LMAP
Internet-Draft
Intended status: Informational
Expires: April 24, 2014

J. Seedorf
NEC
D. Goergen
R. State
University of Luxembourg
V. Gurbani
Bell Labs, Alcatel-Lucent
E. Marocco
Telecom Italia

October 21, 2013

ALTO for Querying LMAP Results
draft-seedorf-lmap-alto-02

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. Finally, we describe our methodology for analyzing the United States Federal Communication Commission's (FCC) Measuring Broadband America (MBA) dataset to derive required topology and cost maps suitable for consumption by an ALTO server.

Copyright Notice

Table of Contents

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: One the one hand, highly aggregated results for certain networks may be made available in the form of PDFs of 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-seperated-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:

- 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.

- 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 or may not be employed depending on the average connectivity type/speed of a user in a given region or with a given access network provider.

- Display estimation of service quality or total download time to users: A webservice could use statistics about average download speeds for a given ISP and/or region to estimate Quality-of-Service for provided services (e.g. to indicate to the user what Quality-of-Experience to expect when clicking on a given link) or to estimate (and display to the user) the total download time for given content.

- 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.

- TBD: add more use cases

3. Advantages of using ALTO

The ALTO protocol [I-D.ietf-alto-protocol] specifies a very lightweigth 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 infromation 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 bottled-up links increasingly 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 asynchrounous 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:

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

- 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. Discussion of Useful ALTO Extensions

The base ALTO Protocol as specified in [I-D.ietf-alto-protocol] can in principle be used to enable a more flexible way to provide access to large-scale network measurement results as discussed in the previous sections of this document. However, certain extensions to the base ALTO Protocol that have recently been proposed in the ALTO WG would allow to better enable the use cases discussed in Section 2:

- Server-initiated Notifications: In [I-D.marocco-alto-ws], it has been proposed to enhance the ALTO protocol such that servers can notify clients about newly available ALTO maps. In the context of this document, this extension would allow applications to be notified when certain new LMAP measurements are available, such as new measurement results on average download speeds. These new results could then be downloaded and used immediately by applications.

- Incremental Updates: In [I-D.schwan-alto-incr-updates], it has been proposed to enhance the ALTO protocol with incremental updates, such that clients can retrieve partial updates for ALTO maps instead of always downloading a full ALTO map (even when only a small fraction of the ALTO map has changed compared to a previous version). When ALTO is used for querying LMAP results, the corresponding ALTO maps may potentially be quite large (e.g. when a webservice queries for particular, detailed results regarding a whole ISP). In this case, incremental ALTO updates would be a very useful mechanism for applications to retrieve updates of ALTO maps, as a reduced amount of data would be needed for transmitting these maps.

6. Case study: Analyzing a large-scale dataset

Measuring broadband performance is increasingly important as communications continue to move towards the Internet. Internet service providers (ISP), national agencies and other entities gather broadband data and may provide some, or all, of the dataset to the public for analysis. As we argue above, there are two extremes prevalent for presenting large-scale data. One is in the form of charts, figures, or summarized reports amenable for easy and quick consumption. The other extreme includes releasing raw data in the form of large files containing tables formatted as values separated by a delimiter. While the former is indispensable to acquire a summary view of the dataset, it does not suffice for additional analysis beyond what is presented. Conversely, the problem with the latter option (raw files) is that the unsuspecting user perusing them is lost in the deluge of data.

We offer the argument that a reasonable medium between the two extremes may be the ALTO protocol [I-D.ietf-alto-protocol]. A necessary prerequisite for using ALTO is abstracting the network information into a form that is suitable for consumption by the protocol. The implication of using ALTO is that data from any large-scale measurement effort must first be distilled in two maps: a topology map and a cost map. Further analysis and ad-hoc queries can be subsequently performed on the normalized dataset.

In the United States, the Federal Communication Commission (FCC) has embarked on a nationwide performance study of residential wireline broadband service [fcc]. Our aim is to use the raw datasets from this study for analysis and to create a topology map and a cost map from this dataset. ALTO queries aimed at these maps will enable users and interested parties to fulfill the use cases listed in Section 2.

