Network Working Group	J. Seedorf	
Internet-Draft	NEC	
Intended status: Informational	E. Burger	
Expires: September 9, 2009	This Space for Sale	
	March 08, 2009	

Application-Layer Traffic Optimization (ALTO) Problem Statement
draft-marocco-alto-problem-statement-05

Status of this Memo

Copyright Notice

Abstract

Peer-to-peer applications, such as file sharing, real-time communication, and live media streaming, use a significant amount of Internet resources. Such applications often transfer large amounts of data in direct peer-to-peer connections. However, they usually have little knowledge of the underlying network topology. As a result, they may choose their peers based on measurements and statistics that, in many situations, may lead to suboptimal choices. This document describes problems related to optimizing traffic generated by peer-to-peer applications and associated issues such optimizations raise in the use of network-layer information.

Table of Contents

---

1. Introduction

TOC

Peer-to-peer (P2P) applications, such as file sharing, real-time communication, and live media streaming, use a significant amount of Internet resources [WWW.cachelogic.picture] (Parker, A., “The true picture of peer-to-peer filesharing,” .) [WWW.wired.fuel] (Glasner, J., “P2P fuels global bandwidth binge,” .). Different from the client/server architecture, P2P applications access resources such as files or media relays distributed across the Internet and exchange large amounts of data in connections that they establish directly with nodes sharing such resources.

One advantage of P2P systems results from the fact that the resources such systems offer are often available through multiple replicas. However, applications generally do not have reliable information of the underlying network and thus have to select among available instances based on information they deduce from empirical measurements that, in some situations, lead to suboptimal choices. For example, one popular metric is an estimation of round-trip time. This choice occurs before actual data transmission begins and thus before the peer can deduce actual throughput. This is one reason why a peer selection algorithm that simply uses round-trip time often results in a sub-optimal choice of peers.

Many of today's P2P systems use an overlay network consisting of direct peer connections. Such connections often do not account for the underlying network topology. In addition to having suboptimal performance, such networks can lead to congestion and cause serious inefficiencies. As shown in [ACM.fear] (Karagiannis, T., Rodriguez, P., and K. Papagiannaki, “Should ISPs fear Peer-Assisted Content Distribution?,” .), traffic generated by popular P2P applications often cross network boundaries multiple times, overloading links which are frequently subject to congestion [ACM.bottleneck] (Akella, A., Seshan, S., and A. Shaikh, “An Empirical Evaluation of WideArea Internet Bottlenecks,” .). Moreover, such transits, besides resulting in a poor experience for the user, can be quite costly to the network operator. Recent studies [ACM.ispp2p] (Aggarwal, V., Feldmann, A., and C. Scheideler, “Can ISPs and P2P systems co-operate for improved performance?,” .) [WWW.p4p.overview] (Xie, H., Krishnamurthy, A., Silberschatz, A., and R. Yang, “P4P: Explicit Communications for Cooperative Control Between P2P and Network Providers,” .) [ACM.ono] (Choffnes, D. and F. Bustamante, “Taming the Torrent: A practical approach to reducing cross-ISP traffic in P2P systems,” .) show a possible solution to this problem. Internet Service Providers (ISP), network operators or third parties can collect reliable network information. This information includes relevant information such as topology or instantaneous bandwidth available. Normally, such information is rather "static", i.e., information which can change over time but on a much longer time scale than information used for congestion control on the transport layer. By providing this information to P2P applications, it would be possible to greatly increase application performance, reduce congestion and optimize the overall traffic across different networks. Presumably both, the application and the network operator, can benefit from the fact that such information is being provided to (and used by) the application. Thus, network operators have an incentive to provide (either directly themselves or indirectly through a third party) such information and applications have an incentive to use such information. This document gives the problem statement of optimizing traffic generated by P2P applications using information provided by a separate party.

Section 3 (The Problem) introduces the problem. Section 4 (Use Cases) describes some use cases where both P2P applications and network operators would benefit from a solution to such a problem. Section 5 (The Problem in Detail) describes the main issues to consider when designing such a solution.

