Internet Engineering Task Force
Internet-Draft
Intended status: Informational
Expires: June 23, 2016

M. Ooki
S. Kamei
NTT Communications

December 21, 2015

Internet Measurement System
draft-ooki-lmap-internet-measurement-system-03

Abstract

This document describes an experience of Japanese Internet measurement system to measure end-to-end performance of user's experience. We have developed the system toward the enhancement of the network performance in an ISP since October 2013. The systems and the considerations about the Internet measurement are introduced along with our current status. This document is expected to be useful for the standardization of Internet measurements.

Status of This Memo

Copyright Notice

Table of Contents

1. Introduction

In Japan, it is common to use a high speed Internet such as 100Mbps and 1Gbps as an ISP's customer connection. Users only know the maximum bandwidth of the last one mile for the ISP connection. The maximum bandwidth value is ranging from 100Mbps to 2Gbps in ISP's price plan as a FTTH connection.

Of course the end-to-end performance of actual Internet connection is below the bandwidth value. Internet users can obtain actual performance depends on various ISP conditions such as congestions. Internet users don't know the performance of the actual network.

On the other hand, ISPs also don’t know the quality that Internet users experience. For the ISP’s point of view, it is important to understand the service quality for its customers in order to design its network properly. For this reason, it is necessary to measure the actual performance of typical Internet users.

The Large-Scale Measurement of Broadband Performance (LMAP) working group is formed to standardize a large scale measurement system to measure broadband network performance. The current LMAP WG focus on the information model, data model language, the protocols in a certain ISP network. However, the LMAP WG does not focus on the measurement of the global end-to-end performance at the moment. We believe that either way someday it will be necessary to establish a method for the Internet measurement and the standardization of the end-to-end performance measurement, that is not closed to a certain ISP.

This document describes the Internet measurement system and our considerations for the end-to-end measurement. Our measurement requirements can be useful for LMAP framework. We have measured the end-to-end performance by using Internet measurement system we have been operating since 2011. We expect the experience of our case can contribute to the standardizations in LMAP WG and the enhancement of network operation from ISP’s perspective.

2. Motivation of Internet Measurement

The LMAP WG describes some use cases for the Large-scale Measurement of Broadband Performance [RFC7536][I-D.deng-lmap-collaboration]. There are three reasons that we, ISPs, need to measure the end-to-end performance of user experience of its access services.

First, ISPs want to keep the customer satisfaction. Typically ISPs provide the list of maximum bandwidth and the service prices, such as the estimated total fee and the discount rate after the result of a cash back campaign. Japanese users select a ISP based on only those information without knowing the end-to-end performance results. The poor performance causes the lower customer satisfaction.

On the other hand, the Ministry of Internal Affairs and Communications in Japan discuss the ideal measurement methods of the end-to-end performance about the mobile network in the research society. The organization has yet to be discussed fixed network. The researchers are planning to measure the mobile network performance on the 1,500 measurement points in the main Japanese area. They use the tool Federal Communications Commission developed for measurement of the end-to-end performance. The mobile network operator in Japan may be required to publish the actual performance in addition to the best effort performance.

Second, contents providers are beyond the control of ISPs. The traffic volume of Contents Delivery Network (CDN) providers such as AKAMAI and LEVEL3 is increasing in the Internet in recent years. How much users are connecting to which contents providers impact the end-to-end performance. ISPs need to understand their behavior to decide ISP’s strategies and operation.

Third, we would like to support the public evaluations of ISPs. Some contents service provider e.g., Google or Sandvine[google][sandvine], presented the reports about Internet traffic and ISP performance based on each criterion recently. The Google report presents the results of multiple ISPs measured in for each locations in USA. People in the world can browse the reports on the Internet. These reports will have huge impact on user’s choice of the ISP selection. We would like to double-check by using our performance data in order to confirm whether the reports can be reliable or not. If we can find the difference of the performance results between our data and the reports, we might be able to review whether our measurement methods are mistaken or not. It is also better for an ISP to investigate and comprehend the status of end-to-end performance between ISPs. So, we have to measure the end-to-end performance by ourself.

Hence, ISPs should measure the end-to-end performances from end users to multiple content providers accurately while comparing with other ISPs’ performance. Then, ISPs can show a performance of the actual network to build brand value compared with other providers.

