

ALTO
Internet-Draft
Intended status: Informational
Expires: September 14, 2011

X. Sun
China Telecom
Y. Yang
Yale University
March 14, 2011

ALTO Deployment Considerations: Configuration and Monitoring by ISPs
draft-sun-alto-ispdeployment-01.txt

Abstract

As ALTO specification continues in the ALTO Working Group and some applications start to conduct integration with ALTO, more ISPs start to evaluate key issues in the deployment of ALTO in their networks. In this document, we discuss key issues that an ISP needs to consider when deploying ALTO. In particular, we discuss issues on how to configure ALTO information as well as how to monitor the effectiveness of ALTO.

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 [1].

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 September 14, 2011.

Copyright Notice

Copyright (c) 2010 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 BSD License.

Table of Contents

- 1. Introduction 2
- 2. ALTO Service Configuration..... 3
 - 2.1. Optimization Area..... 3
 - 2.2. Network and Cost Map Configuration..... 4
 - 2.2.1. Network Map and PID..... 4
 - 2.2.2. Cost Map 5
 - 2.3. Endpoint Service Configuration..... 5
- 3. ALTO Deployment Monitoring..... 6
 - 3.1. Monitoring Metrics..... 6
 - 3.1.1. Network Metrics..... 6
 - 3.1.2. Application Metrics..... 7
 - 3.2. Monitoring Data Sources..... 7
 - 3.2.1. Application Log Server..... 7
 - 3.2.2. P2P Clients..... 7
 - 3.2.3. OAM 8
 - 3.3. Application/ISP Monitoring Integration..... 8
 - 3.3.1. Structure 8
 - 3.3.2. ALTO Monitoring Report Protocol..... 8
- 4. IANA Considerations 8
- 5. Security Considerations..... 9
- 6. References 9
 - 6.1. Normative References..... 9
 - 6.2. Informative References..... 9

1. Introduction

A basic service of ALTO is to provide information from network service providers to applications, in order to improve network efficiency and application performance. Some applications start to or have shown interests to conduct integration with ALTO. Some major ISPs (e.g., China Telecom) are in the process of deploying

ALTO services in their production networks. As a result, more ISPs start to evaluate key issues in the deployment of ALTO in their networks. Thus, a document highlighting some key issues that an ISP should consider in the deployment process can be a highly valuable reference.

The objective of this document is to provide such a reference. The document will try to draw on many valuable discussions in the ALTO mailing list as well as the predecessor p2pi mailing list. In addition, it will try to draw on the trial experiences of multiple ISPs (e.g., [CTTrial,ComcastTrial]).

The deployment of ALTO involves both ISPs and network applications. We can identify four major issues in ALTO deployment:

1. How does an ISP deploy and configure its ALTO servers?
Specifically, an ALTO Server provides the Network Map and the Cost Map. How does an ISP configure these maps? Where does an ISP deploy ALTO servers?
2. Which application entities fetch ALTO information?
3. How does an application integrate ALTO information into its decision process?
4. How does an ISP (potentially with collaboration from applications) monitor the deployment of ALTO, so that the ISP can better understand the status as well as the policy impacts of its ALTO deployment?

This document focuses more on the ISP perspective. Therefore, it focuses more on the first and the fourth issues. There are additional deployment documents in the ALTO working group that focus more on the second issue and the third issue. Our document is complementary to these other documents.

2. ALTO Service Configuration

2.1. Optimization Area

An ISP deploys ALTO service to optimize traffic for a given network area. We define a network area for which traffic need be optimized using the ALTO service as an optimization area. Optimization area definition is very important in ALTO service deployment, because there are multiply types of network and topology infrastructure in real Internet.

Traditional, an optimization area can be an access network, a MAN, or a larger network consisting of both access works and MANs. An

ISP with a relatively small network can define a single optimization area and deploy an ALTO server for the area.

An ISP with a larger network may partition its network into multiple optimization areas. Each optimization area may include one or more MANs. Alternatively, the ISP may choose to use a large optimization area and distribute a group of ALTO servers.

Besides the geographic point of view we can define the different Optimization Areas through the type of network too. The networks based on FTTx, ADSL, wifi and 3G or even 2G wireless technology has different traffic model and infrastructure. ALTO service should represent the difference between these networks.

2.2. Network and Cost Map Configuration

Key components for an ISP to configure when it deploys its ALTO service are the Network Map and Cost Map. They have impacts on both the load and the effectiveness of the service.

2.2.1. Network Map and PID

Different ISPs use different technologies to build their infrastructures. Some ISPs have only a relatively small network, focusing mainly on access. On the other hand, some large ISPs have access networks, MANs, and a Core network. There are tradeoffs when a large ISP defines its Network Map. If the partition of the network in the Network Map is too fine-grained, it may lead to higher complexity and overhead. On the other hand, a too coarse-grained Network Map may lead to suboptimal optimization.

