Network Working Group
Internet-Draft
Intended status: Experimental
Expires: August 18, 2014
S. Randriamasy, Ed.
Alcatel-Lucent Bell Labs
N. Schwan
February 14, 2014

ALTO Cost Schedule
draft-randriamasy-alto-cost-schedule-03

Abstract

The goal of Application-Layer Traffic Optimization (ALTO) is to bridge the gap between network and applications by provisioning network related information. This allows applications to make informed decisions, for example when selecting a target host from a set of candidates. The ALTO problem statement [RFC5693] considers typical applications as file sharing, real-time communication and live streaming peer-to-peer networks. Recently other use cases focused on Content Distribution Networks and Data Centers have emerged.

The present draft proposes to extend the cost information provided by the ALTO protocol. The purpose is to broaden the decision possibilities of applications to not only decide 'where' to connect to, but also 'when'. This is useful to applications that have a degree of freedom on when to schedule data transfers, such as non-instantaneous data replication between data centers or service provisioning to end systems with irregular connectivity. The draft therefore specifies a new cost mode, called the "schedule" mode. In this mode the ALTO server offers cost maps that contain path ratings that are valid for a given timeframe (e.g. hourly) for a period of time (e.g. a day). Besides the functional time-shift enhancement providing multi-timeframe cost values, the ALTO Cost Schedule also allows to save a number of ALTO transactions and thus resources on the ALTO server and clients. Last, guidance to schedule application traffic can also efficiently help for load balancing and resources efficiency.

Requirements Language

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

Status of This Memo

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

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

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

This Internet-Draft will expire on August 18, 2014.

Copyright Notice

Copyright (c) 2014 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

IETF is currently standardizing the ALTO protocol which aims for providing guidance to overlay applications, that need to select one or several hosts from a set of candidates that are able to provide a desired resource. This guidance is based on parameters that affect performance and efficiency of the data transmission between the hosts, e.g., the topological distance. The goal of ALTO is to improve the Quality of Experience (QoE) in the application while simultaneously optimizing resource usage in the underlying network infrastructure.

The ALTO protocol therefore [ID-alto-protocol] specifies a Network Map, which defines groupings of endpoints in a network region (called a PID) as seen by the ALTO server. The Endpoint Cost Service and the Endpoint (EP) Ranking Service then provide rankings for connections between the specified network regions and thus incentives for application clients to connect to ISP preferred endpoints, e.g. to reduce costs imposed to the network provider. Thereby ALTO intentionally avoids the provisioning of realtime information (cmp. ALTO Problem Statement [RFC5693] and ALTO Requirements [RFC5693]), as "Such information is better suited to be transferred through an in-band technique at the transport layer instead". Thus the current Cost Map and Endpoint Cost Service are providing, for a given Cost Type, exactly one rating per link between two PIDs or to an Endpoint. Applications are expected to query one of these two services in order to retrieve the currently valid cost values. They therefore need to plan their ALTO information requests according to the estimated frequency of cost value change. In case these value changes are predicable over a certain period of time and the application does not require immediate data transfer, it would save time to get the whole set of cost values over the period in one ALTO response and using these values to schedule data transfers would allow to optimise the network resources usage and QoE.

In this draft we introduce use cases that describe applications that have a degree of freedom on scheduling data transfers over a period

of time, thus they do not need to start a transfer instantaneously on a retrieved request. For this kind of applications we propose to extend the Cost Map and Endpoint Cost Services by adding a schedule on the cost values, allowing applications to time-shift data transfers.

In addition to this functional ALTO enhancement, we expect to further gain by gathering multiple Cost Values for one cost type as firstly one Cost Map reporting on N Cost Values is less bulky than N Cost Maps containing one Cost value each and secondly, this reduces N ALTO transactions to a single one. This is valuable for both the storage of these ALTO maps and their transfer. Similar gains can be obtained for the ALTO Endpoint Cost Service.