1.1. Research or Engineering?

TOC

The papers [I-D.bonaventure-informed-path-selection] (Saucez, D. and B. Donnet, “The case for an informed path selection service,” February 2008.) and [ACM.ispp2p] (Aggarwal, V., Feldmann, A., and C. Scheideler, “Can ISPs and P2P systems co-operate for improved performance?,” .) [WWW.p4p.overview] (Xie, H., Krishnamurthy, A., Silberschatz, A., and R. Yang, “P4P: Explicit Communications for Cooperative Control Between P2P and Network Providers,” .) are examples of contemporary solution proposals that address the problem described in this document. Moreover, these proposals have encouraging simulation and field test results. These and similar, independent, solutions all consist of two essential parts:

- a discovery mechanism which a P2P application uses to find a reliable information source;
- a protocol P2P applications use to query such sources in order to retrieve the information needed to perform better-than-random selection of the endpoints providing a desired resource.

It is not easy to foresee how such solutions would perform in the Internet, but a more accurate evaluation would require representative data collected from real systems by a critical mass of users. However, wide adoption will probably never happen without an agreement on a common solution based on an open standard.

2. Definitions

TOC

The following terms have special meaning in the definition of the Application-Layer Traffic Optimization (ALTO) problem.

Application: A distributed communication system (e.g., file sharing) that uses the ALTO service to improve its performance (or quality of experience) while optimizing resource consumption in the underlying network infrastructure. Applications may use the P2P model to organize themselves, use the client-server model, or use a hybrid of both.

Peer: A specific participant in an application. Colloquially, a peer refers to a participant in a P2P network or system, and this definition does not violate that assumption. If the basis of the application is the client-server or hybrid model, then the usage of the terms "client" and "server" disambiguates the peer's role.

P2P: Peer-to-Peer.

Resource: Content, such as a file or a chunk of a file or a server process, for example to relay a media stream or perform a computation, which applications can access. In the ALTO context, a resource is often available in several equivalent replicas. In addition, different peers share these resources, often simultaneously.

Resource Identifier: An application layer identifier used to identify a resource, no matter how many replicas exist.

Resource Provider: For P2P applications, a resource provider is a specific peer that provides some resources. For client-server or hybrid applications, a provider is a server that hosts a resource.

Resource Consumer: For P2P applications, a resource consumer is a specific peer that needs to access resources. For client-server or hybrid applications, a consumer is a client that needs to access resources.

Transport Address: All address information that a resource consumer needs to access the desired resource at a specific resource provider. This information usually consists of the resource provider's IP address and possibly other information, such as a transport protocol identifier or port numbers.

Overlay Network: A virtual network consisting of direct connections on top of another network, established by a group of peers.

Resource Directory: An entity that is logically separate from the resource consumer that assists a resource consumer to identify a set of resource providers. Some P2P applications refer to the resource directory as a P2P tracker.

Host Location Attribute: Information about the location of a host in the network topology. The ALTO service gives recommendations based on this information. A host location attribute may consist of, for example, an IP address, an address prefix or address range that contains the host, an autonomous system (AS) number, or any other localization attribute. These different options may provide different levels of detail. Depending on the system architecture, this may have implications on the quality of the recommendations ALTO is able to provide, on whether recommendations can be aggregated, and on how much privacy-sensitive information about users might be disclosed to additional parties.

ALTO Service: Several resource providers may be able to provide the same resource. The ALTO service gives guidance to a resource consumer or resource directory about which resource provider(s) to select, in order to optimize the client's performance or quality of experience while optimizing resource consumption in the underlying network infrastructure.

ALTO Server: A logical entity that provides interfaces to query the ALTO service.

ALTO Client: The logical entity that sends ALTO queries. Depending on the architecture of the application one may embed it in the resource consumer or in the resource directory.

ALTO Query: A message sent from an ALTO client to an ALTO server, which requests guidance from the ALTO Service.

ALTO Response: A message sent from an ALTO server to an ALTO client, which contains guiding information from the ALTO service.

ALTO Transaction: An ALTO transaction consists of an ALTO query and the corresponding ALTO response.

Local Traffic: Traffic that stays within the network infrastructure of one Internet Service Provider (ISP). This type of traffic usually results in the least cost for the ISP.

Peering Traffic: Internet traffic exchanged by two Internet Service Providers whose networks connect directly. Apart from infrastructure and operational costs, peering traffic is often free to the ISPs, within the contract of a peering agreement.

Transit Traffic: Internet traffic exchanged on the basis of economic agreements amongst Internet Service Providers (ISP). An ISP generally pays a transit provider for the delivery of traffic flowing between its network and remote networks that the ISP does not have a direct connection.