6.1. Challenges in data analysis

The FCC Measuring Broadband America (MBA) study consisted of 7,782 volunteers spread across the United States with adequate geographic diversity. Volunteers opted in for the study, however, each of the volunteers remained anonymous. An opaque integral number (unit_id) represented a subscriber in the raw dataset. This unit_id remains constant during the duration of the study in the dataset and uniquely identifies a volunteer subscriber, even if the subscriber switches the ISP. More detail about the methodology used is described in [fcc].

The dataset consisted of 12 tables, each table corresponding to the data drawn from a certain performance test. For the analysis we

present in this document we focus on the "curr_dns" table, which contains the time taken for the ISP's recursive DNS resolver to return a DNS A RR for a popular website domain name. This test was ran approximately every hour in a 24-hour period, and produced about 75-78 million records per month. This resulted in a typical file size in the range of 6-7 GBytes per month. We note that the "curr_dns" table is one of the smaller tables in the dataset.

The first challenge, therefore, was to arrive at computing resources comparable in scale with respect to the dataset consisting of millions of records spread across gigabyte-sized files. To analyze the volume of data we used a canonical Map-Reduce computational paradigm on a Hadoop cluster (more details on the methodology are outlined in Section 6.2).

A second, more pressing challenge, was to identify the geographic location of the unit_ids generating the data. In order to derive a topological map and impose costs on the links, it is important to know the physical locations of the unit_ids that contributed the measurements. However, in the MBA dataset, the population is anonymized and the individual subscriber reporting the measurement data is simply referred to by an opaque integral number. Therefore, an important task was to use the information in the public tables to reveal a coarse location of the subscriber.

We outline the methodology we used to do so in the next section. We stress that this methodology does not identify the specific location of a subscriber, who still remains anonymous. Instead, it simply locates the subscriber in a larger metropolitan region. This level of granularity suffices for our work.

6.2. Geo-locating the units

To geo-locate the units, we simply note that broadband subscriber devices are likely to be configured using DHCP by their ISP. Besides imparting an IP address to the subscriber device, DHCP also populates the DNS name servers the subscriber devices uses for DNS queries. In most installations, these DNS name servers are located in close physical proximity of the subscriber device. The FCC technical appendix states that the DNS resolution tests were targeted directly at the ISP's recursive resolvers to circumvent caching and users configuring the subscriber device to circumvent the ISP's DNS resolvers. Therefore, a reasonable approximation of a subscribers geo-location could be the geographic location of the DNS name server serving the subscriber. We use this very heuristic to geo-locate a subscriber.

Thus our first, and very simple filter consisted of obtaining a

mapping from a unit_id (representing a subscriber) to one or more DNS name servers that the unit_id is sending DNS requests to. It turned out that while this was a necessary condition for advancing, it was not a sufficient one. The raw data would need to be further processed to reduce inconsistencies and remove outliers. A number of interesting artifacts were uncovered during further processing of the data. These artifacts informed the selection of the unit_ids for further analysis.

The artifacts are documented below.

- A handful of unit_ids were geo-located in areas outside the contiguous United States, such as Ukraine, Poland or the United Kingdom. We theorize that the subscribers corresponding to the unit_ids geo-located outside the contiguous United States had simply configured their devices to use alternate DNS servers, probably located outside the United States. We removed these records before conducting our analysis.

- We also observed a reasonable number of non-ISP DNS resolvers, especially Google's 8.8.8.8 and 8.8.4.4 and OpenDNS 208.67.222.222 and 208.67.220.220. These 4 public DNS servers are geo-located in California. We removed these records to ensure that the specific location that these resolvers represented was not oversampled.

- We noticed that a large number of unit_ids were being geo-located in Potwin, Kansas. Intrigued as to why there appeared to be a large population of Internet users being located in a small rural community in Kansas, we investigated further. It appears that Potwin, Kansas is the geographical center of the United States and a number of ISPs have chosen to establish data centers in or around the Potwin area. These ISPs generally locate their primary or secondary DNS name servers in Potwin-area data centers, thus accounting for the popularity of Potwin as an Internet destination. We continue to further investigate on minimizing the impact of such natural aggregation points that, if not accounted for, will skew our results in an unwarranted direction.