The remainder of this draft first provides use cases that motivate the need for a 'schedule' cost mode. It then specifies the needed extensions to the ALTO protocol and details some example messages. Note that the example ALTO transactions are provided with the ALTO syntax as specified in previous ALTO protocol draft versions, see [ID-alto-protocol-13]. The syntax will be updated when the base ALTO protocol will be finalized.

2. Use cases for ALTO Cost Schedule

This section introduces use cases showing the benefits of providing ALTO Cost values in 'schedule' mode. Most likely, the ALTO Cost Schedule would be used for the Endpoint Cost Service where a limited set of feasible non real time application Endpoints is already identified, they need to be accessed neither simultaneously nor immediately and their access can be scheduled within a given time period. The Filtered Cost Map service is also applicable as long as the size of the Map is manageable. An ALTO Cost schedule can be used in several ways:

- the ALTO Server may provide values on past time periods that can be interpreted as historical experience and used to anticipate future cost values in order to schedule transfers of application data or services,

- the ALTO Server may provide stationary values on present or future time periods that can be interpreted as predictions on cost values and used to schedule transfers of application data or services,

- the ALTO Server may provide stationary values on time periods covering the past, present and future and logically be all interpreted as predictions and used to schedule transfers of application data or services.

2.1. Bulk Data Transfer scheduling

Some CDNs are prepopulating caches with content before it actually gets available for the user and thus there is a degree of freedom on when the content is transmitted from the origin server to the caching node. Other applications like Facebook or YouTube rely on data replication across multiple sites for several reasons, such as offloading the core network or increasing user experience through short latency. Typically the usage pattern of these data centers or caches follows a location dependent diurnal pattern.

In the examples above, data needs to be replicated across the various locations of a CDN provider, leading to bulk data transfers between datacenters. Scheduling these data transfers is a non-trivial task as the transfer should not infer with the user peak demand to avoid degradation of user experience and to decrease billing costs for the datacenter operator by leveraging off-peak hours for the transfer. This peak demand typically follows a diurnal pattern according to the geographic region of the datacenter. One precondition to schedule transfers however is to have a good knowledge about the demand and link utilization patterns between the different datacenters and networks.

While this usage data today already is gathered and also used for the scheduling of data transfer, provisioning this data gets increasingly complex with the number of CDN nodes and in particular the number of datacenter operators that are involved. For example, privacy concerns prevent that this kind of data is shared across administrative domains. The ALTO Cost Schedule specified later in this document avoids this problem by presenting an abstracted view of time sensitive utilization maps through a dedicated ALTO service to allow CDN operators a mutual scheduling of such data transfers across administrative domains.

2.2. Endsystems with limited connectivity or access to datacenters

Another use case that benefits from the availability of multi-timeframe cost information is based on applications that are limited by their connectivity either in time or resources or both. For example applications running on devices in remote locations or in developing countries that need to synchronize their state with a data center periodically, in particular if sometimes there is no connection at all. Example applications is enterprise database update, remote learning, remote computation distributed on several data center endpoints.

Wireless connectivity has a variable quality or may even be intermittent. On the other hand, the connectivity conditions are

often predicable. For non real time applications, it is thus desirable to provide ALTO clients with routing costs to connection nodes (i.e. Application Endpoints) over different time periods. This would allow end systems using ALTO aware application clients to schedule their connections to application endpoints.

Another challenge arises with end systems using resources located in datacenters and trading content and resources scattered around the world. For non-real time applications, the interaction with Endpoints can be scheduled at the time slots corresponding to the best possible QoE. For instance, resource Ra downloaded from Endpoint EPa at time t1, Resource Rb uploaded to EPb at time t2, some batch computation involving Ra and Rb done on EPc at time t3 and results R(A,B) downloaded to EPd and EPe at time t4. Example applications are similar to the ones cited in the previous paragraph.

```
  +-----+                                           +-----+
  | EPa |                                           | EPb | <----- Rb
  +-----+                                           +-----+   (t2=50)
     |                   +-------+                     |
     Ra -------------->  | EPc   |                     |
     (time t1=10)        |       |                     |
                         |t3=100 |  <----------------- Rb
                         +-------+
                             | \
                             |  \
                           R(Ra,Rb)
                          (t4=200)
                             |     \
                             |      -------------------.
                             V                         V
                          +-----+                   +-----+
                          | EPd |                   | EPe |
                          +-----+                   +-----+
```

