Network Working Group
Request for Comments: 5693
Category: Informational

J. Seedorf
NEC
E. Burger
Neustar Inc.

October 2009

Application-Layer Traffic Optimization (ALTO) Problem Statement

Abstract

Distributed applications -- such as file sharing, real-time communication, and live and on-demand media streaming -- prevalent on the Internet use a significant amount of network resources. Such applications often transfer large amounts of data through connections established between nodes distributed across the Internet with little knowledge of the underlying network topology. Some applications are so designed that they choose a random subset of peers from a larger set with which to exchange data. Absent any topology information guiding such choices, or acting on suboptimal or local information obtained from measurements and statistics, these applications often make less than desirable choices.

This document discusses issues related to an information-sharing service that enables applications to perform better-than-random peer selection.

Status of This Memo

This memo provides information for the Internet community. It does not specify an Internet standard of any kind. Distribution of this memo is unlimited.

Copyright Notice

Copyright (c) 2009 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

1.1. Overview

Distributed applications, both peer-to-peer (P2P) and client/server used for file sharing, real-time communication, and live and on-demand media streaming, use a significant amount of network capacity and CPU cycles in the routers [WWW.wired.fuel]. In contrast to centralized applications, distributed 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 highly distributed 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 the available peers that provide such replicas randomly or based on information they deduce from partial observations that, in some situations, lead to suboptimal choices. For example, one peer-selection algorithm is based only on the measurements during initial connection establishment between two peers. Since actual data transmission does not begin, the algorithm measures only the

round-trip time and cannot reliably deduce actual throughput between the peers. Thus, such a peer-selection algorithm that simply uses round-trip time may result in a suboptimal 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], traffic generated by popular P2P applications often cross network boundaries multiple times, overloading links that are frequently subject to congestion [ACM.bottleneck]. Moreover, such transits, besides resulting in a poor experience for the user, can be quite costly to the network operator.

Recent studies ([ACM.ispp2p], [WWW.p4p.overview], [ACM.ono]) show a possible solution to this problem. Internet Service Providers (ISPs), network operators, or third parties can collect more reliable network information. This information includes relevant information such as topology or link capacity. Normally, such information changes on a much longer time scale than information used for congestion control on the transport layer. Providing this information to P2P applications can enable them to apply better-than-random peer selection with respect to the underlying network topology. As a result, it may be possible to increase application performance, reduce congestion, and decrease the overall amount of traffic across different networks. Presumably, both applications and the network operator can benefit from such information. Thus, network operators have an incentive to provide, either directly themselves or indirectly through a third party, such information; applications have an incentive to use such information. This document discusses issues related to an information-sharing service that enables applications to perform better-than-random peer selection.

Section 2 provides definitions. Section 3 introduces the problem. Section 4 describes some use cases where both P2P applications and network operators benefit from a solution to such a problem. Section 5 describes the main issues to consider when designing such a solution. Note a companion document to this document, "Application-Layer Traffic Optimization (ALTO) Requirements" [ALTO-REQS], goes into the details of these issues.

1.2. State-of-the-Art

The papers [ACM.ispp2p], [PATH-SEL], and [WWW.p4p.overview] present 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 that a P2P application uses to find a reliable information source, and
- a protocol that 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 clear how such solutions will perform if deployed globally on the Internet. However, wide adoption is unlikely without agreement on a common solution, based upon an open standard.

2. Definitions

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 improving 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 (i.e., a mixture between the P2P model and the client-server model).

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) that 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 and established by a group of peers.

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

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

ALTO Server: A logical entity that provides interfaces to the queries to 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 and/or in the resource directory.

ALTO Query: A message sent from an ALTO client to an ALTO server; it requests guidance from the ALTO service.

ALTO Response: A message that contains guiding information from the ALTO service as a reply to an ALTO query.

ALTO Transaction: A transaction that 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 (ISPs). An ISP generally pays a transit provider for the delivery of traffic flowing between its network and remote networks to which 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 an ALTO client and ALTO server.

