

Network Working Group
Internet-Draft
Intended status: Experimental
Expires: September 15, 2011

S. Randriamasy, Ed.
Alcatel-Lucent Bell Labs
March 14, 2011

Multi-Cost ALTO
draft-randriamasy-alto-multi-cost-02

Abstract

IETF is designing a new service called ALTO (Application Layer traffic Optimization) that includes a "Network Map Service", an "Endpoint Cost Service" and an "Endpoint (EP) Ranking Service" and thus incentives for application clients to connect to ISP preferred Endpoints. These services provide a view of the Network Provider (NP) topology to overlay clients.

The present draft proposes a light way to extend the information provided by the current ALTO protocol. The purpose is to broaden the possibilities of the Application Clients in two ways: firstly by providing a better mapping of the Selected Endpoints to needs of the growing diversity of Content Networking Applications and to the network conditions, secondly by producing a more robust choice of multiple Endpoints, helping thus out for efficient Multi-Path transfer.

There are 2 parts in this draft: the first part proposes protocol extensions to support requests on multiple CostTypes in 1 transaction; the second part proposes additional CostTypes and Cost attributes related to timeframe and validity period.

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 September 15, 2011.

Copyright Notice

Copyright (c) 2011 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 | 4 |
| 2. Scope | 5 |
| 3. Terminology | 5 |
| 4. Proposed ALTO services updates | 6 |
| 4.1. Endpoint Cost Service with multiple Cost Types | 6 |
| 4.2. All Costs Types in one response with vector cost values | 6 |
| 4.3. Proposed additional Cost Types | 7 |
| 4.4. Statistical Costs with a timeframe | 7 |
| 5. Proposed ALTO protocol updates | 8 |
| 5.1. Proposed updates for Multi-Cost ALTO | 8 |
| 5.1.1. Multi-Cost related Attributes | 9 |
| 5.2. Proposed additional Properties and Costs | 9 |
| 5.2.1. Proposed additional Endpoints properties | 9 |
| 5.2.2. Scoping ALTO information | 10 |
| 5.2.3. Proposed additional Cost Types | 10 |
| 5.3. ALTO Status Codes for Multi-Cost ALTO | 11 |
| 5.4. Examples of Multi-Cost ALTO messages | 11 |
| 6. Use case | 11 |
| 6.1. Scenario | 11 |
| 6.2. Illustrative ALTO use case | 12 |
| 7. IANA Considerations | 15 |
| 7.1. Information for IANA on proposed Cost Types | 15 |
| 7.2. Information for IANA on proposed Endpoint Properties | 15 |
| 8. Acknowledgements | 16 |
| 9. References | 16 |
| 9.1. Normative References | 16 |
| 9.2. Informative References | 16 |
| Author's Address | 16 |

1. Introduction

IETF is designing a new service called ALTO that provides guidance to P2P applications, which have to select one or several hosts from a set of candidates that are able to provide a desired resource. This guidance shall be based on parameters that affect performance and efficiency of the data transmission between the hosts, e.g., the topological distance. The ultimate goal is to improve Quality of Experience (QoE) in the application while reducing resource consumption in the underlying network infrastructure. The ALTO protocol conveys the Internet View from the perspective of a Provider Network region that spans from a region to one or more Autonomous System (AS). Together with this Network Map, it provides the Provider determined Cost Map between locations of the Network Map. Last, it provides the Ranking of Endpoints w.r.t. their routing cost.

The term Network Provider in this document includes both ISPs, who provide means to transport the data and Content Delivery Network (CDN) operators who care for the dissemination, persistent storage and possibly identification of the best/closest content copy.

The last ALTO protocol draft see [ID-alto-protocol6], gives the possibility to query multiple Endpoint properties at once (see S.7.7.4.1). However section 7.7.3.2 on Cost Map states about both parameters Cost Type and Cost Mode that: "This parameter MUST NOT be specified multiple times". The ALTO requirements draft, see [ID-ALTO-Requirements7] also states in REQ. ARv05-14: "The ALTO client protocol MUST support the usage of several different rating criteria types". In the current protocol draft, there is no specified way to get values for several Cost Types altogether. Currently, the costs are provided in a scalar form, one by one. So that an ALTO Client wanting information for several Cost Types must place a request and receive a response as many times as desired Cost Types. However, vector costs provide a robust and natural input to multi-path connections and getting all costs in one single query/response transaction saves time and ALTO traffic, thus resources, thus energy.