These examples describe situations where a client has the choice of trading content or resources with several Endpoints and needs to decide with which Endpoint it will trade and at what time. For instance, one may assume that the Endpoints are spread over different time-zones, or have intermittent access. The ALTO Schedule mode specified below allows these clients to retrieve Endpoint cost maps valid for a certain timeframe (e.g. 24 hours), and get a set of values, each applicable on a (e.g. hourly) slot. Thus the application can optimize the needed data transfer according to this information.

Last, the ALTO Cost schedule is beneficial to optimizing ALTO transactions themselves. Indeed, let us assume that an Application Client is located in an end sytem with limited resources and/or has an access to the network that is either intermittent or provides an acceptable QoE in limited but predictable time periods. In that case, it needs to both schedule its resources demanding networking activities and its ALTO requests. Instead of having to figure out when the cost values may change and having to carefully schedule multiple ALTO requests, it could aviod this by relying on Cost Shedule attributes that indicate the time granularity, the validity and time scope of the cost information, together with the time related cost values themselves.

Suppose that for some Cost Types, the ALTO cost values are available in the "schedule" mode. If the values of Cost type 'routingcost' and /or another time-sensitive Cost Type named for example 'pathoccupationcost' are available in the "schedule" mode for the 24 hours following the last update, the ALTO Client embedded in the Application Client may query ALTO information on 'routingcost' or 'pathoccupationcost' for these 24 hours, and get a set of values, each applicable to an hour slot. If appropriate Cost Attributes are provided together with the cost values, the Application client also knows the date of their last update. An example ALTO transaction is provided later in this draft.

2.3. SDN Controller guided access to application endpoints

The Software Defined Networking (SDN), see [sdnrg], is a model that attempts to manage and reconfigure networks in a more flexible way in order to better cope with the traffic challenges posed by nowadays resources greedy applications. To this end, one option is "moving the control plane out of the network elements into "controllers", see [SDN charter, http://www.1-4-5.net/~dmm/sdnrg/sdnrg.html], that implements the network control and management. The SDN Controllers are deemed to gather the network state information and provide it in an abstracted form to SDN aware applications while gathering their requirements in QoE and exchanging other application "management" information and commands.

The relevance of ALTO to perform a number of SDN functions has been recently highlighted. An ALTO Server can assist an SDN Controller by hosting abstracted network information that can be provided to SDN aware applications via an ALTO Client. It can also assist other SDN Control operations using information in and outside the ALTO scope.

In particular, [article-gslh-alto-sdn] identifies SDN Controller functions that ALTO is well suited to perform: the primitives of Abstraction, Get network topology, Get network resources and Event

notification. Additonally, the interaction between ALTO and SDN has been investigated in [draft-xie-alto-sdn] to provide applications with a path selection meeting QoS requirements on bandwidth and delay.

Currently, the base ALTO protocol allows to perform the following SDN services, see [article-gslh-alto-sdn]:

1. Abstraction: through aggregation into PIDs, ranking and a generic cost type.

2. Get network topology: through the Map and the Cost Map Services

3. Get device capabilities: through the Endpoint Property Service.

Another SDN primitive "Get network resources" provides applications with informations allowing them to evaluate the expected QoE. QoE related information includes delay and bandwidth at the application endpoints as well as on the network paths. Such information may be provided via the ALTO Service by proposed extensions of the ALTO protocol that define new ALTO Cost Types allowing to abstract and report QoE to applications.