Specifically, first consider an access network, say an ADSL or Ethernet based access network. A BAS server may be deployed to provide access service for its subscribers. Because all subscribers' traffic must be transmitted through the BAS server, one technique is to identify each such access network by one PID. It is generally unnecessary to further divide such access networks. On the other hand, it can be beneficial to combine several such access networks into a single PID.

A MAN usually consists of several access networks. The MANs are connected to a core network, whose network bandwidth resource may be costly for some networks. Thus, the ISP can define one or several MANS as one PID. It is also possible that the ISP deploys ALTO independently in some MANs.

2.2.2. Cost Map

In ALTO protocol, cost map configuration can be used to tune the traffic model. For example, to one optimization area of mobile network, the wireless uplink is very limited in data transmission, because of current mobile operators' configuration in 3G network. So the cost to this optimization area can be higher, thus this area can provide less traffic to other optimization areas.

In real ALTO service deployment, ISPs can plan its cost map according to its optimization requirement and network conditions.

2.3. Endpoint Service Configuration

In ALTO protocol, Endpoint Property Service allows ALTO clients to look up properties for individual endpoints. The endpoint cost service allows an ALTO service to return either numerical costs or ordinal costs (ranking) directly amongst Endpoints. Using these services, ISP can provide more detail connectivity and cost information of endpoint to upper application. For example in P2P service, ISP can use this service to provide the connectivity type of P2P clients to upper P2P operator.

Also in real deployment of ISP, different networks have different connectivity characters, for example, ADSL network has asymmetrical uplink and downlink bandwidth, as well as mobile network always have small bandwidth compare to fixed network. Because all clients access to Internet through ISPs, so ISPs have detail information of each client's access character. Thus ISP can provide these information to service operator.

In ALTO deployment of ISPs, there are some properties which are necessary to the endpoint property service. P2P or CDN application can benefit from that. We suggest to add the network type, uplinkbandwidth and downlinkbandwidth into endpoint property service. One example of these properties protocol usage is given below.

```
POST /endpoint/prop/lookup?prop=pid HTTP/1.1
Host: alto.example.com:6671
Content-Length: [TODO]
{
  "endpoints" : [ "192.0.2.34", "203.0.113.129" ]
}

HTTP/1.1 200 OK
Content-Length: [TODO]
Content-Type: application/alto
{
```

```
    "meta" : {
      "version" : 1,
      "status" : {
        "code" : 1
      }
    },
    "type" : "endpoint-property",
    "data": {
      "192.0.2.34" : { "pid": "PID1";
"networktype":"ADSL" ; " uplinkbandwidth ":"1M";" downlinkbandwidth
":" 4M"},
      "203.0.113.129" : { "pid": "PID3";
"networktype":"GSM";" uplinkbandwidth ":"0M";" downlinkbandwidth ":"
64K"}
    }
  }
```

3. ALTO Deployment Monitoring

In addition to providing configuration, an ISP providing ALTO may want to deploy a monitoring infrastructure to assess the benefits of ALTO and adjust its ALTO configuration.

To construct an effective monitoring infrastructure, the ISP should (1) define the performance metrics to be monitored; (2) and identify and deploy devices to collect data to compute the performance metrics. We discuss both below.

3.1. Monitoring Metrics

The monitoring of some performance metrics can be dependent on specific applications, and ALTO can be applied to multiple applications such as P2P and CDN. We focus on P2P applications.

3.1.1. Network Metrics

An ISP may monitor the impacts of ALTO on its network through a set of performance metrics. We enumerate some key metrics. We define the term domain as one or many groups of Endpoints. That is, one domain includes one PID or some PIDs. Endpoints and PID are defined in "draft-ietf-alto-protocol-05".

A specific set of metrics measuring the impacts of ALTO on networks can include the following:

- o Inter-domain ALTO-Integrated Application Traffic (Network metric): This metric includes total cross domain traffic generated by applications that utilize ALTO guidance. This

metric evaluates the impacts of ALTO on the inbound and outbound traffic of a domain.

- o Total Inter-domain Traffic (Network metric): This is similar to the preceding but focuses on all of the traffic, ALTO aware or not. One possibility is that some of the reduction of interdomain traffic by ALTO aware applications may This metric is always used with the preceding and the following metrics.
- o Intra-domain ALTO-Integrated Application Traffic (Network metric).
- o Network hop count (Network metric): This metric provides the average number of hops that traffic traverses inside a domain. ALTO may reduce not only traffic volume but also the hops. The metric can also indirectly reflect some application performance (e.g., latency).

3.1.2. Application Metrics

Each specific application can have its specific set of performance metrics. We give one example for file sharing.