The ALTO Problem Statement, see [RFC5693] and the ALTO requirements draft, see [RFC5693] stress that: "information that can change very rapidly, such as transport-layer congestion, is out of scope for an ALTO service. Such information is better suited to be transferred through an in-band technique at the transport layer instead", as "ALTO is not an admission control system "and does not necessarily know about the instant load of endpoints and links. However, longer term statistics or empirical ratings on performance oriented information may still be useful for a reliable choice of candidate endpoints. In addition, given the QoE requirements of nowadays and

future Internet applications, more and more NPs compute and store such information to optimize their traffic. Last, specific ALTO servers can be specified for mobile core networks, which have a smaller scale and can afford and take advantage of using smaller time-scale network information.

Adding QoE-enabling metrics to the Network Provider established routing cost could meet the interests of both the end users and the Providers. Besides, keeping the shortest or cheapest possible path, in addition, saves resources, time and energy.

2. Scope

This draft generalizes the case of a P2P client to include the case of a CDN client, a GRID application client and any Client having the choice in several connection points for data or resource exchange. To do so, it uses the term "Application Client" (AC).

This draft focuses on the use case where the ALTO client is embedded in the Application Client. For P2P applications, the use case where the ALTO Client is embedded in the P2P tracker is also applicable.

It is assumed that Applications likely to use the ALTO service have a choice in connection endpoints as it is the case for most of them. The ALTO service is managed by the Network Provider and reflects its preferences for the choice of endpoints. The NP defines in particular the network map, the routing cost among Network Locations, and which ALTO services are available at a given ALTO server.

The solution proposed in this draft is applicable to fixed networks. It is also meant for smaller networks such as mobile networks.

3. Terminology

Endpoint (EP): can be a Peers, a CDN storage location, a Party in a resource sharing swarm such as Grid or online gaming.

Endpoint Discovery (EP Discovery) : this term embraces the different types of processes used to discover different types of endpoints.

Network provider: includes both ISPs, who provide means to transport the data and Content Delivery Network (CDN) who care for the dissemination, persistent storage and possibly identification of the best/closest content copy.

Application Client (AC): this term generalizes the case of a P2P

client to include the case of a CDN client and of any Client having the choice in several connection points for data or resource exchange.

Traffic Engineered End Point Optimization Tool (TEEPOT): this is a functional entity introduced in this draft, that is linked to an ALTO Client and to an Application Client. Its role is to assist the selection of Endpoints upon Allocation needs and the ALTO responses. It can be a specific group of functions or an already existing function.

4. Proposed ALTO services updates

The currently available ALTO services supporting Endpoint evaluation are: Endpoint Cost Service, Cost Map and Filtered Cost Map. The ALTO client may want to simultaneously use a number $N > 1$ of cost metrics referred to as Cost Types in ALTO. The only possibility in the current ALTO protocol is to sequentially place as many requests as desired cost types. This draft proposes to add the following features:

4.1. Endpoint Cost Service with multiple Cost Types

Some application clients may want to consider several metrics to select the endpoints appropriately w.r.t. the application needs. Clients may also want to use multiple paths for the transfer of particular data bulks, possibly selected with several metrics. Therefore the Endpoint Cost Lookup and the Cost Map Services should have the possibility to handle several metrics.

4.2. All Costs Types in one response with vector cost values

Providing all the numerical costs simultaneously with only one request and response exchange saves time, resources and energy. To avoid overloading the network with ALTO traffic with multiple requests for Cost Types, we propose that the Cost values provided by the ALTO server be arranged in a vector. This requires:

- o to put the requested cost values in an array or vector having a number $N \geq 1$ of components.
- o to define a canonical order that allows to match values in these vectors with Cost Types and Properties.

As specified in the ALTO Requirements [ID-ALTO-Requirements7] "REQ. ARv05-19: The ALTO reply message SHOULD allow the ALTO server to express which rating criteria have been considered when generating

the reply." That is, the ALTO response indicates the mapping between vector components and Cost Types.

Note that in this case, the ALTO client MUST require the Cost Mode "numerical" that is the Mode MUST NOT be "ordinal".

4.3. Proposed additional Cost Types

The current ALTO protocol draft provides examples of metrics in section 5.1.1, that are: air miles, hop-counts or generic routing costs. Statistics or longer term ratings on path bandwidth and latency may also be considered. Additional Endpoint properties may be useful, such as the memory capacity or statistical scores on the load and possibilities of an Endpoint.