One key objective of an SDN controller is the ability to balance the application traffic whenever possible. For non real time applications, data and resources transfer can be time shifted, resources availability may often be predicable and last, strong incentives for applications to time shift their traffic may be given by network operators appropriately setting routing cost values at different time values, according to their policy to cope with network occupation over time.

To achieve this objective, the SDN controller can:

1. get the network state history from its controlled network elements through its southbound API

2. possibly derive an estimation or a prediction of these values over given time frames

3. store their abstraction in an ALTO Server in the form of ALTO Cost Schedule values defined for different time periods

4. deliver these values to the SDN applications via the ALTO Endpoint Cost Service, either as history or prediction or as estimations covering both the past and the future.

This way:

o One one hand, the applications get the best possible QoE, as they can pick the best time for them to access one or more Endpoints,

o One the other hand the SDN controller achieves load balancing as it may guide the application traffic so as to better distribute the traffic over time, and thus optimize its resources usage.

2.4. Large flow scheduling on extended ALTO topologies

[draft-yang-alto-topology-00] presents initial thinking on extending ALTO for topology exposure services, that would provide flexible abstractions based on the raw network topology. Among other features, an ALTO topology may expose several paths between a source (src) and destination (dst), or topology details may be provided on restricted parts. This work was presented to the ALTO WG at IETF88.

The presentation slides [slides-88-alto-5-topology] on [draft-yang-alto-topology-00] expose a use case entitled "Large Flow Scheduling". This case includes a "daylife example" where a Google Map service proposes multiple routes between 2 points A and B, each calculated w.r.t. length and estimated time. For each of these selected paths, the map service exposes a time-sensitive qualitative value taking 4 values between Slow and Fast. A user of this application may thus organize its transfer w.r.t. metrics, paths and time, provided s/he does not have to commute immediately.

The use case on Large flow scheduling on extended ALTO topologies in the present section illustrates one modality of ALTO topology service, that would expose several paths between end to end (src, dst) pairs, computed w.r.t. one of more metrics, possibly under given constraints. On top of this enriched topology service, non real-time applications may also choose the time of data/resources transfer, taking thus advantage of a richer set of decision variables.

The use case "Large Flow Scheduling" of presentation [slides-88-alto-5-topology]can thus be adapted as follows:

o Step1 - obtain the set T transfer tasks {(src, dest, data)}

o Step2 - identify one or more paths for each (src, dst): several information sources exist among which:

   * (a) ALTO CostMap with a "path" metric, // not specified here

   * (b) an ALTO Topology Service providing a path computation hint (e.g. w.r.t. routingcost and/or other metrics)

o Step 3- while T not empty:

* 1 - query for example values for some metric 'available bandwidth' on paths:

  + to this end, query the values in the ALTO 'schedule' Mode: on the selected (src, dst) for a set of time intervals. With this mode, the ALTO client will receive an array of values, each applicable to a time slot .

* 2 - schedule data transfer at the time slots corresponding to the preferred value.

2.5. Providing values for time-sensitve TE metrics

Draft [draft-wu-alto-te-metrics-01], proposes to extend the set of ALTO metrics with traffic engineering (TE) metrics, in order to closely meet applications requirementsand under appropriate trust agreements. This draft exposes a number of TE metrics that are time-sensitive, either by nature such as bandwidth and delay related metrics, or due to "normally" changing network conditions or both.

The draft assumes that the values of ALTO TE metrics are typically collected from routing protocols and provided in a non-real time manner. In "normally" changing network conditions TE metric values remain uniformly distributed over given time intervals and can be aggregated over bigger sample intervals of periodic patterns. For instance an ALTO Server may collect values from a routing protocol produced by measurements done every second over a period of 30 seconds. The ALTO Server may then aggregate these values over ALTO Sample intervals of 60 seconds and every hour, provide the values in 'schedule' mode endcoded as an array of 60 values.

