

ALTO
Internet-Draft
Intended status: Informational
Expires: Sep 11, 2012

Kai Lee
China Telecom
GuangYao.Jian
Xunlei network
March 11, 2012

ALTO and DECADE service trial within China Telecom
draft-lee-alto-chinatelecom-trial-04.txt

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 Sep 11, 2012.

Internet-Draft <ALTO&DECADE trial within China Telecom> March 11 2012

Copyright Notice

Copyright (c) 2012 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to [BCP 78](http://trustee.ietf.org/license-info) 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.

Abstract

This document reports the experience of China Telecom in a recent experiment with the ALTO service and P2P caches deployment. It is found that the deployment of the ALTO service significantly improves the capability of a Service Provider to affect the distribution of P2P traffic. It is also found that a traffic localized ALTO policy may decrease the download speed of a P2P user. However, the deployment of some P2P caches can compensate such influence.

Internet-Draft <ALTO&DECADE trial within China Telecom> March 11 2012

Table of Contents

1.	Introduction	4
2.	High level description of the trial	4
2.1.	Difference between standard ALTO protocol	4
2.2.	Difference between Comcast's trial	5
3.	Trial results	6
3.1.	Trial configuration	6
3.2.	Xunlei traffic distribution before the trial	7
3.3.	ALTO policy test	8
3.4.	P2P cache test	9
4.	Methods of data collection	9
5.	Configurations and algorithms in trial	10
5.1.	Configuration of PID MAP	10
5.2.	Algorithms of Xunlei using ALTO information	10
5.3.	Configuration of cache system	12
6.	Next steps	13
7.	Security Considerations	13
8.	IANA Considerations	13
9.	References	13
	Author's Addresses	13

Internet-Draft <ALTO&DECADE trial within China Telecom> March 11 2012

1. Introduction

Although another trial on P4P, the predecessor of the ALTO, is available by Comcast, the impact of ALTO on a large scale real network has never publicly reported. Such real network should post no limitation on either the number of contents or the number of users. This draft reports the experience of China Telecom in a recent experiment with the deployment of the ALTO service and P2P caches.

With over 60 million fixed-line broadband subscribers, China Telecom is the largest broadband service provider in China. It has one IP backbone network that cover all of the 31 provinces in China and about 200 MAN networks are managed by the provinces respectively. This trial was taken place in one province with 7 million broadband subscribers and about 11 MAN networks.

Xunlei, the cooperator of this trial, is a leading P2P service provider in China. Xunlei supports both file downloads and real time media streaming. In 2009, when was this trail occurring, it serves over 20 million users each day.

This trial is a joint effort of China Telecom and Xunlei. During this trial, China Telecom provided the following devices: an alto server to distribute ALTO information, some P2P caches to improve Xunlei service experience, and network management systems based on SNMP and DPI technology to monitor the traffic load within its backbone. Xunlei provided the P2P client and users. To support this trial, Xunlei modified its platform to support ALTO, and recorded operational information on its platform according to the requirement of China Telecom. Note that the client of Xunlei was not changed.

[2.](#) High level description of the trial

[2.1.](#) Difference between standard ALTO protocol

Note that ALTO protocol is still on progressing, in this trail, some modifications were made to the ALTO.

Firstly a server notification mechanism for the ALTO server is introduced. With this mechanism, the ALTO server notifies its clients the changes of network maps and cost maps. Thus, ALTO clients can make timely respond to the change of traffic optimizing policy(Note:different traffic optimizing policies mean different network maps and cost maps). This mechanism is mainly used to solve the problem related to alto effect evaluation. As we know the traffic within the IP backbone is highly periodical. For example, the traffic on each weekend is higher than which on the workday.As such, the evaluation traffic data to compare should be both collected in the

<Kai.Lee>

Expires Sep 11, 2012

[Page 4]

Internet-Draft <ALTO&DECADE trial within China Telecom> March 11 2012

workday or in the weekend. We need to change the traffic optimizing policy constantly to collect the traffic data with and without effect of alto service to evaluate the effect of ALTO service. That's why we need the server notification in this trial.

