

Network Working Group
Internet Draft
Intended status: standard

Young Lee
Huawei
Greg Bernstein
Grotto Networking
Tae Sang Choi
ETRI
Sreekanth Madhavan
Huawei
Dhruv Dhody
Huawei

January 8, 2013

**ALTO Extensions to Support Application and Network Resource
Information Exchange for High Bandwidth Applications**

[draft-lee-alto-app-net-info-exchange-01.txt](#)

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/ietf/lid-abstracts.txt>

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

This Internet-Draft will expire on July 8, 2011.

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.

Abstract

This draft proposes ALTO information model and protocol extensions to support application and network resource information exchange for high bandwidth applications in partially controlled and controlled environments as part of the infrastructure to application information exposure (i2aex) initiative.

Table of Contents

1.	Introduction.....	3
2.	Problem Statement.....	5
3.	ALTO Constraints Filtering Extension Model.....	7
3.1.	ALTO Query from Application Stratum to Network Stratum....	7
3.2.	ALTO Response from Network Stratum to Application Stratum.	8
3.3.	Information Model of ALTO Query from Application Stratum to Network Stratum.....	9
3.4.	Information Model of ALTO Response from Network Stratum to Application Stratum.....	9
3.5.	ALTO Protocol Extension for Constraints Filtering Mechanism	10
4.	ALTO Protocol Extension for Graph Representation Mechanism....	11
4.1.	Representing bandwidth constraints.....	11
5.	Summary and Conclusion.....	12
6.	Security Considerations.....	12
7.	IANA Considerations.....	12
8.	References.....	13
8.1.	Informative References.....	13
	Author's Addresses.....	14
	Intellectual Property Statement.....	14
	Disclaimer of Validity.....	15

1. Introduction

This draft proposes ALTO information model and protocol extensions to support application and network resource information exchange for high bandwidth applications in partially controlled and controlled environments as part of the infrastructure to application information exposure (i2aex) initiative. The Controlled and partially controlled ALTO environments referred to here are those where general access to a specific ALTO server may be restricted to a qualified list of clients.

This draft is build upon the previously introduced High Bandwidth Use Cases [[HighBW](#)]. In [[HighBW](#)], we have discussed two generic use cases that motivate the usefulness of general interfaces for cross stratum optimization in the network core. In our first use case, network resource usage became significant due to the aggregation of many individually unique client demands. In the second use case where data centers are sending large amount of data with each other, bandwidth usage was already significant enough to warrant the use of traffic engineered "express lanes" (e.g., private line service). We introduce third use case where inter-CDN providers may want to expose controlled network resource usage information so that CDN applications (e.g., request routing server) can utilize such information when appropriate decisions (e.g., request routing redirection) are needed.

These use cases result in optimization problems that trade off computational versus network costs and constraints. Both featured use cases show the usefulness of an ALTO interface between the application and network strata in optimizing the networked applications.

In particular, this draft introduces: (i) enhanced constraints filtering extensions to the ALTO protocol to reduce extraneous information transfer and enhance information hiding from the network's perspective; (ii) constrained cost graph mechanism encoding that enables enhanced application traffic optimization that was introduced by [[HighBW](#)].

In controlled and partially controlled environments in which operators are willing to share additional network stratum resource information such as bandwidth constraints or additional cost types of topology (e.g., graph or summary), it can be useful to reduce the amount of information transferred from the ALTO server to the ALTO client.

In considering information exchange between the application stratum and the network stratum, especially from the network stratum to the application stratum, the degree of information details is one of the key concerns from the network providers' standpoint. On the one hand, the network information has to be useful to the application; on the other hand, the provided network information should hide details about the network. In order to achieve these desired goals, a simple enhancement to ALTO protocol would help significantly both in reducing/filtering the amount of information and in increasing the usefulness of the information from network to application.

Figure 1 shows ALTO Client-Server Architecture for Application-Network information Exchange. Figure 1 shows that ALTO Client in the application stratum can interface with ALTO Server in the network stratum. With this architecture, a simple ALTO query mechanism from application (via ALTO client) to network (via ALTO server) can be implemented. According to this architecture, ALTO Client is assumed to interact with the Application Orchestrator that has the knowledge of the end-user (i.e., source) application requirement, Data Center locations (i.e., destinations) and their resource information.

