

Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: April 26, 2009

S. Das
V. Narayanan
L. Dondeti
Qualcomm, Inc.
October 23, 2008

ALTO: A Multi Dimensional Peer Selection Problem
draft-saumitra-alto-multi-ps-00

Status of this Memo

By submitting this Internet-Draft, each author represents that any applicable patent or other IPR claims of which he or she is aware have been or will be disclosed, and any of which he or she becomes aware will be disclosed, in accordance with [Section 6 of 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/1id-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 April 26, 2009.

Abstract

Bulk data applications are posing serious challenges to the Internet infrastructure and more mass adoption of such applications would only increase that challenge. P2P bulk data applications and other large volume HTTP traffic such as video currently dominate the Internet traffic. These applications do not generally benefit from the traffic engineering techniques developed for the Internet. In the HTTP traffic case, the traffic optimization issues are often addressed using the CDN infrastructure. For P2P bulk data applications, the problem is about picking the right peers to communicate with and the criteria for doing this vary greatly based on the application. Hence, intelligent peer selection is the

fundamental problem to address for these applications. This document explains the multiple dimensions of the peer selection problem with respect to obtaining information from the network as well from other peers and argues for an integrated, common framework to share such information.

Table of Contents

1.	Introduction	3
2.	Terminology	5
3.	Problem Statement	5
3.1.	Network-assisted ALTO	6
3.2.	Peer-assisted ALTO	6
3.3.	The need for multiple criterion	6
3.3.1.	Peer-assisted ALTO alone	7
3.3.2.	Network-assisted ALTO alone	7
3.4.	Applicability and Discussion	8
4.	Design Goals	9
5.	Security Considerations	11
6.	IANA Considerations	11
7.	Acknowledgments	11
8.	References	12
8.1.	Normative References	12
8.2.	Informative References	12
	Authors' Addresses	13
	Intellectual Property and Copyright Statements	14

1. Introduction

A large portion of the Internet traffic today is from P2P (peer-to-peer) applications. Such an architecture for data transfer is attractive because it reduces the bandwidth costs of the content provider since they simply need to seed the content to a few nodes in the network which would then contribute upload bandwidth to assist the content provider's servers in the data transfer. Thus, the single point bottleneck at the content provider is eliminated and both users and content providers benefit.

However such an approach is detrimental to another important entity in the system: the ISPs. While p2p has led to increased popularity of broadband connections and arguably increased subscribers for ISPs; it has also increased traffic costs for the ISPs. This is because P2P applications' peer selection does not consider underlying network topology and link costs in that topology. Most p2p applications typically are only interested in improving their data transfer performance which is satisfied if the download link of the user is saturated. As shown in [2], traffic generated by popular P2P applications often cross network boundaries multiple times, overloading links which are frequently subject to congestion [3]. This happens because most p2p applications simply use random peer selection followed by monitoring and reevaluation. Even if p2p applications perform some network measurements, these typically tend to be round trip time estimation which may or may not lead to peer selection conducive to ISP interests.

This led to ISPs efforts at shaping or blocking P2P traffic on specific ports. In response, p2p applications started using dynamic ports to transfer data following which ISPs had to use deep packet protocol specific information to shape p2p flows. In response, p2p applications have started encrypting their connections. The use of TCP RST messages to deter costly p2p application data transfer has led to some conflicts as well. It is in the ISP's interests to avoid the cat-and-mouse games of protocol-specific detection and mitigation while still not having to increase costs significantly to accommodate p2p traffic.

One way to reduce the impact on ISPs would be the deployment of caching entities in the networks to reduce cross-ISP traffic and network distance of data transfer. However such an approach has several issues:

- o It is not clear who would deploy these high-bandwidth large-storage capable caches since it can be argued that caching pushes the costs from the content provider to the ISPs

- o The caches would need to be able to cache data from all p2p applications and consequently become complicated to deploy and maintain over time as p2p application evolve
- o The use of caches opens up the issue of storage of copyrighted content

Thus, a solution is needed that can allow for peer selection by the p2p application themselves that is friendly to the ISP's network costs while being friendly to the applications' objective of good performance for the user.

Recent studies [4], [5], [6] have suggested that it may be possible to reduce the impact that P2P applications have on ISPs traffic costs. This is mainly achieved by informed peer selection in the P2P applications guided by network level metrics instead of random selection. However, p2p applications do not have a trust relationship with network operators and what may be good for the ISP is not necessarily good for the performance of the p2p applications. These competing interests necessitate a solution for ALTO that addresses the interests of both the entities in the system.