In this trail, ALTO clients were embedded in the trackers of Xunlei, not in the Xunlei clients. The reason for this is mainly for deployment consideration. There are hundreds of millions of Xunlei clients in use, To update these clients as the ALTO client in a short time is not feasible. However, according to the analysis of Xunlei, although both tracker based and tracker-less technology are adopted, the traffic does not controlled by the trackers is less than 15% of its total traffic. Based on this analysis, in this trial, Xunlei clients are not involved in the ALTO service which has negligible influence on the final evaluation of this trial. Such design can also reduce the load on the ALTO server.

Secondly, only map service is provided in this trial. Endpoint property service and Endpoint cost service were not supplied, as they are not essential for this trial.

[2.2.](#) Difference between Comcast's trial

Comcast has a trial with limited swarms, with the cooperation of Pando. According to the [RFC5632](#) "Comcast's ISP Experiences in a Proactive Network Provider Participation for P2P (P4P) Technical Trial", there are five swarms, and overall 57,000 peers are involved in that trial.

There are several differences between CT's trial and Comcast's trial:

1. The scope of trail: CT's trial covers the whole province with over 700 million broadband users. It lasted for over 4 months. There are countless swarms with all kinds of contents. from this point of view, our trial is more realistic than the previous trial from Comcast.
2. The usage of P2P cache: CT's trail differs from the Comcast's trail by the utilization of P2P cache. In this trail, the average download speed of a Xunlei client decreases a little with the degree of traffic localization increased. Thus P2P cache was introduced to compensate the decrease of download speed.
3. The evaluation method: In contrast to that all test data was collected by Pando client in Comcast's trial, we collect test data from two ways. Besides the data from Xunlei P2P client, we simultaneously collect the data from network operator's NMS

system.(such as data from SNMP reports and DPI(deep package inspection) device deployed on backbone). We can do this because Xunlei's p2p traffic occupy about 20% of whole backbone traffic flow in that province. This traffic flow will almost all be affected by alto service and it is big enough to be observed by CT's NMS system.

4. The implementation of ALTO: In this trial, only the P2P trackers are ALTO clients, but not those Xunlei clients. There are some reasons to do this:
 - a) To avoid the update all Xunlei clients and simplify the deployment of trial.
 - b) To lessen the alto server load.

- c) Above 85% of Xunlei traffic flow is controlled by Xunlei tracker, the traffic flow affected by DHT mechanism is less than 15%. An alto server dedicated for Xunlei tracker can affect majority of Xunlei traffic flow.