Application Protocol: A protocol used by the application for establishing an overlay network between the peers and exchanging data on it, as well as for data exchange between peers and resource directories if applicable. These protocols play an important role in the overall ALTO architecture, however, defining them is out of the scope of the ALTO WG.">

ALTO Client Protocol: The protocol used for sending ALTO queries and ALTO replies between ALTO client and ALTO Server.

Provisioning Protocol:
A protocol used for populating the ALTO server with topology-related information.

Inter-ALTO Server Protocol: The protocol used for synchronization, query forwarding, or referral between ALTO servers that have been provisioned with only partial knowledge of the topology-related information (e.g., on a per-domain basis).

```
                                                   +------+
                                               +-----+  |  Peers
   +-----+       +-------+    +=====|     |--+
   |     |.......|       |====+     +--*--+
   +-----+       +-------+    |        *
 Source of        ALTO        |        *
 topological      service     |     +--*--+
 information                  +=====|     | Super-peer
                                    +-----+ (Tracker, proxy)
 Legend:
 === ALTO client protocol
 *** Application protocol (out of scope)
 ... Provisioning or initialization (out of scope)
```

Figure 1 - Overview of protocol interaction between ALTO elements

---

Figure 1 shows the scope of the ALTO client protocol: Peers or super-peers can use such a protocol to query an ALTO-service. The mapping of topological information onto an ALTO service as well as the application protocol interaction between peers and super-peers are out of scope for the ALTO client protocol.

---

3. The Problem

TOC

Network engineers have been facing the problem of traffic optimization for a long time and have designed mechanisms like MPLS (Rosen, E., Viswanathan, A., and R. Callon, “Multiprotocol Label Switching Architecture,” January 2001.) [RFC3031] and DiffServ (Grossman, D., “New Terminology and Clarifications for Diffserv,” April 2002.) [RFC3260] to deal with it. The problem these protocols address consists in finding (or setting) optimal routes for packets traveling between specific source and destination addresses and based on requirements such as low latency, high reliability, and priority. Such solutions are usually implemented at the link and network layers, and tend to be almost transparent. At best, applications can only "mark" the traffic they generate with the corresponding properties.

However, P2P applications that are today posing serious challenges to Internet infrastructures do not benefit much from the above route-based techniques. Cooperating with external services aware of the network topology could greatly optimize the traffic the P2P application generates. In fact, when a P2P application needs to establish a connection, the logical target is not a host, but rather a resource (e.g., a file or a media relay) that is often available in multiple instances on different peers. Selection of the closest one -- or, in general, the best from an overlay topological proximity -- has much more impact on the overall traffic than the route followed by its packets to reach the endpoint.

Optimization of peer selection is particularly important in the initial phase of the process. Consider a P2P protocol such as BitTorrent, where a querying peer receives a list of candidate destinations where a resource resides. From this list, the peer will derive a smaller set of candidates to connect to and exchange information with. In another example, a streaming video client may be provided with a list of destinations from which it can stream content. In both cases, the use of topology information in an early stage will allow applications to improve their performance and will help ISPs make a better use of their network resources. In particular, an economic goal for ISPs is to reduce the transit traffic on interdomain links.

Addressing the Application-Layer Traffic Optimization (ALTO) problem means, on the one hand, deploying an ALTO service to provide applications with information regarding the underlying network and, on the other hand, enhancing applications in order to use such information to perform better-than-random selection of the endpoints they establish connections with.

---

4. Use Cases

TOC

4.1. File sharing

TOC

File sharing applications allow users to search for content shared by other users and download it. Typically, search results consist of many instances of the same file (or chunk of a file) available from multiple

sources. The goal of an ALTO solution is to help peers find the best ones according to the underlying networks.
On the application side, integration of ALTO functionalities may happen at different levels. For example, in the completely decentralized Gnutella network, selection of the best sources is totally up to the user. In systems like BitTorrent and eDonkey, central elements such as trackers or servers act as mediators. Therefore, in the former case, optimization would require modification in the applications, while in the latter it could just be implemented in some central elements.

4.2. Cache/Mirror Selection

TOC

Providers of popular content like media and software repositories usually resort to geographically distributed caches and mirrors for load balancing. Selection of the proper mirror/cache for a given user is today based on inaccurate geolocation data, on proprietary network location systems or often delegated to the user himself. An ALTO solution could be easily adopted to ease such a selection in an automated way.