This document describes the problem of optimizing traffic generated by P2P applications using information provided by third parties, i.e. the other peers in the network or the network operator. The overall goal of optimizing the P2P application is for them to become more network-friendly, while at the same time allowing the networks to remain application-friendly. The following are key arguments that we put forth in this draft:

- o The problem of peer selection needs to take into account the interests of the ISPs, application providers and the peers. The goal should be to reduce the impact on ISP without sacrificing the performance of p2p applications. There are many situations we elaborate upon in [Section 3](#) where simply considering ISP interests leads to poor performance for p2p applications. Information about peer connectivity characteristics is an important component of the ALTO problem of peer selection (in conjunction with ISP routing information) and together with network topology information, can enable peers to optimize their performance as well as take into account ISP interests.
- o An ALTO system with information about ISP routing alone does not provide enough incentives for adoption in p2p applications, due to the competing interests of the two parties and the lack of trust. Combining both elements of peer selection creates an incentive for p2p applications to actually use the ALTO service.

- o Application-specific solutions for sharing peer connectivity characteristics are inefficient and cause unnecessary overhead in the network. A common information plane allows reuse of such information across a wide range of applications ranging from DHT next hop selection, multi-source file download, relay selection, mirror selection, etc., since many of the connectivity characteristics of peers such as upload/download bandwidth, network coordinate, wireless link information are useful to a multitude of p2p applications in peer selection.

The rest of this draft is structured as follows. [Section 3](#) introduces the problem and argues for the multiple criterion that need to be considered when developing solutions, while [Section 4](#) describes the design goals of the system.

2. Terminology

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](#)].

3. Problem Statement

Bulk data applications are posing serious challenges to the Internet infrastructure and more mass adoption of such applications would only increase that challenge. P2P bulk data applications and other large volume HTTP traffic such as video currently dominate the Internet traffic. These applications do not generally benefit from the traffic engineering techniques developed for the Internet. In the HTTP traffic case, the traffic optimization issues are often addressed using the CDN infrastructure. For P2P bulk data applications, the problem is about picking the right peers to communicate with and the criteria for doing this vary greatly based on the application. Hence, intelligent peer selection is the fundamental problem to address for these applications.

One necessary input for intelligent peer selection has been shown to be network topology information. And, accurate network topology information can only be obtained with the help of ISPs. However, there is more to peer selection than just network topology information.

Consider a Client Application at node A which wants to access a service for peer selection such as ALTO. Overall peer selection in ALTO can be performed using two sets of information: (1) One set of information is that from the network, i.e. network-assisted ALTO

which promotes peer selection that is beneficial to the ISP by either reducing the interdomain traffic or reducing congestion on bottleneck links inside the ISP. (2) The other set of information is from the set of potential Client Application Providers are targets for peer connectivity, i.e. peer-assisted ALTO.

3.1. Network-assisted ALTO

Network-assisted application layer traffic optimization refers to the use of network operators in guiding peer selection for p2p applications. Network operators can use network topology and traffic statistics to provide hints to the P2P applications on which hosts it would prefer the application pick for data transfer. For example, peers that involve inter-domain routing would be given lower priority in the response to the querying P2P application. Another example would be that peers whose connectivity is such that their choice would increase the congestion on an already congested link inside the ISP. Such peers would also be given lower priority. Some recent work has proposed solutions in this space [5],[4].

3.2. Peer-assisted ALTO

Peer-assisted application layer traffic optimization refers to the use of published information about peers and their connectivity to the Internet in guiding peer selection for p2p applications. For example, BitTorrent-like data transfer applications would query metrics related to the peer of interest to decide on the peers to select for initial data transfer. This information could be for example the upload bandwidth of the peer. These choices could then be refined or changed by monitoring data connections. Another example would be for VoIP applications to query the ALTO service for peers that it wants to use as relays to find one with low latency. Although latency information is pairwise the peer could publish its network coordinate calculated by a system such as GNP (Global Network Positioning) into the ALTO service which could then be retrieved and used by a querying peer to estimate network latency. Publishing peer connectivity could also involve edge-specific information such as which cell id a peer is connected to in a cellular network so that another peer in the same cell is discouraged from making a connection within the same cell to minimize the congestion and avoid occupying 2 downlink and 2 uplink channels in the same cell. This type of traffic optimization cannot be achieved via network-assisted ALTO.

3.3. The need for multiple criterion

This section justifies the need for multiple criterion in ALTO: i.e. the need for both ISP related information and peer related information. Note that the underlying assumption is that peers and

ISPs may not necessarily trust each other and thus any solution in this space needs to consider the interests of both parties.