3. Framework of Internet Measurement System

We introduce the framework of Internet Measurement System in this section. The words, such as Measurement Agent, Controller, and Collector conforms to the glossary of the LMAP document[I-D.ietf-lmap-framework].

3.1. Measurement Agent

The MA has the functions that receive instructions from Controller Server (described below), performs measurement tasks, and sends the measurement results to the Collector Server.

3.1.1. Specification of the MA

We used a Japanese product, called OpenBlocks [plathome], which is the Linux box with Dual Core Marvell ARMADA XP 1.33GHz, 1GB SDRAM memory. We selected the box as the MA because of the affordable price, software stability, small form factor, flexible functionally, and extendability. The MA needs some CPU power in order to connect PPPoE access line and download tens of contents on the Internet. The OpenBlocks stacks CPU enough to archive them.

3.1.2. Configuration

We introduce information configured on the MA in this section.

- MA’s ID

We have to setup the MA’s ID. The ID has to be a unique among MAs in order for Controller server to distinguish MAs. The information is described in the "/etc/hostname" on Linux File System. The naming rule is based on the location of MA, the types of line, and the plan, etc.

On the other hand, MA doesn’t have a group ID. The ID in our measurement system is under only Controller Server.

- HTTP Get Tasks

MA automatically gets the measurement tasks from Controller Server every five minutes. MA sends the request about the tasks and schedule to Controller Server by HTTP. Controller Server returns the tasks decided based on SCHEDULE Table to the MA.

- Convert measurement results

MA automatically converts from raw data to the type of JSON data. JSON is a lightweight data-interchange format. We selected it as our measurement data model language because of readability, simple format, and easy data cleaning to analyze the measurement data. On the other hand, the LMAP WG selected YANG Data Model [I-D.ietf-lmap-yang]. We need to consider which we should use YANG data model or not. The convert process is executed every one minute. The converted JSON data is written in the measurement result files. An example of JSON data type of the results of wget raw data is as follows.

{"host":"tokyo-xXx01","filename":"tokyo-xXx01_ISP_target_wget_20141203235011.log.ok","result":"ok","line":"2","message":"2014-12-03 23:50:18 (10.1 MB/s) - ‘/dev/null’ saved [67206439/67206439]"}

o Data Collector and Submitting

MA automatically and efficiently collects and submits the measurement data. To make it realized, we selected the fluentd which is an open source data collector, The software lets Collector Server unify the data collection and understanding of data. The reason we selected it is that the software is reliable, stable, and simple of implementation enough to control hundreds of MAs. The data collecting tool of many products and system in Japan is implemented by fluentd. The above measurement result files are submitted to Collector Server by the process as soon as the files are created. An example of the running flutend process is as follows.

```
/usr/bin/ruby1.9.1 /usr/local/bin/fluentd --daemon /var/lib/fluent/
fluentd.pid --user fluent --group fluent --config /etc/fluent/
fluent.conf --log /var/log/fluent/fluent.log -vv
```

o Self Check

The MAs check whether the above processes for measurement service are down or not on the regular interval. If a process is down, the MA transmits the message about the message logs to Controller Server.

3.1.3. Location of the MA

We have distributed MAs on many places all over Japan. The number of locations is approximately 150 in June 2015. The number of our MAs will be increasing in approximately 200 by the end of this year. MAs are located in houses where the residents can respond our requests (e.g., not turning off the power to constantly perform the measurement) to manipulate the device.

3.2. Controller Server

The Controller Server is a Linux server. The Controller Server has the functions to instruct the MAs and receives the HTTP GET requests from MAs. The Controller Server has 3 tables of database implemented by MySQL to instruct MAs.

3.2.1. Control of MAs

The Controller Server manage MAs by using two tables (MA Table and GROUP Table). A example of the MA Table is as follows. MA_ID is a key identifying MA. TYPE is a kind of network type. MODE expresses the type of the measurement.IF MODE is 0, it means MA is the measurement mode. The mode is the status that MAs are performing a measurement task. If MODE is 1, it means MA is the maintenance mode. The mode is the status that MA stop performing a measurement task.