4.4. Statistical Costs with a timeframe

The ALTO Requirements Draft [ID-ALTO-Requirements7] advises against instant performance-related cost metrics as they may be easily captured by online mechanisms and in addition, the ALTO service does not know how a Peer manages its sending rate. Application clients however may have good reasons and wise ways to use performance related information in the mid to long term, on Endpoints that they don't know in advance and on which they therefore cannot plan measurements. Other applications may wisely use static performance indicators such as nominal memory capacity.

Dynamic performance indicators can be represented by scores, reflecting some overall performance, in a static way or with values periodically updated at intervals typically longer than a network layer packet RTT, as assumed in [ID-ALTO-Requirements7].

If statistical Cost Types are available, the following types of information should report on them:

- o their "statistical" nature: for example a mean value, or a median value,
- o their timeframe: that is the period over which statistics were computed and the age of the information. By default this timeframe is supposed permanent, that is, the corresponding EP Cost or Property values are permanent. Timeframe information can be easily recovered by attributes listed in [ID-ALTO-Requirements7] such as 'lifetime' (see REQ. ARv07-29) and an aging mechanism (see REQ. ARv07-29), such as a RFC3339 based TimeStamp.

- o the validity period: indicating the date at which the information can be considered obsolete and updated. This can be easily reflected by the 'age' reflecting the date at which the information was generated and 'lifetime' of this information.

'Lifetime' and 'age' should be also available to other applicable 'non statistical' Cost Types, such as 'OccupationLevel' that can be used to describe an empirical and restricted set of load value ranges.

5. Proposed ALTO protocol updates

This section proposes updates or additions to the ALTO protocol to support Multi Cost ALTO Services or provide additional ALTO information. The applicable ALTO services are:

- o Cost Map Service,
- o Cost Map Filtering Service,
- o Endpoint Property Lookup Service,
- o Endpoint Cost Lookup Service.

5.1. Proposed updates for Multi-Cost ALTO

If an ALTO client desires several Cost Types, instead of placing as many requests as costs, it may request and receive all the desired cost types in one transaction. The correspondence between the components and the cost type MUST be indicated in the ALTO request.

The ALTO server then, provided it supports the desired cost, and provided it supports the vector cost values, sends one single response where for each {source, destination} pair, the cost values are arranged in a vector, whose component each corresponds to a specified Cost Type. The correspondence between the components and the cost types MUST be indicated in the ALTO response.

The following ALTO protocol services and features need to be updated to enable Multi Cost ALTO transactions.

- o Endpoint (EP) Property (see [ID-alto-protocol6])
- o Endpoint (EP) Cost (see [ID-alto-protocol6]).
- o Cost attributes (see [ID-alto-protocol6]).

- o Cost Map (see [ID-alto-protocol6]):
 - * between Network Locations (that are groups of 1 or several endpoints).
- o Cost Map filtering: need the same updates as for the Cost Map.

5.1.1. Multi-Cost related Attributes

To enable Multi-Cost ALTO Cost Services, we propose the following updates to the Cost Attributes, described in [ID-alto-protocol6] .

- o extension of the attribute Cost Type from a single value to a vector of $N \geq 1$ values. If $N > 1$, then the values WILL be interpreted as numerical values.
- o addition of definitions that list and identify the Cost Types supported by the acting ALTO server. These definitions will be formulated according to the syntax defined in Section 7.7 of [ID-alto-protocol6],
- o definition of the correspondence between an index "i_typecost" in [1,N] in a cost vector and the ID of the defined cost types and properties.
- o optional association of a validity timeframe, indicating how long the information can be considered as up to date.
 - * by default the validity timeframe WILL be considered infinite

To the attribute Cost Mode in S.5.1: addition of a rule stipulating that when multiple cost types are requested, then the requested Cost Mode MUST be numerical. If the attribute Cost Length is > 1 and the Cost Mode is set to "ordinal", then one option is that the ALTO Server returns the 'Sucess' code "E_INVALID_COST_TYPE".

5.2. Proposed additional Properties and Costs

5.2.1. Proposed additional Endpoints properties

The Endpoint Properties given as example in [ID-alto-protocol6] S.3.2.3 mostly apply to fixed end nodes. We propose to add other properties, that are static, contribute to reflect the potential physical abilities of end nodes and therefore may guide their selection. In addition, these properties apply to end nodes connected by any access technology. Example additional properties include:

- o EP capacity in memory,
- o EP nominal bandwidth,
- o EP access technology.

Note that if this service is not supported, it is possible although less convenient to get the information at the overlay level, thus without the ALTO server.