3.3.1. Peer-assisted ALTO alone

It is already well known that using peer related information alone is not enough. Simply randomly selecting peers and trying them out and reevaluating the choices based on observed performance is suboptimal as it takes time to converge on a good set of peers as well as causes unnecessary overhead in the network. Incorporating locality into peer selection using either active measurements or some sort of information plane (e.g. [7]) has been shown to improve the download completion time performance of BitTorrent. For example, [7] showed a 35% reduction in download time for 60% of the peers in a swarm. Other studies [5] showed this gain to be around 10-25%. However, [5] also showed that such an approach increases the traffic on congestion on bottleneck links by 69%. In summary, while peer related information is powerful in improving performance, it affects ISP performance and can lead to shaping and throttling.

3.3.2. Network-assisted ALTO alone

On the other hand there are several situations where relying on ISP information by itself is not sufficient. If ISPs simply return a rank ordered list of node identifiers in response to a list of node identifiers sent in a query, the querying node may not be optimizing its performance. It is possible that the peers in a swarm that are within the ISP are all DSL hosts which can upload at 128Kbps while some of the other peers that may involve interdomain routing are actually high bandwidth fiber or cable connected hosts with high upload bandwidth. In such cases, the p2p application user would sustain high download times to meet the ISP's objective. Similarly, peers that are nearby geographically (which is correlated with latency typically) may be preferred for relays in VoIP but do not meet the ISP's objective since they use a different access network.

The reasons for these problems are two-fold: (1) The ISPs' objectives in general may not always match the users objective of improved performance, and (2) The ISP can only provide an indication of its preference for connectivity to certain peers outside its domain but it does not have any notion of end-to-end performance provided by these outside peers.

Thus neither network-assisted ALTO or peer-assisted ALTO alone help in bringing together these competing interests together. Thus, in this draft we argue for a combined solution that provides a true information plane for peer selection, i.e. one that does not sacrifice but, in fact improves the performance of p2p applications

while getting cooperation from p2p applications for the ISP. ISP information is only one of the two pillars of the ALTO service - it is an important one, given that the ISPs have knowledge of internal congestion, interdomain peering policies and connectivity information. Peer information is the other pillar of the ALTO service; and by providing it as a service to all p2p applications we reduce the overhead and inefficiency of each application performing their own measurements which is inefficient for the network and complicated for application developers.

3.4. Applicability and Discussion

The problem of peer selection has been studied at various dimensions and has pretty broad applicability. At the very basic level, intelligent peer selection can be applied, in file sharing networks, to the problem of choosing the right source to download a file from. However, there are a variety of other applications for such intelligent peer selection. For example, locality aware structured overlays are built with topology-assisted neighbor selection models or topology-based identity generation schemes. The more accurate the topology information is, the more efficient such schemes will be. Very fundamentally, intelligent neighbor selection schemes aim at improving the query latency in such systems. Peer selection is also applicable to other problems such as identifying the best super peer to contact in a system, using the best relay for NAT traversal, identifying the best next hop for a query based on several criteria (e.g., quality, reputation, semantic expertise, etc.), etc.

A lot of study has been done on the applicability of the peer selection problem. Some examples include [\[8\]](#), [\[9\]](#), [\[10\]](#), [\[11\]](#), [\[12\]](#), [\[13\]](#).

Two observations can be made from these applicability aspects of peer selection. One is that the problem is applicable to a wide range of scenarios with some possible generalization - in all cases, while the criteria for peer selection and the context in which it is done may be different, the basic model for peer selection revolves around ranking a list of peers based on information collected about the peers in accordance with the given criteria. A second observation is that the same piece of information collected could be put to use in very different ways in different scenarios - e.g., the use of topology information in determining source peers to download content from and neighbor peers to make connections with are very different applications.

Peer selection is also equally important for application providers, ISPs and the peers themselves for very different reasons. ISPs have an incentive to keep their operational costs to a minimum, while

ensuring their customers a good experience. Application providers are interested in keeping their own operational costs to a minimum, which by nature, conflicts with the ISP's interests of keeping the costs low. Application providers are also interested in ensuring that the application performs well to keep the peers interested in the application. For the peers themselves, the emphasis is on performance as well as their own access costs.