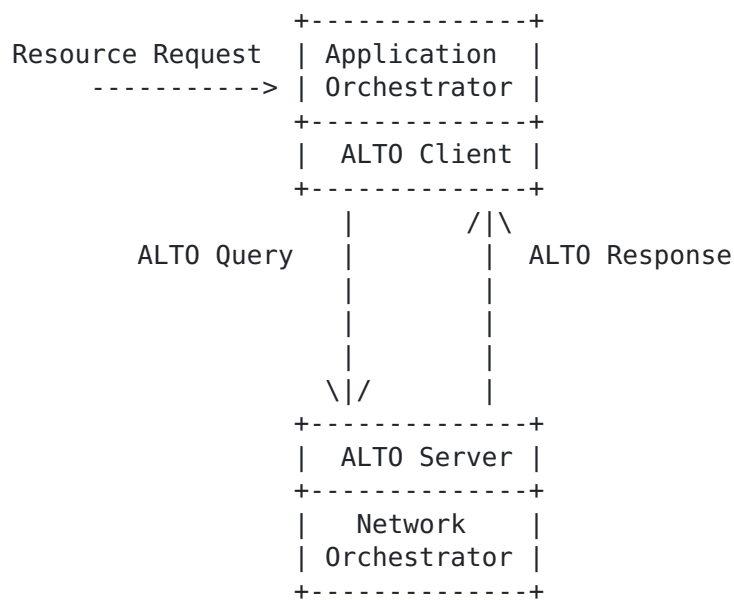


Figure 1 ALTO Client-Server Architecture for Application-Network information Exchange

The Application Orchestration functions depicted in Figure 1 interfacing data centers and end-users are out of the scope of this document. For simplicity purpose, Figure 1 doesn't depict the detailed relationship between ALTO client and server. In fact, both client and server don't need to be in the same administration boundary. It can be inter-operator and one to many relationships. For example, in the cases of inter-CDN environment or generic multi-domain environment, ALTO client represents a request routing server of upstream CDN operator and it interacts with multiple downstream CDN operators for their network resource information to make efficient decisions for desired quality CDN services. Interaction methods can either iterative or recursive. That is, ALTO client can interact with multiple ALTO servers directly or ALTO client can interact with one representative ALTO server and subsequent interaction is done by the representative one to rest of ALTO servers.

The organization of this document is as follows. [Section 2](#) discusses the ALTO architecture in the context of the application and network strata interaction. [Section 3](#) provides ALTO Information model and protocol extension to support ALTO Constraints Filtering Mechanism. [Section 4](#) provides ALTO information model and the protocol extension to support ALTO Constrained Cost Graph Mechanism.

[2. Problem Statement](#)

One critical issue in Application-Network information exchange in ALTO is the amount of information exchanged between the application and network strata. The information provided by network providers can be not so useful to the application stratum unless such information is abstracted into an appropriate level the that application stratum can understand.

In partially controlled and controlled environments, network providers can furnish appropriately abstracted and pruned information to the application stratum with the cooperation of the application stratum that can indicate some level of filtering and pruning in its query.

To reduce extraneous information this draft allows for "filtering" (or "pruning") of the following information in query/response of the ALTO pull model:

- . Topology Filtering - reduction of the results to only those specified set of source(s) and destination(s) instead of the entire network cost map. Note that this mechanism is not new in the current ALTO protocol. In the context of application-network information exchange, this topology filtering can be

- of a tremendous help in reducing the amount of data exchanged between application and network.
- . Constraint Filtering on paths or graphs (e.g., bandwidth, latency, hop count, packet loss, etc.) - reduction of results to only those that meet ALTO client specified cost bounds.

As discussed in [[HighBW](#)], in a controlled environment optimization is significantly enhanced by sharing data related to bandwidth constraints and additional cost measures [[MultiCost](#)]. Such information may be considered sensitive to the network provider just as application data, e.g., usage, demand, etc., may be considered sensitive to an application provider. [Section 3](#) provides ALTO information model and protocol extensions to support topology/constraints filtering mechanism.