In case of MODE 1, the MA automatically connect to Controller Server by using ssh protocol. We can login to the MA of MODE 1 from Controller Server and change the configuration. We can manage the behavior for MAs by switching the MODE.

MA_ID	TYPE	AREA	OS_TYPE	MODE	GET_SCHEDULE_TIME
tokyo-nFh04	flets	tokyo	Debian 7	1	2014-12-08 23:21:00
osaka-nFs01	flets	osaka	Debian 7	1	2014-12-08 23:22:00

Table 1: MA Table

A example of the GROUP Table is as follows. The GROUP_ID is a key record grouping MAs. MAs are sure to belong one group at least. MA, tokyo-nFh04, belongs to the group-id1. The GROUP_INFO is the remarks. We can set the information of the group which MAs belong in the column

GROUP_ID	MA_ID	GROUP_INFO
group-id1	tokyo-nFh04	Group 01
group-id2	osaka-nFs01	Group 02

Table 2: GROUP Table

3.2.2. Control of the Assigned ISP

We set the information of ISP accounts to assign MAs on the Controller Server. MAs automatically download the information from Controller Server. Of course, assigned ISP account is unique of all ISP accounts for the measurements. The table includes the column of multiple assigned information so that the duplicate use does not happen. A example of the ISP Table is as follows. ASSIGN_ID is a key record within the table. MEASURE_ISP is the service name of ISP. ISP_ID is the unique ID to connect to the ISP network. PASSWORD is the password of the ISP_ID. ASSIGN_STATUS, ASSING_MA, and ASSIGN_TIME are the assigned information at that time. If the ASSIGN_STATUS is 1, that means a MA use the ISP_ID.

ASSIGN_ID	MEASURE_ISP	ISP_ID	PASSWORD	ASSIGN_STATUS	ASSIGN_MA	ASSIGN_TIME
1	OCN	abc123@ocn.ne.jp	abc123def	1	tokyo-nFh04	2014-12-08 23:21:05
2	OCN	ghi456@ocn.ne.jp	ghi456jkl	0		

Table 3: ISP Table

3.2.3. Setting the Measurement Task and Measurement Schedule

We set the measurement tasks to instruct MAs on the Controller Server. MAs automatically download the task from Controller Server by the fixed time. We need to set a measurement schedule with the measurement task at the same time. A example of the MEASUREMENT SCHEDULE Table is as follows. SCH_ID is a key record within the table. LINE_TYPE is a type of network provided by a network service. ISP is a Internet service provider to perform the measurement tasks. SCRIPT is the script file of the measurement tasks described by some programming languages. PARAM is a parameter file required for performing measurement tasks. START_TIME is the time when MAs start performing a measurement task. END_TIME is the time when MAs stop performing a measurement task.

SCH_ID	GROUP_ID	LINE_TYPE	ISP	SCRIPT	PARAM	START_TIME	END_TIME
1	group-id1	flets	ISP1	measure1.sh	param1	00:00:00	00:00:00
2	group-id2	flets	ISP2	measure2.sh	param2	00:00:00	00:00:00

Table 4: MEASUREMENT SCHEDULE Table

3.2.4. Receiving the Requests

On the Controller Server, a httpd program is running as a daemon that executes continuously in the background to handle requests. The Controller Server returns the appropriate measurement tasks and measurement schedules to MAs in response to HTTP GET requests. The MA which complete own measurement task receives a new measurement task continuously. MAs can start performing the next measurement tasks continuously.

3.3. Collector Server

The Collector Server receives the measurement results from MAs through fluentd process. The fluentd process is running as a daemon that executes continuously in the background to handle the measurements data. The details of the measurement results received by fluentd process are listed below.

20141214230628+0900 measure.tokyo-nFh04 {"host":"tokyo-nFh04","filename":"tokyo-nFh04_ISP_DEST_wget_20141214230450.log.ok","result":"ok","message":"2014-12-14 23:06:18 (745 KB/s) - ‘/dev/null’ saved [67206439/67206439]"}

20141214230902+0900 measure.tokyo-nFh04 {"host":"tokyo-nFh04","filename":"tokyo-nFh04_ISP_DEST_wget_20141214230731.log.ok","result":"ok","message":"2014-12-14 23:08:52 (811 KB/s) - ‘/dev/null’ saved [67206439/67206439]"}

A example of the directory structure of the stored measurement results is as follows. A MA’s measurement result file is created by the day.