The varied scenarios of applicability of peer selection information, the diverse interests of various parties in peer selection and the underlying common nature of these models suggest that the problem should be tackled in a uniform manner, allowing for generality and extensibility of information. The scenarios also broadly fall under two buckets - one, where information may be obtained from specific hosts (e.g., ISP hosts, reputation tracking servers, etc.) and another, where information may be obtained from any number of peers. The final peer selection algorithm may be a combination of various pieces of such information in a manner that fits a given scenario and criterion. Solving any one piece of this puzzle separately is likely to lead to multiple different protocols and mechanisms for the same problem. Further, any one such protocol is unlikely to result in a dramatic improvement of the scenario (be it bandwidth utilization or latency reduction, etc.). For instance, as argued earlier in this draft, network-assistance and peer information are both important to consider for peer selection. Hence, the adoption of any one of these protocols will be less incentivized in such a fragmented model.

The ALTO service, given a set of host location identifiers for peers, may annotate them with ISP information (such as the p-distance from P4P) as well as peer connectivity information (which is variable and could be related to bandwidth, network coordinates, wireless link information etc) and return the annotated set to the p2p application. The p2p application can use this data to perform peer selection based on its requirements. We argue that the peer connectivity information be a generic container that can contain different types of information. However the request response protocol can be unified as an ALTO protocol for retrieving the annotated information for each destination peer. Note that the data pertaining peer-assisted ALTO and network-assisted ALTO may be stored differently but the interface to p2p applications can be kept consistent.

4. Design Goals

The fundamental goal of ALTO, extracting from what has been described thus far, has to do with allowing the use of both network-assisted and peer-assisted information exchanges for various categories of

applications that can use it. The main idea is to model ALTO as a service that these applications may use to apply the peer selection intelligence in the specific context of interest to the application. The following are design goals for the ALTO service framework:

1. The ALTO service should accommodate specific hosts providing network topology information.

Network topology information may be provided by hosts in an ISP network with much greater accuracy than can be done via other schemes. Hence, a model that allows for such information exchange from specific hosts is useful. The discovery of such hosts capable of providing this information falls in this territory as well.

2. The ALTO service should allow peers to publish ALTO related information in an application agnostic manner.

Peers may publish information of various kinds that helps in peer selection. A primary example of such information is the connectivity characteristics of a peer. This type of information allows other criteria to be taken into account for peer selection, such as the uplink/downlink bandwidth of peers, the wireless nature of links, etc. This information, coupled with topology or proximity information, will allow exchange of bulk data in a manner that benefits the ISPs and the peers.

3. The ALTO service should allow for both centralized and distributed service models.

ISP hosts and any servers that are queried for ALTO related information may be offering the ALTO service in a centralized model. On the other hand, peer information related to ALTO may be obtained in a distributed fashion, e.g., by querying an overlay. It is important to cater to both models and also allow for the two models to be linked. For instance, a query on an overlay for ALTO related information may result in a different node on the overlay querying the ISP's ALTO host.

4. The ALTO protocols should be agnostic to the actual peer selection algorithm in use.

The peer selection algorithm itself is contextual and depends on different factors for different scenarios. For instance, source selection for bulk data downloads involves optimal bandwidth usage and performance as criteria, neighbor selection in structured overlays may be a function of hop

count and delay, relay selection for NAT traversal may be a function of the relay capacity and connectivity characteristics and so on. Hence, the ALTO protocols should be agnostic to the specific peer selection algorithm for any given instantiation.

5. The ALTO protocols should be generic to allow any type of information useful for performing ALTO services to be exchanged. The ALTO protocols should be extensible to allow carrying new types of information that may be defined at a future point.

As in the case of the algorithm, the types of information exchanged for peer selection are also contextual to various scenarios. For instance, bandwidth or latency based peer selection may involve topology information and connectivity characteristics of peers, relevance based peer selection may involve the semantic expertise of various peers, and so on. Hence, the ALTO protocols should simply be a container to transport information that assists in peer selection without being dependent on the actual type of information exchanged.

5. Security Considerations

Information exchange, as facilitated by ALTO, may be sensitive and may require secure communication. Hence, the ALTO protocols must provide for channel security properties such as confidentiality and integrity protection. Further, the exchange of some of this information may need to be limited to certain entities. This calls for authentication and authorization of entities prior to exchange of ALTO information. Hence, the ALTO framework should provide for entity authentication and authorization as part of the service. When various types of information to be carried in the ALTO protocols are standardized, a call must be made about whether it is subject to the authorization model. Further, some types of information exchanges could lead to privacy issues to various parties and the implications of that need to be studied prior to standardization.

6. IANA Considerations

This document does not have any IANA considerations.

7. Acknowledgments

This draft has benefited from the insights presented in several IETF drafts namely the ALTO problem statement, requirements and survey

drafts. The research work done on IPlane, P4P, Taming the torrent, and ISP and P2P oracles as well as the various measurement studies on P2P applications impact on ISPs performed by the research community have helped inform us in creating this draft.