While it is important to reduce and filter the information amount from network to application, for some applications that require stringent QoS objectives (e.g., bandwidth and latency), simple summary source-destination network resource information (i.e., summary form of topology) may not provide sufficient details to the application stratum. For example, suppose that a multiple number of large concurrent flows need to be scheduled from application to network. In such a case, a summary form of network topology that only shows source-destination bandwidth availability may not show the bottleneck links over which more than one flow may compete for the link bandwidth resource. This problem was indicated by [[HighBW](#)]. The following are the excerpts from [[HighBW](#)].

Consider the network shown in Figure 2, where DC indicates a datacenter, ER an end user region (as in the end user aggregation use case), N a switching node of some sort, and L a link. The link capacities and costs are also shown on the figure as well as a cost map between [ER1, ER2] and [DC1, DC2, DC3]. Since the network has a tree structure (very unusual but easier to draw in ASCII art), the cost map is unique.

As an illustration, assume that the maximum available capacity between any individual end region and a data center is 5 units (i.e., $L1=L2=L5=L6=5$). However, link L3 (capacity 8 units) represents a bottle neck to all the data centers (L3 is on all the paths to DC1, DC2, or DC3 from all end regions, ER1 and ER2).

ALTO Cost Map is shown in the lower right corner of Figure 2. This summary cost map does not provide enough details on the bottle necks. The lower left corner shows Link Capacity Cost, from which

the bottle necks can be shown such that multi-flow commodity scheduling can be made possible to avoid such bottle necks.

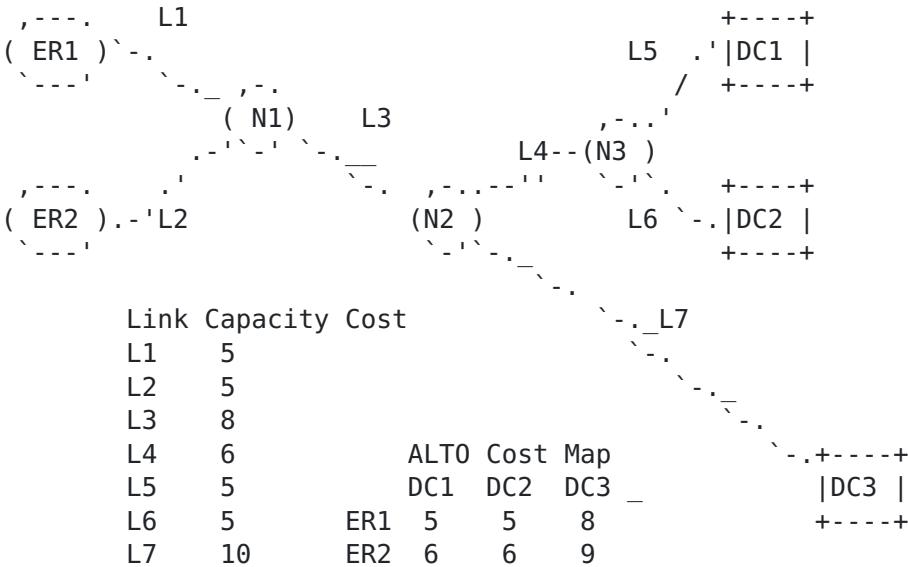


Figure 2. Example network illustrating bottlenecks

With the current ALTO cost map structure, the least cost path from ER1 would be either to DC1 or DC2. However, with the proposed capacitated cost map, the connection from ER1 to DC3 could be a better choice than the rest depending on the relative cost of network resources to data center resources.

A more general and relatively efficient alternative is to provide the requestor with a capacitated and multiply weighted graph that approximates and abstracts the capabilities of the network as seen by the source and destination location sets. This document provides ALTO information model and protocol extensions to support the graph model in [Section 4](#).

3. ALTO Constraints Filtering Extension Model

3.1. ALTO Query from Application Stratum to Network Stratum

In order for the network stratum to provide its resource information, the application stratum needs to provide the End Point

Cost Map to the network stratum. The End Point Cost Map should include the following information at a minimum:

- . End Point Source Address(es) /* these are the respective addresses of the nearest PE's to the user/client location */
- . End Point Destination Address(es) /* these are the respective addresses of the nearest PE's to a set of the candidate Data Center locations that can provide service to the user request. */

Note that how ALTO client derives the End Point Source/Destination addresses in terms of the nearest PE's is beyond the scope of this document.