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

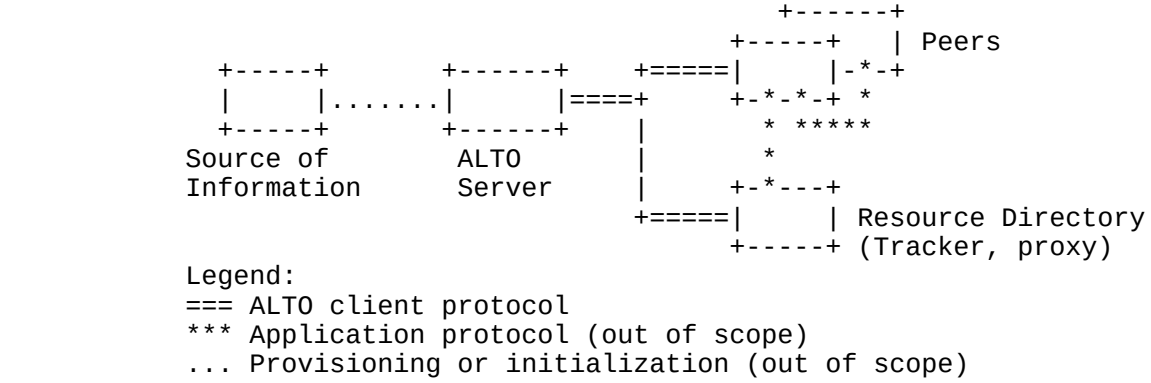

Figure 1: Overview of Protocol Interaction between ALTO Elements

Figure 1 shows the scope of the ALTO client protocol: peers or resource directories can use such a protocol as ALTO clients to query an ALTO server. The mapping of topological information onto an ALTO

service as well as the application protocol interaction between peers and resource directories are out of scope for the ALTO client protocol.

3. The Problem

Network engineers have been facing the problem of traffic optimization for a long time and have designed mechanisms like MPLS [RFC3031] and Diffserv [RFC3260] to deal with it. The problem these protocols address consists in finding (or setting) optimal routes (or optimal queues in routers) for packets traveling between specific source and destination addresses. Solutions are 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.

However, distributed applications in general and, in particular, bandwidth-greedy P2P applications that are used, for example, for file sharing, cannot directly use the aforementioned techniques. By cooperating with external services that are aware of the network topology, applications could greatly improve the traffic they generate. In fact, when a P2P application needs to establish a connection, the logical target is not a stable host, but rather a resource (e.g., a file or a media relay) that can be available in multiple instances on different peers. Selection of a good host from an overlay topological proximity has a large impact on the overall traffic generated.

> Note that while traffic considerations are important, several other factors also play a role on the performance experienced by users of distributed applications. These include the need to avoid overloading individual nodes, fetching rare pieces of a file before those pieces are available at a multiplicity of nodes, and so on. However, better information about topological conditions does improve the overall selection algorithm on an important aspect.

Better-than-random peer selection is helpful in the initial phase of the process. Consider a P2P protocol in which 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 with which they establish connections.

4. Use Cases

4.1. File sharing

File-sharing applications allow users to search for content shared by other users and to download respective resources from other users. For instance, search results can 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, improvement would require modification in the applications, while in the latter it could just be implemented in some central elements.

4.2. Cache/Mirror Selection

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

4.3. Live Media Streaming

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 a peer to find effective communicating peers that exchange the media content.

4.4. Real-Time Communications

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. Unfortunately, however, 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 public addresses. An ALTO solution needs to help peers find the best relays.

4.5. Distributed Hash Tables

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

5. Aspects of the Problem

This section introduces some aspects of the problem that some people may not be aware of when they first start studying the problem space.

5.1. Information Provided by an ALTO Service

The goal of an ALTO service is to provide applications with information they can use to perform better-than-random peer selection. In principle, there are many types of information that can help applications in peer selection. However, not all of the information to be conveyed is amenable to an ALTO-like service. More specifically, 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 of an ALTO-like, out-of-band technique at the application layer. An ALTO solution for congestion will either have outdated information or must be contacted too frequently by applications. And finally, information such as end-to-end delay and available bandwidth can be more accurately measured by applications, themselves.

The kind of information that is meaningful to convey to applications via an out-of-band ALTO service is any information that applications cannot easily obtain themselves and that changes on a much longer time scale than the instantaneous information used for congestion control on the transport layer. Examples for such information are operator’s policies, geographical location or network proximity

(e.g., the topological distance between two peers), the transmission costs associated with sending/receiving a certain amount of data to/from a peer, or the remaining amount of traffic allowed by a peer's operator (e.g., in case of quotas or limited flat-rate pricing models).

5.2. ALTO Service Providers

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

1. Network operators. Network operators usually have full knowledge of the network they administer and are aware of their network topology and policies.

2. Third parties. Third parties are entities separate from network operators but that may either have collected network information or have arrangements with network operators to learn the 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.

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

5.3. ALTO Service Implementation

It is important for the reader to understand there are significant user communities that expect an ALTO server to be a centralized service. Likewise, there are other user communities that expect the ALTO service be a distributed service, possibly even based on or integrating with a P2P service.

As a result, one can reasonably expect there to be some sort of service-discovery mechanism to go along with the ALTO protocol definition.

5.4. User Privacy

On the one hand, there are data elements an ALTO client could provide in its query to an ALTO server that 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. On the other hand,

applications might consider such information private. In addition, some applications may not know a priori what kind of request they will be making.

5.5. Topology Hiding

Operators, with their intimate knowledge of their network topology, can play an important role in addressing the ALTO problem. However, operators often consider revealing details of such network information to be confidential.