5.2.2. Scoping ALTO information

One way to moderate the ALTO traffic load while maintaining some reliability is to associate the following attributes to the applicable ALTO information:

- o an age attribute indicating when the information was generated.
- o for statistical costs a time period attribute indicating over which period the statistics were collected.
- o a lifetime attribute as proposed in [ID-ALTO-Requirements7] . By default, this parameter can be set to infinity.

The Time related values can be used by the aging mechanism as proposed in REQ ARv05-28 of [ID-ALTO-Requirements7] for a better synchronization of Cost Information collected at various times and places.

5.2.3. Proposed additional Cost Types

Additional Cost Types may be used in either the Cost Map or the Endpoint Cost Lookup Services and include:

- o Endpoint availability: indicating how often an Endpoint is reachable, preferably as a percentage. To be further specified. Possibly with associated Time frame and Time To Expire.
- o Endpoint reliability: indicating how easily an Endpoint is reachable, and / or the degree of continuity of its reachability, preferably as a percentage. To be further specified.
- o Endpoint Load: indicating the average load, preferably as a percentage, or a quantitative coarse grain index indicating whether this Endpoint is in a rush period or calm period. To be further specified.

- o Path robustness: one or more timeframed indicators related to statistical evaluations of the path performance on bandwidth, delay, packet loss, or other such metrics. This Cost can also be represented by a quantitative coarse grain index indicating whether this Endpoint is in a rush period or calm period. To be further specified.

5.3. ALTO Status Codes for Multi-Cost ALTO

If the vector cost structure is not supported, then the ALTO server sends an ALTO status code 7 corresponding to HTTP status code 501 indicating "Invalid cost structure". The ALTO client may then needs to place as many requests as needed Cost Types, and the ALTO server sends as many cost maps or EP cost as needed.

To the attribute Cost Mode in S.5.1 should be associated a rule stipulating that when multiple cost types are requested, then the requested Cost Mode MUST be numerical. If the attribute Cost Length is > 1 and the Cost Mode is set to "ordinal", an option is that the ALTO Server returns the 'Sucess' code "E_INVALID_COST_TYPE".

5.4. Examples of Multi-Cost ALTO messages

Request and Response syntax. To be further specified.

6. Use case

6.1. Scenario

A Multi-Cost ALTO transaction is illustrated in a simple scenario, where an application client in a terminal wants to use several paths for a data transfer. This scenario applies to a terminal having access to the network via one or several interfaces.

The application client wants for example 3 paths per transfer:

- o 1 path optimising the Cost Type 'routingcost',
 - o 2 other paths optimizing 2 metrics: the Cost Type 'routingcost' and an Endpoint property named 'EP memory'.
- * The application client in addition wants these 2 paths to optimize the first criterion with a weight W_PATH_LENGTH equal for example to 0.4 and the second criterion with a weight W_EP_MEMORY equal to 0.6.

- * If the EP Property Service provides the information on Endpoint Load, then the application client wants this information in the available lifetime closest to 1 hour.

A TEEPOT connected with the ALTO Client and the Application Client takes in the list of candidate Endpoints from the Application Client and prepares for the ALTO Client the request to the ALTO Server, in particular the following information: vector TimeFrame[EP Cost Length], with components equal to either a value or an indication of "not applicable".

- o The list of requested EP Cost Types, that are identified by their index I_CostType.
- o

6.2. Illustrative ALTO use case

Figure 1 shows the example scenario in the last IETF ALTO protocol draft, where the ALTO client is embedded in the P2P Client and requires an ALTO server servicing its own ISP to provide the Endpoint Cost for a list of gathered peers.

As written in [ID-alto-protocol6], the use case proceeds as follows:

1. The P2P Client discovers peers from sources such as Peer Exchange (PEX) from other P2P Clients, Distributed Hash Tables (DHT), and P2P Trackers.
2. The P2P Client queries the ALTO Server's Ranking Service, including discovered peers as the set of Destination Endpoints, and indicates the 'ordinal' Cost Mode. The response indicates the ranking of the candidate peers.
3. The P2P Client connects to the peers in the order specified in the ranking.

