

Network Working Group  
Internet-Draft  
Intended status: Standards Track  
Expires: January 7, 2016

Z. Dulinski  
Jagiellonian University  
P. Wydrych  
R. Stankiewicz  
AGH University  
July 6, 2015

Inter-ALTO Communication Problem Statement  
draft-dulinski-alto-inter-problem-statement-02

Abstract

The Application-Layer Traffic Optimization (ALTO) protocol has been standardized in RFC7285 to ease better-than-random overlay connection management. Due to various reasons, optimization capabilities of ALTO servers are limited by the fact that it may not be possible for an ALTO server to compute costs for source/destination pairs correctly if a source and/or a destination is outside the administrative domain it belongs to. In other words, an ALTO server may be generally unable to optimize the traffic with only locally available topology information.

Therefore, it is proposed that operators of distinct administrative domains may agree on topology- and policy-related information exchange through a standardized inter-ALTO communication framework. This document provides a rationale for design, standardization, and implementation of the inter-ALTO communication framework.

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 January 7, 2016.

## Copyright Notice

Copyright (c) 2015 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 . . . . .	2
2. Definitions . . . . .	3
3. Motivation for Inter-ALTO Communication . . . . .	4
3.1. Remote View of Topology . . . . .	6
3.2. Detailed Information on Remote Topology . . . . .	7
3.3. Partitioned Networks . . . . .	7
4. IANA Considerations . . . . .	8
5. Security Considerations . . . . .	8
6. References . . . . .	8
6.1. Normative References . . . . .	8
6.2. Informative References . . . . .	9
Appendix A. Acknowledgments . . . . .	9
Authors' Addresses . . . . .	9

## 1. Introduction

As stated in [RFC5693], information on network topologies and routing policies in today Internet is not generally available to the application layer. At the same time, a lot of new overlay applications creating their own topologies on top of the physical one emerge. As both network operators and users of the overlay applications may benefit from better-than-random overlay connection management, the ALTO protocol has been standardized in [RFC7285].

Topology- and policy-related information may be supplied through ALTO in a proactive or in a reactive way. In the former case, an ALTO server distributes network and cost maps through which a client may compute costs of sending the traffic, while in the latter case, a client may request a server to provide rating of explicitly specified source and destination pairs using the endpoint cost service. As a server provides information composed only from information that can

be gathered by it and applies policies defined by operator of the administrative domain it belongs to, the returned information has always local meaning to the server - regardless of the method used by a client to request guidance from it. Moreover, an ALTO client may not be able to receive ALTO information from outside of its administrative domain.

Due to various reasons, it may be generally not possible to optimize the traffic with only locally available topology information. Therefore, in this document, it is proposed that operators of distinct administrative domains may agree on topology- and policy-related information exchange through a standardized inter-ALTO communication framework.

Among others, the inter-ALTO communication framework may allow an operator of one administrative domain to apply its policies to topology information gathered from other administrative domains. Moreover, a server may have separate security contexts for processes responsible for server-to-client (i.e., ALTO) and server-to-server (i.e., inter-ALTO) communication to ensure that no sensitive topology- and policy-related information is distributed in an uncontrolled way.

The requirements for the inter-ALTO communication framework, its detailed specification and issues related to server discovery [RFC7286] are out of scope of this document.

## 2. Definitions

This document uses the following terms defined in [RFC5693] and [RFC7285]: ALTO Server, ALTO Client, Endpoint, Peering Traffic, Transit Traffic. Moreover, the following terms have the special meaning in the definition of the inter-ALTO communication problem.

**Local Administrative Domain:** The administrative domain which the ALTO client belongs to.

**Remote Administrative Domain:** An administrative domain other than the one which the ALTO client belongs to.

**Local ALTO Server:** An ALTO server belonging to the local administrative domain.

**Remote ALTO Server:** An ALTO server belonging to a remote administrative domain.

**Local Endpoint:** An endpoint belonging to the local administrative domain.

Remote Endpoint: An endpoint belonging to a remote administrative domain.

### 3. Motivation for Inter-ALTO Communication

Optimization capabilities of ALTO servers are limited by the fact that they use information available locally only. It can be shown that without additional information on remote endpoints, routing paths, or remote administrative domains' preferences, rating provided by an ALTO server may be sub-optimal for both sides. Data from remote administrative domains obtained from an ALTO server holding authoritative information about those domains' topologies may have a substantial significance for the traffic management.

The ALTO Protocol specification [RFC7285] states that:

It may also be possible for an ALTO server to exchange network information with other ALTO servers (either within the same administrative domain or another administrative domain with the consent of both parties) in order to adjust exported ALTO information. Such a protocol is also outside the scope of this specification.

This document provides a rationale for design, standardization, and implementation of the inter-ALTO communication framework, described as "External Interface" in Figure 1 of [RFC7285]. The requirements and detailed specification are out of scope of this document.