/data/MA’s-ID/measure_result_MA’s-ID.DATETIME.log

When the Collector Server receives the measurement results, the server creates the directory of the MA-ID of MA and the measurement result files. The measurement results are stored in the data directory.

3.4. Architecture

The architecture of the measurement system is composed of MAs, Controller Server, and, Collector Server.

```
   +-----------------+                +-----------------+
   |   Controller    |                |    Collector    |
   |                 |                |                 |
   |     Server      |                |     Server      |
   +-----------------+                +-----------------+
          ^                                       ^
          |                                       |
          | Get Measurement              Sending  |
          | Tasks and Measurement    the Results  |
          | Schedules                             |
          |            +-------------+            |
```

```
             |            | Measurement |            |
             +----------  |             | -----------+
                          |   Agents    |
                          +-------------+
                            |        ^
               Perform Only |        | Obtain
         Active Measurement |        | Results
                            v        |
      +---------------------------------------------------+
      |                                                   |
      |                  FTTH Access Line                 |
      |                                                   |
      +---------------------------------------------------+
                            |       ^
                            |       |
                            |       |
                            v       |
      +---------------------------------------------------+
      |                                                   |
      |                    ISP network                    |
      |                                                   |
      +---------------------------------------------------+
                            |       ^
                            |       |
                            |       |
                            v       |
      +---------------------------------------------------+
      |                                                   |
      |                      Internet                     |
      |                                                   |
      +---------------------------------------------------+
                            |       ^
                            |       |
                            |       |
                            v       |
                     +----------------------+
                     |   Target Contents    |
                     +----------------------+
```

Figure 1: Architecture of the Internet Measurement System

We need to import Controller Server the record of MA’s configuration in MA Table, GROUP Table, and MEASUREMENT SCHEDULE Table before the MA is powered on.

When a MA is powered on, it tries to establish the FTTH access PPPoE connection with the ISP. After obtaining an IP address, it automatically connect the Controller Server and gets the configuration by HTTP. If the value of MODE column in MA Table is 1, the MA automatically gets the maintenance mode. If the value of that is 0, the MA automatically gets the measurement mode and start downloading the measurement tasks that is configured in MEASUREMENT SCHEDULE Table.

The MA prepares for the measurement tasks, performs the tasks for Measurement Target Contents actively, and collects the measurement results.

After the completion of the measurement tasks, the MA sends the measurement results to Collector Server using fluentd process. and submits the request for downloading the next measurement tasks to Controller Server using HTTP.

When the specification of the LMAP WG’s protocol and framework is finished, we will deploy the protocol in our measurement system.

4. Operation of Internet Measurement System

We introduce the operation of Internet Measurement System we have been operating since October 2013 in this section.

4.1. Measurement Performance Metrics

The MAs perform only active measurements for Target Contents. Examples of the Target Contents include Video Streaming files, OS update files, and the test server for performance measurement in a local ISP network.

In our measurements, the measurement performance metrics are as follows.

o Round Trip Time (RTT)

This is the response time between the submission of the ICMP echo request packet and the reception of the ICMP echo reply packet. The metric can also be regarded as the round-trip delay time. This is measured by the ping command. We take the min/avg/max time and the loss rate based on measuring this metrics one hundred times by the measurement task. An example of the command MAs execute is as follows.

```
MA-ID $ ping -i 0.05 -c 100 {contents_ip_address}
```

o Hop Count and Network Path

This metric refers to the number of intermediate devices (like routers) through which the data must pass between the MA and the Target Contents. This metric is regarded as the network distance between the MA and the Target Contents. This is measured by the traceroute command. Actually, MA submits the ICMP echo requests three times. Afterwards, by checking the hop counts and network path, we can find the change of the network routing on the Internet. An example of the command MAs execute is as follows.

```
MA-ID $ traceroute -nIq 3 {contents_ip_address}
```