8. References

8.1. Normative References

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

8.2. Informative References

- [2] Karagiannis, T., Rodriguez, P., and K. Papagiannaki, "Should ISPs fear Peer-Assisted Content Distribution?", IMC Proceedings of the Internet Measurement Conference, 2005.
- [3] Akella, A., Seshan, S., and A. Shaikh, "An Empirical Evaluation of WideArea Internet Bottlenecks", Proceedings of ACM SIGCOMM, 2003.
- [4] 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, 2006.
- [5] Xie, H., Krishnamurthy, A., Silberschatz, A., and R. Yang, "P4P: Explicit Communications for Cooperative Control Between P2P and Network Providers", Proceedings of ACM SIGCOMM, 2008.
- [6] Choffnes, D. and F. Bustamante, "Taming the Torrent: A practical approach to reducing cross-ISP traffic in P2P systems", Proceedings of ACM SIGCOMM, 2008.
- [7] Madhyastha, H., Isdal, T., Piatek, M., Dixon, C., Anderson, T., Krishnamurthy, A., and A. Venkataramani, "iPlane: an information plane for distributed services", Proceedings of USENIX OSDI, 2006.
- [8] Lo, V., Liu, Y., GauthierDickey, C., and J. Li, "Scalable Leader Selection in Peer-to-Peer Overlay Networks", Proceedings of Second International Workshop on Hot Topics in Peer-to-Peer Systems (Hot-P2P), 2005.
- [9] Xhafa, F., Barolli, L., Fernandez, R., and T. Daradoumis, "An Experimental Study on Peer Selection in a P2P Network over PlanetLab", Proceedings of International Conference on

Parallel Processing Workshops, 2007.

- [10] Yang, X. and G. Veciana, "Service Capacity of Peer to Peer Networks", Proceedings of IEEE INFOCOM, 2004.
- [11] Habib, A. and J. Chuang, "Service Differentiated Peer Selection: An Incentive Mechanism for Peer-to-Peer Media Streaming", Proceedings of IWQoS, 2004.
- [12] Tang, S., Wang, H., and P. Meigham, "The Effect of Peer Selection with Hopcount or Delay Constraint on Peer-to-Peer Networking", Proceedings of IFIP NETWORKING, 2008.
- [13] Haas, P. and R. Siebes, "Peer Selection in PeertoPeer Networks with Semantic Topologies", Proceedings of WWW, 2004.

Authors' Addresses

Saumitra M. Das
Qualcomm, Inc.
3195 Kifer Road
Santa Clara, CA
USA

Phone: +1 408-533-9529
Email: saumitra@qualcomm.com

Vidya Narayanan
Qualcomm, Inc.
5775 Morehouse Dr
San Diego, CA
USA

Phone: +1 858-845-2483
Email: vidyan@qualcomm.com

Lakshminath Dondeti
Qualcomm, Inc.
5775 Morehouse Dr
San Diego, CA
USA

Phone: +1 858-845-1267
Email: ldondeti@qualcomm.com

Full Copyright Statement

Copyright (C) The IETF Trust (2008).

This document is subject to the rights, licenses and restrictions contained in [BCP 78](#), and except as set forth therein, the authors retain all their rights.

This document and the information contained herein are provided on an "AS IS" basis and THE CONTRIBUTOR, THE ORGANIZATION HE/SHE REPRESENTS OR IS SPONSORED BY (IF ANY), THE INTERNET SOCIETY, THE IETF TRUST AND THE INTERNET ENGINEERING TASK FORCE DISCLAIM ALL WARRANTIES, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO ANY WARRANTY THAT THE USE OF THE INFORMATION HEREIN WILL NOT INFRINGE ANY RIGHTS OR ANY IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

Intellectual Property

The IETF takes no position regarding the validity or scope of any Intellectual Property Rights or other rights that might be claimed to pertain to the implementation or use of the technology described in this document or the extent to which any license under such rights might or might not be available; nor does it represent that it has made any independent effort to identify any such rights. Information on the procedures with respect to rights in RFC documents can be found in [BCP 78](#) and [BCP 79](#).

Copies of IPR disclosures made to the IETF Secretariat and any assurances of licenses to be made available, or the result of an attempt made to obtain a general license or permission for the use of such proprietary rights by implementers or users of this specification can be obtained from the IETF on-line IPR repository at <http://www.ietf.org/ipr>.

The IETF invites any interested party to bring to its attention any copyrights, patents or patent applications, or other proprietary rights that may cover technology that may be required to implement this standard. Please address the information to the IETF at ietf-ipr@ietf.org.