- We observed some unit_ids changing ISPs during the observation period. This is a normal occurrence and to the extent that the unit_id is geo-located in the same geographical area after the change in ISP, we do not exclude such unit_ids from further analysis.

Subsequent filters extracted the stable unit_ids from our dataset. In order to determine which unit_id are stable, i.e., remain constant with respect to their geographic location over the observation period from January to December 2012, we extracted for each unit_id the IP

address of each DNS name server it consulted. This is obtained by applying the map reduce paradigm on the DNS dataset. We extracted for each unit_id the triggered DNS servers and obtained the individual DNS servers accessed by a unit_id. This was repeated for each month of the observation period. The resulting sets were cleaned up of private IP addresses and other artifacts discussed above. The cleaned set consisted of about 8000 distinct unit_id.

In order to determine the stability of each unit_id we proceeded to sum up the occurrences of IP addresses over the whole observation period separated in monthly files. If the IP address of a DNS server occurred 12 times this meant that the unit_id always accessed the same DNS server and therefore remained stable over the observation period. The obtained stable unit_ids, around 1500, will be used for further analysis. Assuming a 99% confidence level and +/- 3 point margin of error, we will require a sample of 1494 unit_ids. With our stable unit_id set of 1500 unit_ids, we are now positioned to perform further analysis on the dataset to create the full topology and cost maps.

Table 1 presents a sample of the geographic location data that we have uncovered for unit_ids. A complete list of identified units superimposed on the geographical map of the United States is available at http://cdb.io/13UOHgD.

Unit ID	City, State	Latitude/Longitude
872	Morganville, NJ	40.35950089,-74.26280212
885	Madison, WI	43.07310104,-89.40119934
898	Foley, AL	30.40660095,-87.68360138
7969	Manteca, CA	37.79740143,-121.2160034
8024	Quincy, MA	42.25289917,-71.00229645

Sample unit identification tuples

Table 1

7. Security considerations

There are no security artifacts invalidated due to our analysis in Section 6. All of our analysis was performed on publicly available data. However, we do note that some privacy may have been lost based on our analysis. In the raw dataset, the unit identifiers are opaque strings with no immediate correlation with a geographic location. After our analysis, while the unit identifiers still remain opaque, they are nonetheless correlated to a specific, though coarse, geographic location.

8. IANA considerations

This document does not contain any IANA considerations.

9. Conclusion

This document argues that, compared to existing solutions, there may be a need for a more flexible way to provide access to large-scale network measurement results. Further, the document argues that the ALTO protocol is a good candidate to enable querying for specific, possibly aggregated, measurement results in a flexible way. Examples of how such a flexible query meachnism for large-scale measurement results could look like based on ALTO are given.

With respect to the case study in Section 6, identification of the geographic location of the unit_ids generating the performance data is essential in order to continue the work. We have presented a methodology and some early results in identifying a geographic location. This location, although coarse, suffices for our future work that will consist of further data mining and analysis to create appropriate ALTO network and cost maps.

10. References

10.1. Normative References

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

10.2. Informative References

[I-D.ietf-alto-protocol]
Alimi, R., Penno, R., and Y. Yang, "ALTO Protocol", draft-ietf-alto-protocol-20 (work in progress), October 2013.

[I-D.marocco-alto-ws]
Marocco, E. and J. Seedorf, "WebSocket-based server-to-client notifications for the Application-Layer Traffic Optimization (ALTO) Protocol", draft-marocco-alto-ws-01 (work in progress), July 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.

[I-D.schwan-alto-incr-updates]
Schwan, N. and B. Roome, "ALTO Incremental Updates", draft-schwan-alto-incr-updates-02 (work in progress), July 2012.

[fcc] United States Federal Communications Commission, "Measuring Broadband America", Accessed July 12, 2013, http://www.fcc.gov/measuring-broadband-america.

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

David Goergen
University of Luxembourg

Email: david.goergen@uni.lu

Radu State
University of Luxembourg

Email: radu.state@uni.lu

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