- Throughput

This metric refers to how much data can be transferred from the MA to the Target Contents in a given amount of time. This is regarded as the bandwidth. We can understand how fast we can get the contents the on the network. Currently, this is measured by the wget command. A MA receives URL of the measurement targets and start downloading the contents using HTTP GET. When the download is completed, the value divided the contents size by the download complete time is regarded as the performance metrics of throughput. An example of the command MAs execute is as follows.

```
MA-ID $ wget -T 300 -dvO /dev/null {wgetopts} {contents_url}
```

In addition to above three performance metrics, we are studying the change of destination IP address of the Internet contents distributed by some contents service providers. It is important for ISPs to know the mechanism of the contents delivery networks. MAs resolve some FQDN and gets the destination IP address. We are studying the mechanism of the contents delivery networks based on the response results.

4.2. Measurement Target Contents

The selection of the Target Contents is important for the Internet measurement; the type, the length, and the number of the contents. We need to measure the representative contents on the Internet. In order to find such contents, we have selected contents based on two viewpoints.

One is the volume of transferred data of network traffic in an ISP. We obtained partial traffic data on multiple prefectures in Japan. We selected the Target Contents which were higher in the transferred traffic volume ranking. Examples of such target Contents are Youtube Video Streaming files and Mac OS update file on AKAMAI and so on.

Second is the number of access to the contents in the Internet. For example, the portal sites such as Google or Yahoo!, etc. and the shopping sites such as Amazon and iTunes, etc. are always the higher in the number of access to the contents in the Internet.

In another viewpoint, we need to change the target contents according to the purpose of analysis. If our purpose is to measure an event traffic, e.g., the download traffic concerning iOS update or the access traffic concerning the special winning sale of the professional baseball team, etc., we need to measure the related contents.

4.3. Measurement Schedule

On receiving the measurement tasks from Controller Server, MAs start measurement tasks. MAs used to perform the measurement task by thirty minutes. When the measurement completes, MAs wait the next scheduled time and do not perform the next measurement tasks. On the contrary, when the measurement does not complete before the next scheduled time, the MA kills the measurement process and moves to the next measurement. The current system in that point is flexible because the MA can start the next task as soon as a measurement task is completed. We can collect more kinds of data than before.

4.4. Applications of Measurement

Using the data collected by our measurement system we have studied how to comprehend and analyze end-to-end performance more accurately than ever before. A example of analysis is the difference of network performance between Japanese ISPs, based on the combinations of Target Contents, measurement time, and areas. We took the measurement results in consideration to ISP network design and ISP operation as a reference information.

Furthermore, we have studied the analysis method based on the combination of our customer feedbacks, remarks on social network service, and customer voices on our callcenters in addition to the collected measurement data. As the result of combined analysis, we expect to find new and useful knowledge we have never found before.

5. Issues of Internet Measurement System

We introduce the issues of Internet Measurement System we have been operating in this section. The issue section is divided into three parts: Architecture Issue, Operation Issue, and Security Issue.

5.1. Architecture Issue

o Scalability

The Controller Server is connected to receive HTTP GET requests from multiple MAs. This means that the Controller Server needs to process as many HTTP GET requests as the number of MAs. The number of MAs can easily grow beyond the number of HTTP GET requests that a Controller Server can process. If we place hundreds of MAs all over Japan, we will need to improve the scalability of our system.

o IPv6 Support

IPv6 network is constructed totally independently from IPv4 network. Hence, the performance of the IPv6 network is highly likely different from that of the IPv4 network.

Although the IPv6 network is not the majority yet, it is growing. NTT EAST and WEST provided only 2.7% in NGN (Next Generation Network) on December 2013. The rate of IPv6 enabled network in Japan is 27% in June 2014[IPv6-Promotion Council]. NTT EAST and WEST presented the IPv6 support in PPPoE connection on March 2014. All CPE devices for NTT access line already support IPv6 tunneling, allowing users to adapt IPv6 easily.

In order to achieve the broad applicability of our measurement results, we will need to investigate the IPv6 performance also.

o Data Reliability

We need many kinds of data in order to improve the reliability of data analysis. If there are many kinds of data, the reliability of our analysis results will be improved and the analysis results might be statistically significant. We need to develop the architecture to collect as many different types data as possible. When a MA completed all the instructed measurement tasks, we are creating the measurement system that the MA performs other measurement tasks being high priority as soon as possible when a MA completed all the instructed measurement tasks.