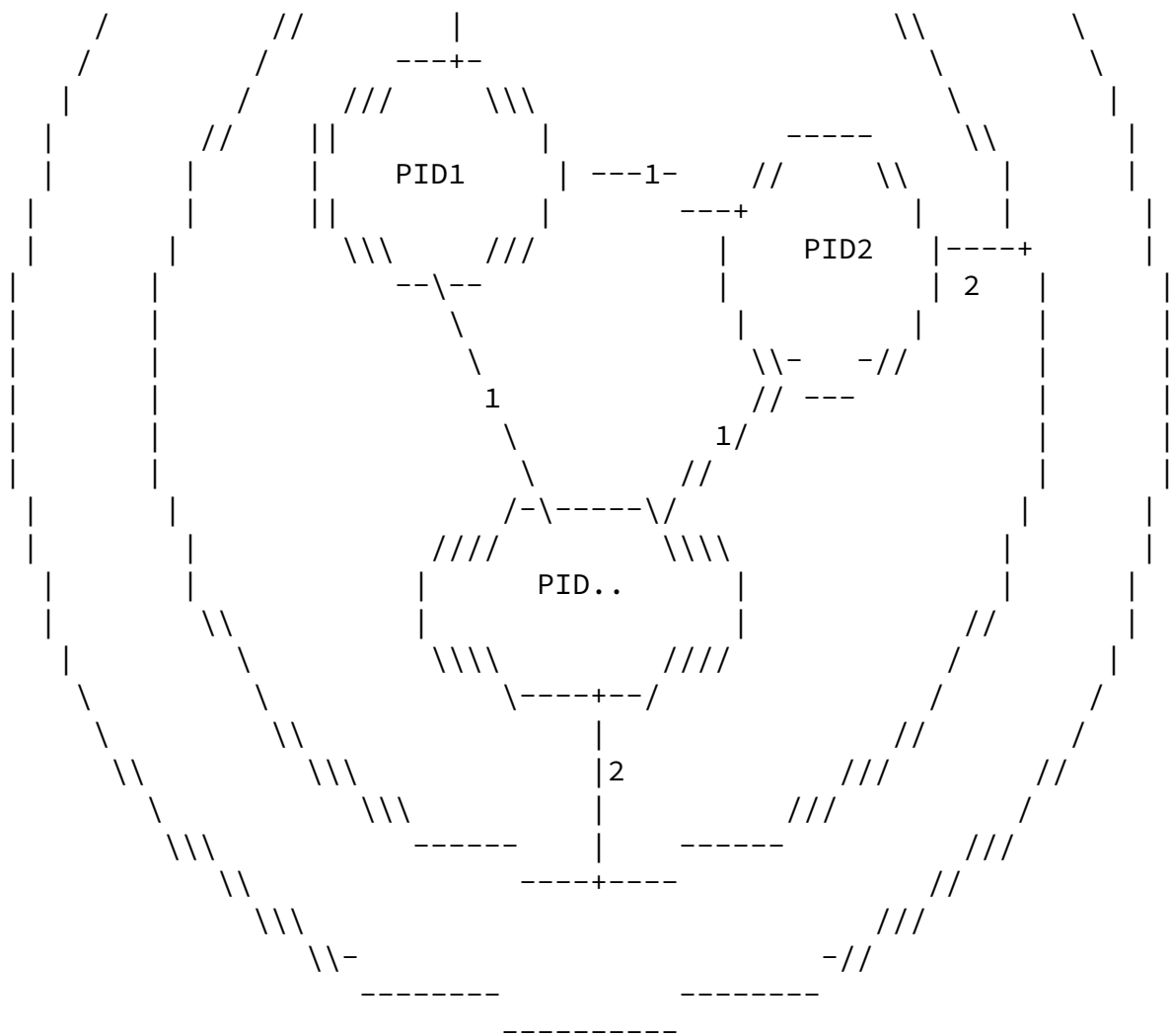
3. Trial results

3.1. Trial configuration

All Xunlei p2p client in the province and all contents that are requested or served by Xunlei P2P client in the province are involved in this trial. The trial environment is more realistic than Comcast's. A primary objective of this trial is to measure the effects of traffic localization and change of users download speed in comparison to those without alto service.

The test process is divided into two parts: first part is just applied the ALT0 server to measure the effects of traffic localization and change of P2P user experience. The second part is to introduce the P2P cache to the trial, to measure the improvement of user download speed, the bandwidth consumption and relationship with the scale of p2p cache and average download speed.

Our trial starts at 2009.6.12 and ends at 2009.10.18, lasting nearly four months. We do this trial by applying different traffic control policy to Xunlei tracker through ALTO service. There are two kinds of traffic control policy: One is optimized policy and the other is un-optimized policy. corresponding to the different traffic control policy there are two kinds of network map and cost map the ALTO server need to create. The optimized policy tries to localize the traffic by utilizing the information from ALTO server. Here is the network map and cost map corresponding to the optimized policy below.



The normal policy will just use the original Xunlei peer selection and traffic control rules though . The corresponding network map just has one PID with all IPs.