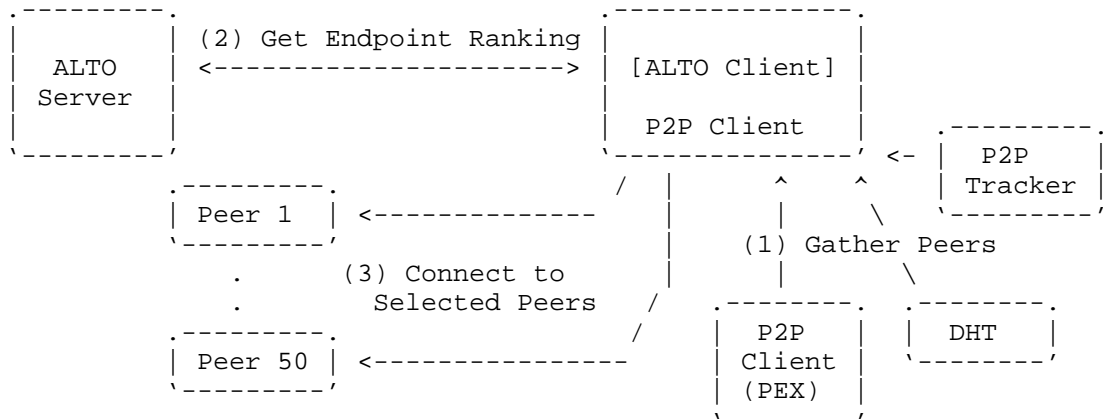


Figure 1: example scenario in the last IETF ALTO protocol draft, where the ALTO client is embedded in the P2P Client

Figure 2 depicts the features and mechanisms added to the current ALTO scenario for Multi-Cost ALTO services, for the use case of Figure 1. The EPs have already been discovered. In this figure, the term Peer is replaced by the term Endpoint (EP), the term P2P Client by Application Client and an Endpoint Tracker for resource Sharing Applications is added to the tools involved in Step (1) Gather Endpoints .

We focus on the ALTO use case where the ALTO client is co-located with an Application client in a terminal node, as not all P2P systems use a P2P tracker for peer discovery and selection as written in section 9.2 of [ID-alto-protocol6]. In Figure 2, the entity called P2P Client mentioned in the current protocol draft is zoomed to an entity called in this draft "Client Block" and that links: the Application Client (AC), its ALTO Client and the Traffic Engineered EP Optimization Tool (TEEPOT).