4.3. Live Media Streaming

TOC

P2P applications for live streaming allow users to receive multimedia content produced by one source and targeted to multiple destinations, in a real-time or near-real-time way. This is particularly important for users or networks that do not support multicast. Peers often participate in the distribution of the content, acting as both receivers and senders. The goal of an ALTO solution is to help peers to find the best sources and the best destinations for media flows they receive and relay.

4.4. Realtime Communications

TOC

P2P real-time communications allow users to establish direct media flows for real-time audio, video, and real-time text calls or to have text chats. In the basic case, media flows directly between the two endpoints. However, unfortunately a significant portion of users have limited access to the Internet due to NATs, firewalls or proxies. Thus, other elements need to relay the media. Such media relays are distributed over the Internet with a public addresses. An ALTO solution needs to help peers to find the best relays.

4.5. Distributed Hash Tables

TOC

Distributed hash tables (DHT) are a class of overlay algorithms used to implement lookup functionalities in popular P2P systems, without using centralized elements. In such systems, peers maintain addresses of other peers participating in the same DHT in a routing table, sorted according to specific criteria. An ALTO solution will provide valuable information for DHT algorithms.

5. The Problem in Detail

TOC

This section introduces some aspects to keep in consideration when designing an ALTO service to provide applications with information they can use to perform better-than-random peer selection.

5.1. ALTO Service Providers

TOC

At least three different kinds of entities can provide ALTO services:

1. Network operators: usually have full knowledge of the network they administer and are aware of the topology and policies that transit and peering traffic are subject to;

2. Third parties: are entities different from the network operators, but which may have collected network information. Examples of such entities are content delivery networks like Akamai, which control wide and highly distributed infrastructures, or companies providing an ALTO service on behalf of ISPs (and thus acquire the information from the ISPs themselves);

3. User communities: run distributed algorithms, for example for estimating the topology of the Internet.

TOC

5.2. Discovery of ALTO servers

As a direct consequence of the totally decentralized architecture of the Internet, it seems almost impossible to centralize all information P2P applications may need to optimize traffic they generate. Therefore, any solution for the ALTO problem will need to specify a mechanism for applications to find a proper ALTO server to query.
It is important to note that, depending on the implementation of the ALTO service, an ALTO server could be a centralized entity, for example deployed by the network operator, as well as a ephemeral node participating in a distributed algorithm.

5.3. User Privacy

TOC

Information provided by the ALTO client querying the ALTO server could help increase the level of accuracy in the replies. For example, if the querying client indicates what kind of application it is using (e.g. real-time communications or bulk data transfer), the server will be able to indicate priorities in its replies accommodating the requirements of the traffic the application will generate. However, it is important that for using an ALTO service the application does not have to disclose information it may consider sensitive.

5.4. Topology Hiding

TOC

Operators can play an important role in addressing the ALTO problem, but they generally consider network information they own to be confidential. Therefore, in order to succeed and achieve wide adoption, any solution should provide a method to help P2P applications in peer selection without explicitly disclosing topology of the underlying network.

5.5. Coexistence with Caching

TOC

Caching is a common approach to optimizing traffic generated by applications that require large data transfers. In some cases, such techniques have proven to be extremely effective in both enhancing user experience and saving network resources. However, they have two main limits in respect to the solutions based on the provision of topology information:

1. Application specificity: since a cache is meant to replace the source of the content being accessed -- either explicitly or transparently -- it must be able to speak the same protocol with the querying peer. For this reason, caching solutions can be reasonably adopted only for the most popular applications, such as HTTP and BitTorrent.

2. Content awareness: since caches need to store the content being delivered, they are subject to legal issues whenever the user does not have the right to access or distribute such content. This limitation makes caching approaches that do not (or cannot) support digital rights management unusable for distributing copyrighted material. Since, it is very difficult for an abstract file sharing proxy to know all of the legal parameters around distributing content, this makes caching unusable for many file-sharing systems. Since this is a legal and not technical issue, the solution would be at the legal, not network, layer.

In general, solutions based on provision of topology information need not interfere with caching. In fact, if the ALTO service used by applications is aware of the presence of caches, the service can indicate this in its response, marking them with higher priorities to achieve greater optimization.