5.6. Coexistence with Caching

Caching is an approach to improving traffic generated by applications, and it requires large amounts of data transfers. In some cases, such techniques have proven to be extremely effective in both enhancing user experience and saving network resources.

A cache, either explicitly or transparently, replaces the content source. Thus, a cache must, in principle, use and support the same protocol as the querying peer. That is, if a cache stores web content, it must present an HTTP interface to the web client. Any cache solution for a given protocol needs to present that same protocol to the client. Said differently, each caching solution for a different protocol needs to implement that specific protocol. For this reason, one can only reasonably expect caching solutions for the most popular protocols, such as HTTP and BitTorrent.

It is extremely important to realize that caching and ALTO are entirely orthogonal. ALTO, especially if it is aware of caches, can in fact direct clients to nearby caches where the user could get a much better quality of experience.

6. Security Considerations

This document is neither a requirements document nor a protocol specification. However, we believe it is important for the reader to understand areas of security and privacy that will be important for the design and implementation of an ALTO solution. Moreover, issues such as digital rights management are out of scope for ALTO, as they are not technically enforceable at this level.

Some environments and use cases of ALTO may require client or server authentication before providing sensitive information. In order to support those environments interoperably, the ALTO requirements document [ALTO-REQS] outlines minimum-to-implement authentication and other security requirements.

Applications can decide to rely on information provided by an ALTO server to enhance the peer-selection process. In principle, this enables the ALTO service that provides such information to influence the behavior of the application, basically letting a third-party -- the ALTO service provider -- take an important role in a distributed system it was not previously involved in.

For example, in the case of an ALTO server deployed and run by an ISP, the P2P community might consider such a server hostile because the operator could:

- use ALTO to prevent content distribution and enforce copyrights;
- redirect applications to corrupted mediators providing malicious content;
- track connections to perform content inspection or logging;
- apply policies based on criteria other than network efficiency. For example, the service provider may suggest routes suboptimal from the user's perspective in order to avoid peering points regulated by inconvenient economic agreements.

It is important to note there is no protocol mechanism to require ALTO for P2P applications. If, for some reason, ALTO fails to improve the performance of P2P applications, ALTO will not gain popularity and the P2P community will not use it.

At the time of this writing, the privacy issues described in Section 5.4 are relevant for an ALTO solution. Users may be reluctant to disclose sensitive information to an ALTO server. Operators, on the other hand, may not wish to disclose information that would expose details of their interior topology. When exploring the solution space in detail, one needs to consider these issues so that an ALTO protocol does not presume mandatory information disclosure, by either clients or servers.

7. Contributors

This document was initially edited by Enrico Marocco and Vijay Gurbani. In the role of Working Group chairs, they have continued to provide significant edits and inputs to the current authors.

8. Acknowledgments

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, Greg DePriest, and many others provided insightful discussions, specific comments, and much needed corrections.

Jan Seedorf and Sebastian Kiesel are partially supported by the NAPA-WINE project (Network-Aware P2P-TV Application over Wise Networks, http://www.napa-wine.org), a research project supported by the European Commission under its 7th Framework Program (contract no. 214412). The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of the NAPA-WINE project or the European Commission.

Thanks in particular to Richard Yang for several reviews.

9. Informative References

[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?", ACM USENIX IMC, Berkeley 2005.

[ACM.ispp2p]
Aggarwal, V., Feldmann, A., and C. Scheideler, "Can ISPs and P2P systems co-operate for improved performance?", 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.

[ALTO-REQS]
Kiesel, S., Popkin, L., Previdi, S., Woundy, R., and Y. Yang, "Application-Layer Traffic Optimization (ALTO) Requirements", Work in Progress, April 2009.

[PATH-SEL]
Saucez, D. and B. Donnet, "The case for an informed path selection service", Work in Progress, February 2008.

[RFC3031] Rosen, E., Viswanathan, A., and R. Callon, "Multiprotocol Label Switching Architecture", RFC 3031, January 2001.

[RFC3260] Grossman, D., "New Terminology and Clarifications for Diffserv", RFC 3260, April 2002.

[WWW.p4p.overview] Xie, H., Krishnamurthy, A., Silberschatz, A., and R. Yang, "P4P: Explicit Communications for Cooperative Control Between P2P and Network Providers", <http://www.dcia.info/documents/P4P_Overview.pdf>.

[WWW.wired.fuel] Glasner, J., "P2P Fuels Global Bandwidth Binge", April 2005, <http://www.wired.com>.

Authors' Addresses

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
Neustar Inc.
46000 Center Oak Plaza
Sterling, VA 20166-6579
USA

Phone:
Fax: +1 530 267 7447
EMail: eburger@standardstrack.com
URI: http://www.standardstrack.com