We usually change the alto policy in midnight of a day and send a notification to Xunlei tracker with notification mechanism. (<http://tools.ietf.org/id/draft-sun-alto-notification-02.txt>)

3.2. Xunlei traffic distribution before the trial

Before we do the trial , we collect the information about Xunlei'S

No	Data Item	Description	The way of collection
1	Peer distribution	24.6% is within the province, 75.4% is out of the province	Random sampling by Xunlei tracker 24 times one day
2	Traffic distribution	76.9% is inter-province traffic 23.1% is intra-province traffic	Random selecting peers to report their traffic flow

3.3. ALTO policy test

After we applied the alto optimized policy about 60% inter-province traffic has become The intra-province traffic. Below is the result that we observed on china telecom's network NMS system:

No	Data Item	Description	The way of collection
1	Outbound bandwidth	Decreased 42.77Gbps, about 50.61% of total Xunlei outbound traffic	Collecting max average outbound traffic of a day from the DPI system
2	Inbound/outbound bandwidth	outbound bandwidth decreased 31.58Gbps inbound bandwidth decreased 10.46Gbps	Collecting max average inbound/outbound traffic of a day from the snmp system

3.4. P2P cache test

In this trial we deployed 16 cache devices, each with 1.8TB SAS hard disks. The P2P cache system has 15Gbps links connected to the Internet. We cached the content according to its popularity.

No	Data Item	Description	The way of collection
1	Outbound bandwidth	Decreased 40Gbps, about 54.47% of total Xunlei outbound traffic	Collecting max average outbound traffic of a day from the DPI system
2	Inbound/outbound bandwidth	outbound bandwidth decreased 39.18Gbps inbound bandwidth decreased 28.3 Gbps	Collecting max average inbound/outbound traffic of a day from the snmp system
3	Average download speed	From 279KBps up to 294.5KBps	Collection from Xunlei OAM system

The P2P cache system occupancy ratio is about 80%. Bandwidth consumed is about 4-5Gbps. After deployed the P2P cache system, the traffic flow in the province has decreased a lot. Meanwhile the average download speed of Xunlei client has been increased.

4. Methods of data collection

In this trial we have two ways for information collection; one is to collect from p2p service provider such as Pando and Xunlei just like Comcast's trial. The other is to collect from ISP's network OAM system. Because the Xunlei's inter-province traffic flow is about 80Gbps that is large enough to be observed by ISP's network OAM system

1. Information from ISP's network OAM system and DPI system

- a) Inbound/outbound traffic flow statistic
- b) Xunlei traffic flow detected by DPI system. The DPI system just monitored the uplink of the province to China telecom's backbone.

2. Information from Xunlei

<Kai.Lee>

Expires Sep 11, 2012

[Page 9]

Internet-Draft <ALTO&DECADE trial within China Telecom> March 11 2012

- a) Inter-province/intra-province traffic flow.
- b) User average download speed.

[5.](#) Configurations and algorithms in trial

[5.1.](#) Configuration of PID MAP

- a) PID Map: We define 11 PIDs PID1-PID11 represent the 11 MANs of the trial network PID12 represents rest of the Internet
- b) Cost Map: Bidirectional cost between any PIDs from PID1 to PID11 has the same value 1 Bidirectional cost between PID12 and PIDi ($1 \leq i \leq 11$) has the same value 2

[5.2.](#) Algorithms of Xunlei using ALTO information

Xunlei is a hybrid application utilizing both trackers and DHT, About 85% of Xunlei traffic controlled by Xunlei trackers. In this trail ALTO clients just include the Xunlei trackers not include the Xunlei client. Just the traffic controlled by Xunlei tracker has been affected.