All time sensitive ALTO TE metrics are potentially applicable to a cost schedule. The real applicability actually depend on the network topology and structure. A small topology with low density and capacity that carries inpredictable heavy and bursty traffic has few chances to exhibit stationary TE metric value patterns over large periods and would benefit to use the ALTO Schedule over smaller time slots.

3. ALTO Cost Schedule extension

One example of non-real time information that can be provisioned in a 'schedule' is the expected path bandwidth. While the transmission rate can be measured in real time by end systems, the operator of a data center is in the position of formulating preferences for given paths, at given time periods of given time scales, for example to avoid hotspots due to diurnal usage patterns. The entity managing the ALTO Server values can decide to integrate path bandwidth in the

ALTO 'routingcost' metric. However to better highlight the purpose of the cost schedule the remainder of this document will use a Cost Type named 'pathoccupationcost' and assumed to report an abstracted form of available bandwidth. A definition and usage of such a Cost-Type is proposed in [draft-randriamasy-multi-cost-alto].

The usage of a time related cost is rather proactive in that it can be used like a "time table" to figure out the best time to schedule data transfer and also anticipate predictable events including predictable flash crowds. An ALTO Cost Schedule should be viewed as a synthetic abstraction of real measurements that can be historic or be a prediction for upcoming time periods.

3.1. Cost Schedule Attributes

Specifications on the cost "schedule" are proposed here and will be completed in further versions of this draft.

3.1.1. ALTO Cost-Mode: Schedule

The "schedule" mode applies to Costs that are eligible for a single-valued Cost Mode and can also be expressed as such. In that sense, when the "numerical" mode is available for a Cost-Type, the cost expressed in the "schedule" mode is an extension of its expression from one value in the "numerical" mode to an array of several values varying over time.

Types of Cost values such as JSONBool can also be expressed in the "schedule" mode, as states may be "true" or "false" depending on given time periods. It may be expressed as a single value which is either "true" or "false" following a decision rule outside the ALTO protocol.

3.2. ALTO Capability: Cost-Scope

To ensure that the application client uses the NP provided information in the cost schedule in an unambiguous way we define the Cost Scope capability, which defines the validity of the "scheduled" cost values.

For Cost Types whose values are provided in a mode different than 'schedule', the Cost Scope capability is specified by the string "permanent". The Cost Scope attributes provided for the 'schedule' mode are listed below. The reference time zone for the provided values is UTC.

- o Unit: expresses the time interval applicable to each value. A two element array where the first element is the time unit, ranging

from "second" to "year", and the second one the number of units of this duration. For example: '["minute", 5]' means that each value is provided on a time interval lasting 5 minutes.

- Size: the number of values of the cost schedule array,
- Begin: the index of the first unit in the array,
- Reference time zone: set to "UTC",
- Next update: the date at which the sample will be re-computed,
- Last update: the last re-computation date.

The reference time zone is UTC.

Attributes 'Last update 'and 'Next update' report on the update frequency and age of the information.

3.2.1. Example of time scope for a cost schedule

Let us assume that the metric 'pathoccupationcost' (POC for short) is computed for 24 hours, on time intervals lasting 2 hours, with the first interval starting at 0h00. The ALTO Server thus provides an array 12 values. This information is then used to enable applications to see which time intervals in a day are the most favorable to operate, and which "busy " time intervals should be avoided. If the "Begin" date is past, the application can also use the information to compute statistics or infer a some customized prediction.

3.3. Example of scheduled information resources in the IRD

The example IRD given in this Section includes 2 particular URIs:

- "http://alto.example.com/endpointcost/lookup", in which the ALTO Server offers several Endpoint Cost Types, including a Cost called "pathoccupationcost" for which the "schedule" Cost Mode is available. The Endpoint Costs available are the "hopcount", "routingcost" and "pathoccupationcost" Cost Types, with the two first ones in the "numerical" Cost Mode and "pathoccupationcost" in the "schedule" Cost Mode.