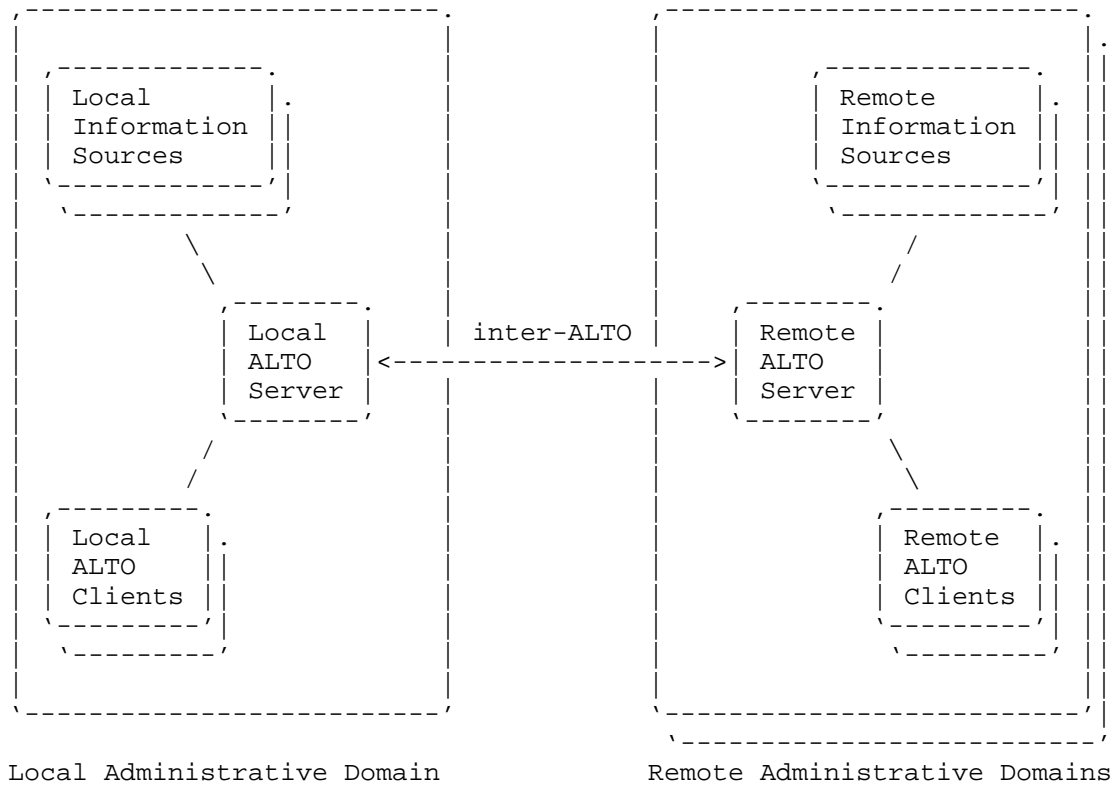


Figure 1: Inter-ALTO communication architecture.

The architecture of the inter-ALTO communication framework is shown in Figure 1. Both ALTO servers gather the information from their information sources like routing protocols, provisioning policies, or dynamic network information sources using respective provisioning protocols. To provide (using the ALTO Protocol) better guidance for its clients, the local ALTO server needs to communicate with remote ALTO servers to obtain information which is available only at the entities in the remote administrative domains.

In particular, two general use-cases of inter-ALTO communication may be distinguished:

- o providing the local ALTO server with information on remote view of the multi-domain topology;

- o providing the local ALTO server with information on remote network topology, more accurate than available from the local information sources.

Moreover, for partitioned networks run by one organization or a consortium, both use-cases are valid.

### 3.1. Remote View of Topology

Due to the complex topology of the current Internet and independent application of routing policies in autonomous systems, the communication between two endpoints does not need to follow the same path in the opposite directions. Moreover, there is a significant disproportion between availability of information on upstream and downstream paths. While the current local information sources can easily export data on the former, the latter are generally unknown to ISPs. Depending on the deployment-specific use-case and cost metric applied, this fact may indeed affect the ALTO server's ability to calculate the costs.

For a set of reasons (e.g., application performance or quality perceived by end-users) an ALTO server may suggest its clients to connect to endpoints located in their proximity. One of the simplest measures of proximity is the number of AS hops, as announced by BGP. As indicated above, due to the route asymmetry, the number of AS hops between two communicating endpoints may significantly differ between the upstream and downstream paths.

Under other circumstances, an ISP may prefer to reduce the transit traffic by increasing the peering one and it may configure its ALTO server to differentiate endpoints into classes taking into account through which links the traffic is exchanged and what are its business relationships with its neighbors. Still, as the Internet routes are asymmetrical, a reply for request sent to an endpoint through a peering link may return via a transit one (or vice versa).

Therefore, an ALTO server may easily calculate costs of sending the traffic from the local administrative domain to remote ones, while calculation of correct costs of receiving the content from the remote administrative domains may be not possible at all. To mitigate this situation, the inter-ALTO communication framework may be used to exchange information on downstream paths between two interested parties. If a local ALTO server would be able to gather such information, a risk of suboptimal endpoint rating may be greatly reduced.