Before the trial Xunlei tracker peer selection algorithm is: Xunlei Peer selection algorithm depends on two properties: ISP ID and UC (upload capability), the peer selection priority is :

Same ISP ID > different ISP ID

Higher UC > lower UC

The peers with same ISP ID with the requesting peer have higher priority than those with different ISP ID. If peers have same ISP ID then the peers with higher UC have higher priority than those with lower UC.

After applying the ALTO information into the Xunlei peer selection algorithm. Xunlei changed his Peers select mechanism. All Xunlei

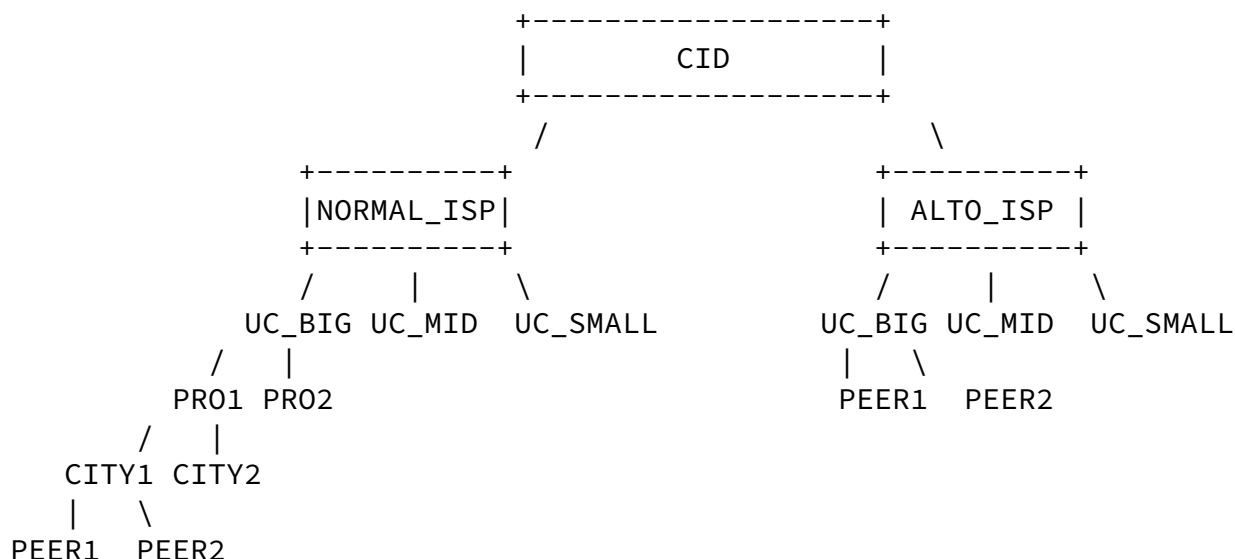
peers are organized in a tree structure which is indexed by CID(content ID), in the second level ALTO_ISP and normal_ISP represent the network of ISP with and without alto information. In this trial 11 MANs in trial province became 11 ALTO_ISPs. The third level is defined by different upload capability(UC) of peers. The fourth level of normal_ISP branch is the different provinces(PRO1,PRO2) of ISP, the fifth level of the normal_ISP is different city of ISP.

<Kai.Lee>

Expires Sep 11, 2012

[Page 10]

Internet-Draft <ALTO&DECADE trial within China Telecom> March 11 2012



The algorithms of cost between origination peer(peer_o) and destination peer(peer_d) is :

If (peer_o and peer_d both from ALTO_ISP)

 If (peer_o and peer_d in the same ALTO_ISP) then cost = 0;

 Else cost = 100000;

Else if (peer_o from ALTO_ISP and peer_d from normal_ISP) cost = 100000;

Else if (peer_o from normal_ISP and peer_d from ALTO_ISP) cost = 1000;

Else if (peer_o and peer_d both from normal_ISP){

```

If (peer_o and peer_d from different normal_ISP) cost =1000;

Else if (peer_o and peer_d from different province) cost = 100;

Else if (peer_o and peer_d from different city) cost = 10;

Else cost =0;

}