- o Application download rate (Application metric): This metric measures application performance directly. Download means inbound traffic to one user. Global average means the average value of all users' download rates in one or more domains.
- o Application Client type audit(Application metric): this metric gives the audit of client types in ALTO service. The current types include fixed network client and mobile network client.

3.2. Monitoring Data Sources

The preceding metrics are derived from data sources. We identify four data sources.

3.2.1. Application Log Server

Many P2P applications deploy Log Servers to collect data.

3.2.2. P2P Clients

Some P2P applications may not have Log Servers. When available, P2P client logs can provide data.

3.2.3. OAM

Many ISPs deploy OAM systems to monitor IP layer traffic. An OAM provides traffic monitoring of every network device in its management area. It provides data such as link physical bandwidth and traffic volumes.

3.3. Application/ISP Monitoring Integration

3.3.1. Structure

As discussed in the preceding section, some data sources are from ISP while some others are from application. When there is a collaboration agreement between the ISP and an application, there can be an integrated monitoring system as shown in the figure below. In particular, an application developer may deploy Monitor Clients to communicate with Monitor Server of the ISP to transmit raw data from the Log Server or P2P clients of the application to the ISP.

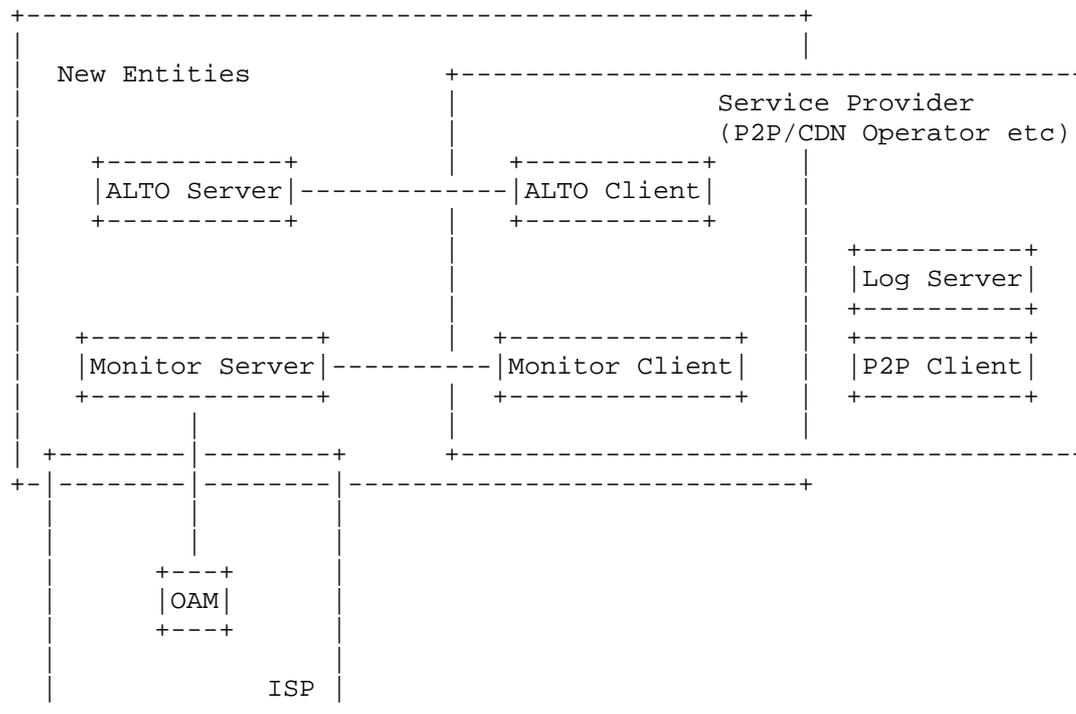


Figure 1

3.3.2. ALTO Monitoring Report Protocol

A potential report message format from the Monitor Client to the Monitor Server can be:

```

HTTP/1.1 200 OK
Content-Length: [TODO]
Content-Type: application/alto
{
  "meta" : {
    "version" : 1,
    "status" : {
      "code" : 1
    }
  },
  "metric1 name" : "value",
  "metric2 name" : "value",
}

```

4. IANA Considerations

This document makes no request of IANA.

Note to RFC Editor: this section may be removed on publication as an RFC.

5. Security Considerations

Multiple documents in the ALTO WG discuss security perspectives.

These documents complement this document.

6. References

6.1. Normative References

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

6.2. Informative References

[2] H. Xie, Y.R. Yang, A. Krishnamurthy, Y. Liu, and A.Silberschatz., "P4P:", In SIGCOMM 2008.

Appendix A. Acknowledgments

We thank the discussions with Kai Li.

Authors' Addresses

Xianghui Sun

China Telecom

Email: sunxh@ctbri.com.cn

Yang Richard Yang

Yale University

Email: yry@cs.yale.edu