- "http://custom.alto.example.com/endpointcost/schedule/lookup", in which the ALTO Server provides the 'routingcost' in both "numerical" and "schedule" modes. This resource is accessible via a separate subdomain called "custom.alto.example.com". The ALTO Client may either get the last update of the 'routingcost' value

or request for a previsonal sample of 24 values established each for 1 hour. An ALTO Client can discover the services available at "custom.alto.example.com" by successfully performing an OPTIONS request to "http://custom.alto.example.com/endpointcost".

```
GET /directory HTTP/1.1
   Host: alto.example.com
   Accept: application/alto-directory+json,application/alto-error+json


HTTP/1.1 200 OK
   Content-Length: [TODO]
   Content-Type: application/alto-directory+json

   {

      ... usual ALTO resources ...

    "resources" : [
      {
         "uri" : "http://alto.example.com/endpointcost/lookup",
         "media-types" : [ "application/alto-endpointcost+json" ],
         "accepts" : [ "application/alto-endpointcostparams+json" ],
         "capabilities" : {
           "cost-constraints" : true,
           "cost-modes" : [ "numerical", "numerical", "schedule" ],
           "cost-types" : [ "routingcost", "hopcount", "pathoccupationcost" ],
           "cost-scope":  [ "permanent", "permanent",
                            {"unit": ["hour", 1], "size": 24, "begin": 0,
                             "time zone": "UTC",
                             "lastupdate": mm/hh/dd/mm/yyyy,
                             "nextupdate": mm/hh/dd/mm/yyyy}
           ]
         },
      {
         "uri" : "http://custom.alto.example.com/endpointcost/schedule/lookup",
         "media-types" : [ "application/alto-endpointcost+json" ],
         "accepts" : [ "application/alto-endpointcostparams+json" ],
         "capabilities" : {
           "cost-constraints" : true,
           "cost-modes" : [ "numerical", "schedule" ],
           "cost-types" : [ "routingcost", "routingcost" ],
           "cost-scope":  [ "permanent",
                            {"unit": ["hour", 1], "size": 24, "begin": 0,
                             "time zone": "UTC",
                             "lastupdate": mm/hh/dd/mm/yyyy,
                             "nextupdate": mm/hh/dd/mm/yyyy}
           ]
         }
       }
     ]
   }
```

3.3.1. Example scenario and ALTO transaction with a Cost Schedule

Let us assume an Application Client located in an end sytem with limited resources and having an access to the network that is either intermittent or provides an acceptable quality in limited but possibly predictable time periods. Therefore, it needs to both schedule its resources demanding networking activities and minimize its ALTO transactions.

The Application Client has the choice to trade content or resources with a set of Endpoints of moderate 'routingcost', and needs to decide with which Endpoint it will trade at what time. For instance, one may assume that the Endpoints are spread on different time-zones, or have intermittent access. In this example, the 'routingcost' is assumed constant for the scheduling period and the time sentitive decision metric is the path bandwidth reflected by a Cost type called 'pathoccupationcost'.

The ALTO Client embedded in the Application Client queries ALTO information on 'pathoccupationcost' for the 24 hours following (implicitely) the date of "lastupdate", as this resource is listed in the IRD.

```
POST /endpointcost/lookup HTTP/1.1
  Host: alto.example.com
  Content-Length: [TODO]
  Content-Type: application/alto-endpointcostparams+json
  Accept: application/alto-endpointcost+json,application/alto-error+json

  {
    "cost-type" : ["pathoccupationcost"],
    "cost-mode" : ["schedule"],
    "endpoints" : {
      "srcs": [ "ipv4:192.0.2.2" ],
      "dsts": [
        "ipv4:192.0.2.89",
        "ipv4:198.51.100.34",
        "ipv4:203.0.113.45"
      ]
    }
  }


  HTTP/1.1 200 OK
  Content-Length: [TODO]
  Content-Type: application/alto-endpointcost+json

  {
    "meta" : {},
    "data" : {
      "cost-type" : ["pathoccupationcost"],
      "cost-mode" : ["schedule"],
      "map" : {
        "ipv4:192.0.2.2": {
          "ipv4:192.0.2.89"    : [7, ... 24 values],
          "ipv4:198.51.100.34" : [4, ... 24 values],
          "ipv4:203.0.113.45"  : [2, ... 24 values]
        }
      }
    }
  }
```