```

The peer select mechanism is lower cost peers will have higher priority

The updated peer selection mechanism is not the best mechanism. For example a peer in MAN2 is supposed to be better choice than the peers which not located in china telecom's network when a peer in

<Kai.Lee>

Expires Sep 11, 2012

[Page 11]

Internet-Draft <ALTO&DECADE trial within China Telecom> March 11 2012

MAN1 send a content request to tracker. But this mechanism will select the peer out of china telecom's network first then select the peer in the MAN2. Before we defined the network map with 12 PIDs. We first defined a network map with just 2 PIDs. PID1 represent the trial province and PID2 represent the other network to test the backbone traffic saving effect of ALTO service. The test result show that the network map with 12PIDs has almost same backbone traffic saving effect compared to the network map with 2 PIDs. So in the trial we deployed this mechanism.

The other change is the number of returned peers from Xunlei tracker . If a listing request is from the trial province, the maximum # of returned peers from Xunlei tracker is set to 120, not the normal case of 500.

[5.3](#). Configuration of cache system

Before we deploy the cache system we have made some statistics about relationship of content popularity and network traffic caused by content with different popularity in trial province.

+-----+	+-----+	+-----+	+-----+
content	total	total	proportion
popularity	size(GB)	traffic	of total
		(Gbps)	traffic(%)

top 10	18.9	1.34	9.3
top 20	29.3	1.68	11.7
top 50	51.8	2.28	15.9
top 100	93.6	2.89	20.1
top 500	418.7	4.74	33
top 1000	812.4	5.88	40.9
top 2000	1518.6	7.16	49.8
top 5000	3551	8.89	61.9

Our cache system has limited storage and access bandwidth so we need to know which content is most "valuable" to be cached. According the statistics from Xunlei if a downloading task is fed over 100 peers , this task always can get the maximum download speed(this speed depends on the peer's access network, in the trial the average access speed of user is about 2Mbps). The top 2000 popular content almost all have over 100 seeds in trial province. That means the top 2000 popular contents don't need be cached. Our cache policy is just cache the content which's popularity rank behind 2000.

6. Next steps

The alto mechanism is very effective to optimize the traffic flow.If alto can cooperate with p2p cache or other service performance enhancement mechanism, it will be more practical. The ALTO service's effect depends on the SP such as Xunlei,pando how to use it. The mechanism such as peer selection mechanism and content cache mechanism need to be studied.

7. Security Considerations

High-level security considerations can be found in the [[draft-ietf-alto-problem-statement](#)].

8. IANA Considerations

This document requests the registration of a new media type:
"application/alto"

9. References

- [RFC5693] Seedorf, J. and E. Burger, "Application-Layer Traffic Optimization (ALTO) Problem Statement", [RFC 5693](#), October 2009.
- [RFC5632] Griffiths, C., Livingood, J., Popkin, L., Woundy, R., and Y. Yang, "Comcast's ISP Experiences in a Proactive Network Provider Participation for P2P (P4P) Technical Trial", [RFC 5632](#), September 2009.
- [I-D.ietf-alto-reqs] Kiesel, S., Popkin, L., Previdi, S., Woundy, R., and Y. Yang, "Application-Layer Traffic Optimization (ALTO) Requirements", [draft-ietf-alto-reqs-01](#) (work in progress), July 2009.
- [I-D.penno-alto-protocol] Penno, R. and Y. Yang, "ALTO Protocol", [draft-ietf-alto-protocol-01](#) (work in progress), July 2009.

Author's Addresses

Kai Lee
China Telecom Beijing Research Institute
Email: leekai@ctbri.com.cn

GuangYao Jian
Xunlei Network
Email: jianguangyao@xunlei.com

<Kai.Lee>

Expires Sep 11, 2012

[Page 13]

Internet-Draft <ALTO&DECADE trial within China Telecom> March 11 2012

