER: The opinions expressed are those of the author and do not necessarily represent the views of the Federal Communications Commission or the United States Government

10. References

10.1. Normative References

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

10.2. Informative References

- [2] Almes, G., Kalidindi, S., and M. Zekauskas, "A Round-trip Delay Metric for IPPM", RFC 2681, September 1999.
- [3] Federal Communications Commission, "Measuring Broadband America", September 2012.
- [4] Broadband Forum, "Liaison Statement: New Project - Broadband Access Service Attributes and Performance Measures", August 2012.
- [5] Almes, G., Kalidindi, S., and M. Zekauskas, "A One-way Delay Metric for IPPM", RFC 2679, September 1999.
- [6] Mathis, M. and M. Allman, "A Framework for Defining Empirical Bulk Transfer Capacity Metrics", RFC 3148, July 2001.
- [7] Claise, B., "Cisco Systems NetFlow Services Export Version 9", RFC 3954, October 2004.
- [8] Barnes, R. and M. Lepinski, "Using Imprecise Location for Emergency Context Resolution", Internet draft draft-ietf-ecrit-rough-loc-05, July 2012.

Authors' Addresses

Henning Schulzrinne
Federal Communications Commission - See Disclaimer
445 12th Street SW
Washington, DC 20554
USA

Phone: +1 202 418 1544
Email: henning.schulzrinne@fcc.gov

Walter Johnston
Federal Communications Commission - See Disclaimer
445 12th Street SW
Washington, DC 20554
USA

Phone: +1 202 418 0807
Email: walter.johnston@fcc.gov

James Miller
Federal Communications Commission - See Disclaimer
445 12th Street SW
Washington, DC 20554
USA

Phone: +1 202 418 7351
Email: James.Miller@fcc.gov

LMAP
Internet-Draft
Intended status: Informational
Expires: August 22, 2013

J. Seedorf
NEC
V. Gurbani
Bell Labs, Alcatel-Lucent
E. Marocco
Telecom Italia
February 18, 2013

ALTO for LMAP
draft-seedorf-lmap-alto-00

Abstract

In the context of Large-Scale Measurement of Broadband Performance (LMAP), measurement results are currently made available to the public either at the finest granularity level (e.g. as a list of results of all individual tests), or in a very high level human-readable format (e.g. as PDF reports).

This document argues that there is a need for an intermediate way to provide access to large-scale network measurement results, flexible enough to enable querying of specific and possibly aggregated data. The Application-Layer Traffic Optimization (ALTO) Protocol, defined with the goal to provide applications with network information, seems a good candidate to fulfill such a role.

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 August 22, 2013.

Copyright Notice

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

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

Table of Contents

1. Introduction	3
2. Example Use Cases	5
3. Advantages of using ALTO	6
4. Examples	7
4.1. Download speeds	7
4.1.1. Network map	8
4.1.2. Cost map	9
5. References	10
5.1. Normative References	10
5.2. Informative References	10
Appendix A. Acknowledgment	11
Authors' Addresses	12

1. Introduction

Recently, there is a discussion on standardizing protocols that would allow measurements of broadband performance on a large scale (LMAP [I-D.schulzrinne-lmap-requirements]). In principle, the vision is that "user networks gather data, either on their own initiative or instructed by a measurement controller, and then upload the measurement results to a designated measurement server."

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

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

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

2. Example Use Cases

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

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

3. Advantages of using ALTO

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

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

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

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

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

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

4. Examples

[NOTE: syntax most certainly wrong!]

4.1. Download speeds

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

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

4.1.1. Network map

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

      "MSMNT-CL1" : {
        "ipv4" : [ "192.168.0.0/30" ]
      },
      "TOTAL" : {
        "ipv4" : [ "0.0.0.0/0" ]
      }
    }
  }
}
```

4.1.2. Cost map

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


5. References

5.1. Normative References

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

5.2. Informative References

- [I-D.ietf-alto-protocol]
Alimi, R., Penno, R., and Y. Yang, "ALTO Protocol",
draft-ietf-alto-protocol-13 (work in progress),
September 2012.
- [I-D.schulzrinne-lmap-requirements]
Schulzrinne, H., Johnston, W., and J. Miller, "Large-Scale
Measurement of Broadband Performance: Use Cases,
Architecture and Protocol Requirements",
draft-schulzrinne-lmap-requirements-00 (work in progress),
September 2012.

Appendix A. Acknowledgment

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

Authors' Addresses

Jan Seedorf
NEC
Kurfuerstenanlage 36
Heidelberg 69115
Germany

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

Vijay K. Gurbani
Bell Labs, Alcatel-Lucent

Email: vkg@bell-labs.com

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

Email: enrico.marocco@telecomitalia.it