### 3.2. Detailed Information on Remote Topology

The second use-case stems from the fact that a lot of topology information is lost when a prefix is being announced via BGP by one AS to others. Moreover, prefix aggregation often used to reduce the size of routing tables causes that a number of networks of different characteristics are announced as one network. Therefore, a local ALTO server may consider - regardless of the cost metric used - two remote endpoints as equal, while they should be in fact differentiated.

For instance, a remote administrative domain may comprise (among others) of a set of networks connected to its core by expensive links or containing endpoints of worse capabilities than those in the rest of its network. Then, inter-ALTO communication may be used to denote such a fact and to characterize endpoints properly. Similarly, when some remote endpoints stand out above the rest, they may be promoted by the local ALTO server.

A cost metric may also take into account congestions on intra- and inter-domain links or another exhaustive consumption of some resources. When a link becomes congested for a longer period of time, it may be desirable to promote endpoints reachable through lightly loaded links. Likewise, a set of endpoints providing a content or a service may be overloaded and clients should be discouraged from using them. Regardless of the reason for endpoint differentiation, a local ALTO server may be informed by a remote one about remote domain's preference in endpoint selection.

### 3.3. Partitioned Networks

Means for exchange of detailed information on view of network topology may be also important for partitioned networks, run by one organization or a consortium of organizations fully trusting each other. To optimize the traffic flowing within partitions and between them, ALTO servers located in each partition may exchange detailed network topology information. In principle, ALTO servers may be deployed hierarchically or in a mesh. When a hierarchical architecture is used, a central ALTO server may gather a view of topology information from the rest of ALTO servers, merge the information, calculate the costs for all source/destination pairs, and distribute the merged network and cost maps to the ALTO servers and/or serve ALTO clients from all partitions. ALTO servers may be as well set up in each partition independently, connected to each other, gathering the network topology information from other partitions, and serving their own clients.

In both deployment schemes, both remote view of the inter-partition topology (see Section 3.1) and detailed view of remote partition topology (see Section 3.2) may bring a lot of benefit to the organization/consortium. The former can be used to optimize the traffic flowing between the partitions, while the latter may allow an ALTO server to differentiate endpoints within one partition.

#### 4. IANA Considerations

This document does not define any new media types nor does it introduce any new IANA considerations.

#### 5. Security Considerations

In general, communication between ALTO servers run by distinct parties and exchange of information on their topologies may require a formal agreement between them, mutual authentication, and authorization. Since sensitive data may be exchanged, a secure deployment of inter-ALTO communication framework may require setting up encrypted tunnels or using SSL between the ALTO servers.

The inter-ALTO communication may allow ALTO servers to exchange any parameters which allow them to manage traffic in an optimal way for mutual benefit. In order to achieve these results a set of administrative domains may exchange sensitive data which should be kept confidential. They should be used to calculate the cost maps, but should not be revealed directly to a third party, e.g., an ALTO client. To implement such a differentiation, an ALTO server may need to use separate security contexts for client-to-server and server-to-server communication.

Moreover, to ease secure environment deployment, administrative domains may form ALTO server communities, i.e., groups of ALTO servers trusting each other and working for common benefit.

#### 6. References

##### 6.1. Normative References

- [RFC5693] Seedorf, J. and E. Burger, "Application-Layer Traffic Optimization (ALTO) Problem Statement", RFC 5693, October 2009.
- [RFC7285] Alimi, R., Penno, R., Yang, Y., Kiesel, S., Previdi, S., Roome, W., Shalunov, S., and R. Woundy, "Application-Layer Traffic Optimization (ALTO) Protocol", RFC 7285, September 2014.



## 6.2. Informative References

[RFC7286] Kiesel, S., Stiernerling, M., Schwan, N., Scharf, M., and H. Song, "Application-Layer Traffic Optimization (ALTO) Server Discovery", RFC 7286, November 2014.

## Appendix A. Acknowledgments

P. Wydrych was supported by the grant no. 15.11.230.190.

Z. Dulinski and R. Stankiewicz were supported by the SmartenIT project, funded by the EU FP7 Program under contract no. FP7-2012-ICT-317846.

The authors would like to thank Y. Richard Yang (Tongji/Yale University), Wendy Roome (Alcatel-Lucent/Bell Labs), and Chen Guohai (Huawei) for their valuable comments.

## Authors' Addresses

Zbigniew Dulinski  
Jagiellonian University  
ul. Lojasiewicza 11  
Krakow 30-348  
Poland

Email: [dulinski@th.if.uj.edu.pl](mailto:dulinski@th.if.uj.edu.pl)

Piotr Wydrych  
AGH University of Science and Technology  
al. Mickiewicza 30  
Krakow 30-059  
Poland

Phone: +48 12 617 48 17  
Email: [piotr.wydrych@agh.edu.pl](mailto:piotr.wydrych@agh.edu.pl)  
URI: <http://wydrych.net/>

Rafal Stankiewicz  
AGH University of Science and Technology  
al. Mickiewicza 30  
Krakow 30-059  
Poland

Email: [rstankie@agh.edu.pl](mailto:rstankie@agh.edu.pl)