- . Cost Type:= 'routingcost' as defined by base specification. Additional cost (ex. latency, hopcount) are defined in [\[MultiCost\]](#).
- . Cost Mode :={summary, graph} /* the cost map can be either a summary form or a graph form */
 - o Cost Mode: summary

This cost mode is indicated by string 'summary'. This mode indicates that the returned costs contain constraints values which can be used by application stratum for better selection of resources.
 - o Cost Mode: graph

This cost mode is indicated by string 'graph'. [TBD]
- . Constraints /* a set of constraints that apply to the requested path summary or graph for filtering. For instance, constraints can be like bandwidth greater than 'x', latency less than 'y', hopcount less than 'z', packetloss less than 'a' etc. */
- . Objective-function: The summary or the graph should be computed based on optimizing which parameter - IGP cost, latency, residual bandwidth, etc. This is for future use.

[3.2.](#) ALTO Response from Network Stratum to Application Stratum

In response to the ALTO Query from the Application Stratum, the Network Stratum needs to provide the filtered Cost Map Data of the

feasible path found. The Filtered End Cost Map Data should include the following information at a minimum:

- . The list of feasible Source-Destination pair and its Cost Type
- . For each feasible S-D pair, indicate the following:
- . Constraints Values /* indicate the actual values of each constraint requested */

3.3. Information Model of ALTO Query from Application Stratum to Network Stratum

Alto query:

```
object {
  CostMode          cost-mode;
  CostType          cost-type;
  JSONString        constr<0..*>;          [OPTIONAL]
  JSONString        ObjectiveFunction <0..*>; [OPTIONAL]
  EndpointFilter    endpoints;
} CsoReqEndpointCostMap;
```

3.4. Information Model of ALTO Response from Network Stratum to Application Stratum

Alto response:

```
object {
  JSONNumber hopcount;
  JSONNumber latency;
  JSONNumber pktloss;
  JSONNumber cost;      /* the cost/rank is determined based on
the filtering done using Objective function parameter and mode*/
} DstCostsConstraints;

object CsoEndpointDstCosts {
  DstCostsConstraints[TypedEndpointAddr];    ...
};
object {
  CsoEndpointDstCosts [TypedEndpointAddr]<0..*>;
  ...
}
```

```
} CsoEndpointCostMapData;  
  
object {  
    CostMode          cost-mode;  
    CostType          cost-type;  
    CsoEndpointCostMapData map;  
} CsoInfoResourceEndpointCostMap;
```

For each source endpoint, a CsoEndpointDstCosts object denotes the associated "cost + constraints" to each destination endpoint specified in the input parameters; the name for each member in the object is the TypedEndpointAddr string identifying the corresponding destination endpoint. This can be viewed as a two dimensional array which holds cost + constraints value.

3.5. ALTO Protocol Extension for Constraints Filtering Mechanism

This section provides the ALTO protocol extensions based on the information model discussed in Sections [3.3.](#) and 3.4. The scenario provided in this section is that the ALTO client in the Application Stratum requests the summary cost map from the network with one source and three destinations.

In this particular example, the ALTO client requests the filtered summary of the network path subject to: bandwidth ≥ 20 , latency < 10 , hop count < 5 and packet loss < 0.03 .

The ALTO server provides the resulted network paths in summary.

```
POST /endpointcost/lookup HTTP/1.1  
Host: alto.example.com  
Content-Length: [TODO]  
Content-Type: application/alto-csoendpointcostparams+json  
Accept: application/alto-csoendpointsummary+json,application/alto-error+json  
{  
    "cost-mode" : "summary",  
    "cost-type" : "routingcost",  
    "constraints": ["bw gt 20", "latency lt 10", "hopcount lt 5",  
    "pktloss lt 0.03"],  
    "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-csoendpointsummary+json

```
{
  "meta" : {},
  "data" : {
    "cost-mode" : "summary",
    "cost-type" : "routingcost",
    "map" : {
      "ipv4:192.0.2.2": {
        "ipv4:192.0.2.89" : [ "latency eq 5",
                              "hopcount eq 8", "pktloss eq 0.01" ],
        "ipv4:18.51.100.34" : [ "latency eq 9",
                                "hopcount eq 10", "pktloss eq 0.02" ],
        "ipv4:203.0.113.45" : [ "latency eq 40",
                                "hopcount eq 12", "pktloss eq 0.02" ]
      }
    }
  }
}
```

4. ALTO Protocol Extension for Graph Representation Mechanism

4.1. Representing bandwidth constraints