5.2. Operation Issue

o Selection of the Products as Measurement Agent

We used a Japanese product, called OpenBlocks. However, some issues happen by using this product. The product sometimes generates high heat (e.g., a certain hot day). The heat is hot enough to feel like getting burned. In fact, the maximum degree of the heat reaches about eighty degrees Celsius. Some people were afraid to set up MA in their house. We are looking for more efficient product satisfying our requirements than this model.

- Selection of Measurement Target Contents

It is difficult to decide what contents should be measured to present the representative performance. There are many kinds of contents on the Internet.

This time we have selected the Target Contents based on the volume of transferred data at some points in an ISP. However, there are more metrics to consider, such as the number of accesses to that contents, rather than the transferred volume. Other metrics are not studied in this document.

- Stable Operation

We had experiences where the measurement results were not sent immediately, and the measurements for some Target Contents were failed. Although the actual causes of these difficulties vary (e.g., accidentally disconnected LAN cable or power cable), we could easily respond to those issues using informations (e.g., time and place) contained in the centralized logs in the Collector Server. Another difficulty is the change in the settings of the contents provider. For example, wget command for a video content has not worked due to a change in a setting in the contents provider. This issue is difficult to tackle and is left for future work.

5.3. Security Issue

- Measurement between ISPs and CSPs

If we continuously measure the performance about the contents in the Internet, from the point of the contents service provider, it can be obstacle to provide the stable service due to the traffic volume for measurement. However, we need real situations for the customer to measure the performance as correctly as possible. The service provider may request the limitation, e.g., volume and the number of access, for measurement to MA. In terms of the combination between ISPs and CSPs, we may need the condition for measurement.

- DDoS Attack

We placed approximately 150 MAs all over Japan. These MAs may become DDoS attackers by wrong commands from the Controller Server. From this reason, the list of commands MAs can perform should be restricted. And also, the MAs must deny illegal accesses and logins.

MAs should permit only access through instruction from the Controller Server.

6. Security Considerations

As described in Setion 5.3, security consideration for Internet measurement must be considered.

7. IANA Considerations

No need to describe any request regarding number assignment.

8. References

8.1. Normative References

[RFC7536] Linsner, M., Eardley, P., Burbridge, T.,and Sorensen, F., "Large-Scale Broadband Measurement Use Cases, May 2015", .

[I-D.deng-lmap-collaboration]
Deng., L., Huang, R.,and Duan, S., "Use-cases for Collaborative LMAP, draft-deng-lmap-collaboration-05 (work in progress), June 2015", .

[I-D.ietf-lmap-framewark]
Eardley, P., Morton, A., Bagnulo, M., Burbridge, T., Aitken, P., and Akhter, P., "A framework for large-scale measurement platforms(LMAP),draft-ietf-lmap-framework-14 (work in progress), April 2015", .

[I-D.ietf-lmap-yang]
Schoenwaelder, J.,and Bajpai, V., "A YANG Data Model for LMAP Measurement Agents, draft-ietf-lmap-yang-00 (work in progress), April 2015", .

8.2. URL References

[MIC] MIC Report, "http://www.soumu.go.jp/main_sosiki/kenkyu/speed_measurement/", .

[sandvine]
Sandvine Report, "https://www.sandvine.com/pr/2014/5/14/sandvine- report-netflix%E2%80%99s-british-invasion.html", .

[google] Google Report, "http://www.google.com/get/videoqualityreport/", .

[plathome]
OpenBlocks,
"http://openblocks.plathome.com/products/ax3/", .

[IPv6-Promotion-Council]
Japanese IPv6-Promotion,
"http://v6pc.jp/jp/spread/ipv6spread_02.phtml", .

Authors' Addresses

Motoyuki Ooki
NTT Communications
GranPark 16F
3-4-1 Shibaura, Minato-ku, Tokyo
108-8118,Japan

EMail: m.ooki@ntt.com

Satoshi Kamei
NTT Communications
GranPark 16F
3-4-1 Shibaura, Minato-ku, Tokyo
108-8118,Japan

EMail: skame@nttv6.jp