4. IANA Considerations

Information for the ALTO Endpoint property registry maintained by the IANA and related to the new Endpoints supported by the acting ALTO server. These definitions will be formulated according to the syntax defined in Section on "ALTO Endpoint Property Registry" of [ID-alto-protocol],

Information for the ALTO Cost Type Registry maintained by the IANA and related to the new Cost Types supported by the acting ALTO server. These definitions will be formulated according to the syntax defined in Section on "ALTO Cost Type Registry" of [ID-alto-protocol],

4.1. Information for IANA on proposed Cost Types

When a new ALTO Cost Type is defined, accepted by the ALTO working group and requests for IANA registration MUST include the following information, detailed in Section 11.2: Identifier, Intended Semantics, Security Considerations.

4.2. Information for IANA on proposed Endpoint Propeeries

Likewise, an ALTO Endpoint Property Registry could serve the same purposes as the ALTO Cost Type registry. Application to IANA registration for Endpoint Properties would follow a similar process.

5. Acknowledgements

Thank you to D. Lopez, R. Yang, H. Peng, Q. Wu and the ALTO WG for fruitful discussions.

6. References

6.1. Normative References

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

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

6.2. Informative References

[ID-alto-protocol-13]
R.Alimi, R. Penno, Y. Yang, Eds., "ALTO Protocol, draft-ietf-alto-protocol-13.txt", September 2012.

[ID-alto-protocol]
R.Alimi, R. Penno, Y. Yang, Eds., "ALTO Protocol, draft-ietf-alto-protocol-25.txt", January 2014.

[article-gslh-alto-sdn]
V. Gurbani, M. Scharf, T.Lakshman, and V. Hilt, , "Abstracting network state in Software Defined Networks (SDN) for rendezvous services, IEEE International Conference on Communications (ICC) Workshop on Software Defined Networks (SDN)", June 2012.

[draft-jenkins-alto-cdn-use-cases-01]
B. Niven-Jenkins (Ed.), G. Watson, N. Bitar, J. Medved, S. Previdi, , "Use Cases for ALTO within CDNs, draft-jenkins-alto-cdn-use-cases-01", June 2011.

[draft-randriamasy-multi-cost-alto]
S. Randriamasy, Ed., B. Roome, N. Schwan, , "Multi-Cost ALTO, draft-randriamasy-alto-multi-cost-07", October 2012.

[draft-wu-alto-te-metrics-01]
Q. Wu, Y. Yang, Y. Lee, D. Dhody, S. Randriamasy, , "ALTO Traffic Engineering Cost Metrics (work in progress)", February 2014.

[draft-xie-alto-sdn]
H. Xie, T. Tsou, D. Lopez, H. Yin, , "Use Cases for ALTO with Software Defined Networks, draft-xie-alto-sdn-extension-use-cases-00", June 2012.

[draft-yang-alto-topology-00]
Y. Yang, , "ALTO Topology Considerations (work in progress)", July 2013.

[sdnrg] "Software Defined Network Research Group, http://trac.tools.ietf.org/group/irtf/trac/wiki/sdnrg", .

[slides-88-alto-5-topology]
G. Bernstein, Y. Lee, Y. Yang, , , "ALTO Topology Service: Use Cases, Requirements and Framework (presentation slides IETF88 ALTO WG session), http://tools.ietf.org/agenda/88/slides/slides-88-alto-5.pdf", November 2013.

Authors' Addresses

Sabine Randriamasy (editor)
Alcatel-Lucent Bell Labs
Route de Villejust
NOZAY 91460
FRANCE

Email: Sabine.Randriamasy@alcatel-lucent.com

Nico Schwan

Email: ietf@nico-schwan.de