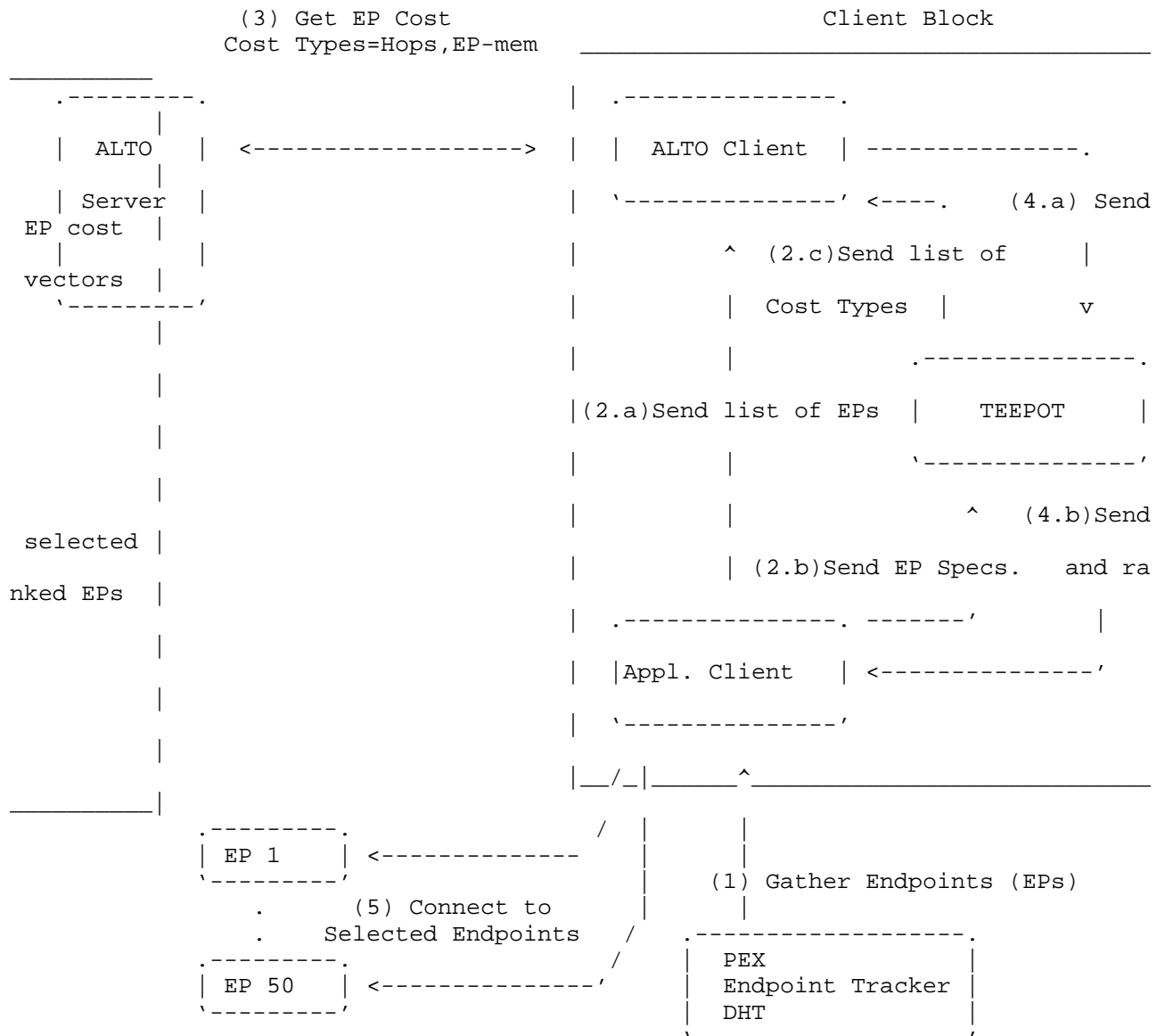


Figure 2: features and mechanisms added to the current ALTO scenario for Multi-Cost ALTO services

The use case in Figure 2 proceeds as follows:

1. The Application Client discovers Endpoints (EPs) from sources such as Peer Exchange (PEX) from other P2P Clients, Distributed Hash Tables (DHT), P2P Trackers or other types of EP trackers.
2. In the "Client Block" gathering the Application Client (AC), its ALTO Client and the Traffic Engineered EP Optimization Tool (TEEPOT):
 - A. the Application Client (AC) sends to the ALTO Client the list of the discovered peers as the set of Destination Endpoints.
 - B. the Application Client (AC) sends to the TEEPOT the specifications on the EPs to select, according to the needs of the application. For example, AC needs 3 EPs, with 1 EP optimizing the Path Length Metric and 2 EPs optimizing the Path Length and the EP Memory Capacity Score, with respective

weights of 0.4 and 0.6.

- C. the TEEPOT indicates to the ALTO Client that the Service to request is EP Cost, with the Cost Mode set to "Numerical", and the Cost Dimension equal to the number of requested metrics and with the index of the requested Cost Types.
- 3. The ALTO Client queries the ALTO Server's EP Cost Service, sends the list of the discovered peers as the set of Destination Endpoints and the index of requested metrics, corresponding in this example to: "Path Length" and "EP Memory Capacity Score". As the number of requested metrics is > 1, the Cost Mode is implicitly set to 'numerical'. The ALTO Server response contains the set of metric values associated to each EP.
- 4. In the Client block:
 - A. The ALTO Client hands to the TEEPOT the list of EPs and their associated value set.
 - B. The TEEPOT ranks the EPs with some smart algorithm, given the metric weights and then sends the ranked list to the Application Client.
- 5. The Application Client connects to the selected EPs.

7. IANA Considerations

The current ALTO protocol version includes a Section 11 entitled IANA considerations, where the ALTO Cost Type registry is defined in Section 11.2

7.1. 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.

7.2. Information for IANA on proposed Endpoint Properties

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.

8. Acknowledgements

Thank you to Richard Alimi whose reviewing of the previous version of this draft and advises provided a valuable input to its updates.

9. References

9.1. Normative References

- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.
- [RFC5693] "Application Layer Traffic Optimization (ALTO) Problem Statement", October 2009.

9.2. Informative References

- [ID-ALTO-Requirements]
"draft-ietf-alto-reqs-05.txt", June 2010.
- [ID-ALTO-Requirements7]
"draft-ietf-alto-reqs-07.txt", January 2011.
- [ID-alto-protocol5]
"ALTO Protocol" draft-ietf-alto-protocol-05.txt",
July 2010.
- [ID-alto-protocol6]
, eds., "ALTO Protocol" draft-ietf-alto-protocol-06.txt",
October 2010.

Author's Address

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

Email: Sabine.Randriamasy@alcatel-lucent.com