---

6. Security Considerations

TOC

The approach proposed in this document asks P2P applications to delegate a portion of their routing capability to third parties. This gives the third party a significant role in P2P systems.
In the case where the network operator deploys an ALTO solution, it is conceivable that the P2P community would consider it hostile because the operator could, for example:

- redirect applications to corrupted mediators providing malicious content;

- track connections to perform content inspection or logging; and

- apply policies based on criteria other than network efficiency. For example, the service provider may suggest routes sub-optimal from the user's perspective to avoid peering points regulated by inconvenient economic agreements.

It is important to note that ALTO is completely optional for P2P applications and its purpose is to help improve performance of such applications. If, for some reason, it fails to achieve this purpose, it would simply fail to gain popularity and the P2P community would not use it.

Even in cases where the ALTO service provider maliciously alters results returned by queries after ALTO has gained popularity (i.e., the service provider plays well for a while to become popular and then starts misbehaving), it would be easy for P2P application maintainers and users to revert to solutions that are not using it.

---

7. IANA Considerations

TOC

None.

---

8. Acknowledgments

TOC

The basis of this document is draft-marocco-alto-problem-statement, written by Enrico Marocco and Vijay Gurbani. The authors of this draft continued editing the previous version in agreement with the original authors.

Vinay Aggarwal and the P4P working group conducted the research work done outside the IETF. Emil Ivov, Rohan Mahy, Anthony Bryan, Stanislav Shalunov, Laird Popkin, Stefano Previdi, Reinaldo Penno, Dimitri Papadimitriou, Sebastian Kiesel, and many others provided insightful discussions, specific comments and much needed corrections.

Thanks in particular to Richard Yang for several reviews.

---

9. Informative References

TOC

[ACM.bottleneck]	Akella, A., Seshan, S., and A. Shaikh, "An Empirical Evaluation of WideArea Internet Bottlenecks," Proceedings of ACM SIGCOMM, October 2003.
[ACM.fear]	Karagiannis, T., Rodriguez, P., and K. Papagiannaki, "Should ISPs fear Peer-Assisted Content Distribution?," In ACM USENIX IMC, Berkeley 2005.
[ACM.ispp2p]	Aggarwal, V., Feldmann, A., and C. Scheideler, "Can ISPs and P2P systems co-operate for improved performance?," In ACM SIGCOMM Computer Communications Review (CCR), 37:3, pp. 29-40.
[ACM.ono]	Choffnes, D. and F. Bustamante, "Taming the Torrent: A practical approach to reducing cross-ISP traffic in P2P systems," Proceedings of ACM SIGCOMM, August 2008.
[I-D.bonaventure-informed-path-selection]	Saucez, D. and B. Donnet, "The case for an informed path selection service," draft-bonaventure-informed-path-selection-00 (work in progress), February 2008 (TXT).
[RFC3031]	Rosen, E., Viswanathan, A., and R. Callon, "Multiprotocol Label Switching Architecture," RFC 3031, January 2001 (TXT).
[RFC3260]	Grossman, D., "New Terminology and Clarifications for Diffserv," RFC 3260, April 2002 (TXT).
[SIGCOMM.resprox]	Gummadi, K., Gummadi, R., Ratnasamy, S., Gribble, S., Shenker, S., and I. Stoica, "The impact of DHT routing geometry on resilience and proximity," Proceedings of ACM SIGCOMM, August 2003.
[WWW.cachelogic.picture]	Parker, A., "The true picture of peer-to-peer filesharing."
[WWW.p4p.overview]	Xie, H., Krishnamurthy, A., Silberschatz, A., and R. Yang, "P4P: Explicit Communications for Cooperative Control Between P2P and Network Providers."
[WWW.wired.fuel]	Glasner, J., "P2P fuels global bandwidth binge."

Authors' Addresses

TOC

	Jan Seedorf
	NEC Laboratories Europe, NEC Europe Ltd.
	Kurfuersten-Anlage 36
	Heidelberg 69115
	Germany
Phone:	+49 (0) 6221 4342 221
Email:	jan.seedorf@nw.neclab.eu
URI:	http://www.nw.neclab.eu
	Eric W. Burger
	This Space for Sale
	New Hampshire
	USA
Phone:	
Fax:	+1 530 267 7447
Email:	eburger@standardstrack.com
URI:	http://www.standardstrack.com