```
object {
  LinkEntry [LinkName]<0..*>;
} CostConstraintGraphData;

object {
  PIDName:    a-end; // Node name at one side of the link
  PIDName:    z-end; // Node name at the other side of the link
  Weight:     wt;
```

```
JSONNumber: latency;
Capacity:    r-cap; // Reservable capacity
} LinkEntry;
```

Where a link name is formatted like a PIDName (but names a link), and PID names are used for both provider defined location and provider defined internal model node identification. A graph representation of the network of Figure 2 might look like:

```
{
  "meta" : {},
  "data" : {
    "graph": {
      "L1": {"a-end":"ER1", "z-end":"N1", "wt":1,"r-cap":5},
      "L2": {"a-end":"ER2", "z-end":"N1", "wt":2,"r-cap":5},
      "L3": {"a-end":"N1", "z-end":"N2", "wt":1,"r-cap":8},
      "L4": {"a-end":"N2", "z-end":"N3", "wt":2,"r-cap":6},
      "L5": {"a-end":"N3", "z-end":"DC1", "wt":1,"r-cap":5},
      "L6": {"a-end":"N3", "z-end":"DC2", "wt":1,"r-cap":5},
      "L7": {"a-end":"N2", "z-end":"DC3", "wt":6,"r-cap":10}
    }
  }
}
```

[5. Summary and Conclusion](#)

TBD

[6. Security Considerations](#)

TBD

[7. IANA Considerations](#)

TBD

8. References

8.1. Informative References

- [HighBW] G. Bernstein and Y. Lee, "Use Cases for High Bandwidth Query and Control of Core Networks," [draft-bernstein-alto-large-bandwidth-cases](#), work in progress.
- [MultiCost] S. Randriamasy, Ed., "Multi-Cost ALTO," [draft-randriamasy-alto-multi-cost](#), work in progress.

Author's Addresses

Young Lee
Huawei Technologies
1700 Alma Drive, Suite 500
Plano, TX 75075
USA
Phone: (972) 509-5599
Email: ylee@huawei.com

Greg M. Bernstein
Grotto Networking
Fremont California, USA
Phone: (510) 573-2237
Email: gregb@grotto-networking.com

Sreekanth Madhavan
Huawei Technologies, India
Email: sreekanth.madhavan@huawei.com

Dhruv Dhody
Huawei Technologies, India
Email: dhruv.dhody@huawei.com

Tae-Sang Choi
ETRI
161 Gajong-Dong, Yusong-Gu
Daejeon, Republic of Korea
Phone: (8242) 860-5628
Email: choits@etri.re.kr

Intellectual Property Statement

The IETF Trust takes no position regarding the validity or scope of any Intellectual Property Rights or other rights that might be claimed to pertain to the implementation or use of the technology described in any IETF Document or the extent to which any license under such rights might or might not be available; nor does it represent that it has made any independent effort to identify any such rights.

Copies of Intellectual Property disclosures made to the IETF Secretariat and any assurances of licenses to be made available, or

the result of an attempt made to obtain a general license or permission for the use of such proprietary rights by implementers or users of this specification can be obtained from the IETF on-line IPR repository at <http://www.ietf.org/ipr>

The IETF invites any interested party to bring to its attention any copyrights, patents or patent applications, or other proprietary rights that may cover technology that may be required to implement any standard or specification contained in an IETF Document. Please address the information to the IETF at ietf-ipr@ietf.org.

Disclaimer of Validity

All IETF Documents and the information contained therein are provided on an "AS IS" basis and THE CONTRIBUTOR, THE ORGANIZATION HE/SHE REPRESENTS OR IS SPONSORED BY (IF ANY), THE INTERNET SOCIETY, THE IETF TRUST AND THE INTERNET ENGINEERING TASK FORCE DISCLAIM ALL WARRANTIES, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO ANY WARRANTY THAT THE USE OF THE INFORMATION THEREIN WILL NOT INFRINGE ANY RIGHTS OR ANY IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

Acknowledgment

Funding for the RFC Editor function is currently provided by the Internet Society.