ALTO Server Load Reduction
draft-stiemerling-alto-load-reduction-00

Abstract

   Traffic localization for peer-to-peer applications is currently
   discussed by the IETF ALTO working group. The protocol specification
   itself is discussed and some issues about load reduction for ALTO/P4P
   servers are also discussed, but mainly on the transport level (i.e.,
   caching of traffic localization data) and some data aggregation by
   using macros that comprise a number of IP prefixes. However, there
   are no considerations about reducing the query load to the ALTO
   server. This memo aims at giving informational guidance to P2P
   applications implementers on how reduce the query load for ALTO.

Status of this Memo

   This Internet-Draft is submitted to IETF in full conformance with the
   provisions of BCP 78 and BCP 79.

   Internet-Drafts are working documents of the Internet Engineering
   Task Force (IETF), its areas, and its working groups. Note that
   other groups may also distribute working documents as Internet-
   Drafts.

   Internet-Drafts are draft documents valid for a maximum of six months
   and may be updated, replaced, or obsoleted by other documents at any
   time. It is inappropriate to use Internet-Drafts as reference
   material or to cite them other than as "work in progress."

   The list of current Internet-Drafts can be accessed at
   http://www.ietf.org/ietf/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 September 2, 2010.

Copyright Notice

Table of Contents

[1](#). Introduction

   Peers in current p2p systems are keeping track of their neighbor
   peers (neighbors in the overlay) in a peer list. This peer list
   typically divided in active peers (they deliver data/chunks),
   candidate peers (peers that have been recently learned by the peer)
   and dead peers (peers that have performed bad in the last epoch and
   will be discarded at some point). The currently deployed peer-to-
   peer systems work on time epochs, i.e., they exchange data with other
   peers for a fixed time period, and after this period the peers
   evaluate the usefulness of each peer in terms of throughput and if
   the peer did have the requested chunk. A typical value for an epoch
   is 30 seconds.

   The current proposal in traffic localization (P4P/ALTO) assumes that
   peers do always send the complete candidate set or the active set to
   the server, asking for traffic localization guidance on that set.
   However, the active and the candidate sets are changing frequently.
   Every peer needs to query the ALTO server quite frequently with the
   updated candidate or active set.

   It is foreseeable that ALTO servers will need to be able to handle a
   query load that is proportional to the number of peers (ALTO
   clients), the query rate of the peers, and the amount of peers
   included in each query. P2P systems maintain an active and a
   candidate peers set that should be evaluated by the ALTO server, so
   that it can give guidance about what peers are to be preferred. This
   would require querying the ALTO server every epoch even though it is
   not clear whether the candidate peers can actually deliver the
   requested content at a desired throughput. Second, new peers can
   also be learned on a peer base, i.e., not in a bulk via a resource
   directory. In this case, the peer would query for single peers at
   the ALTO server.

   The steps are:

   1. The peer obtains the set of new peers and adds them to its
       candidate set (either via a resource directory (tracker) or via a
       peer exchange protocol);

   2. The peer queries the ALTO server with the candidate set;

   3. The peer takes peers preferred by the ALTO server out of its
       candidate sets and starts data exchange with them;

   4. The peer moves a candidate peer to the active set, if the peers
       has the data of interest and if the peer delivers sufficient
       throughput (typically above a certain threshold);

5. The peer moves a candidate peer to the dead set of drops
    immediately if the data of interest is not available or if the
    throughput is below a certain threshold.

Imagine now that every peer queries for a complete candidate set
(100s of peers) every 30 seconds and that there are thousands of
peers within a single ISP. This will cause a high load on the ALTO
servers that need to handle the requests of all ALTO clients.
Furthermore, the peer will query for candidate peers although they
might not have the data of interest or they might not able to deliver
the desired throughput, causing unnecessary queries or entries in
queries.

We call the set of peers send to the ALTO server the query set, as it
may contain other peers as the actual candidate set, i.e., more or
less entries and completely other entries.

Comments and discussions about this memo should be directed to the
ALTO working group: alto@ietf.org.

[2](#). **Guidance on Load Reduction**

   The known candidate set is divided into two sets called Pending 1 and
   Pending 2. The candidate set in regular p2p systems is comparable
   with the Pending 1 set, as it takes a full list of candidate peers
   that have been rated (i.e., sorted by costs and expected performance)
   by the ALTO server in the ALTO query 1. This track in Figure 1 XXX
   is mainly performed for candidate peers that have been obtained from
   a resource directory (e.g. a tracker or a DHT). Peers in the Pending
   1 set are evaluated in their order of ALTO rating and good peers are
   transferred to the active peers sets. Peers that do not perform well
   in the Pending 1 set are transferred to the dead peers set. However,
   this track is as in today's p2p systems.

   P2P systems use also a peer exchange (via gossiping) to learn other
   new candidate peers that can be added to the candidate set. Each of
   the new peer would have to be rated by the ALTO server for it costs
   and expected performance. To avoid querying the ALTO server too
   much, we introduce the track 2 that is depicted on the right side of
   Figure 1 XXX. Once the peer has learned a new candidate peer, the
   peer evaluates if the new peer is required right now (e.g. throughput
   is insufficient). If not, the candidate peer is collected in a
   buffer (top of figure). The buffer content is pushed to track 1 is
   some candidate peers have been accumulated.

   If a new candidate peer is required right now, the peer will first
   test the connectivity to the candidate peer, without asking for the
   ALTO server rating. The peer is transferred to the dead peer set if
   it performs badly or if it is unreachable. If the candidate peer
   performs well it is queried at the ALTO server. Depending on the
   ALTO rating it is either transferred into the active peer set (cheap
   candidate peer) or it is transferred to the dead peers set if it is
   expensive and if there is no urgent need to get new peers. The
   urgent need for peers can be, for instance, if the throughput of all
   peers in the active set drops below a threshold. In this case, the
   candidate peer is anyhow transferred to the active peers set despite
   the ALTO rating of being expensive. This will avoid that the peer is
   disconnected from the p2p network and that the ISP network will get
   fragmentation, i.e., disconnected from the rest of the p2p network.

   An extension of the above proposed track 2 is shown in Figure 2 XXX.
   In this all candidate peers are subject to track 2. This will lower
   the query load on the server, as first all candidate peers are
   subject to the connectivity test in Pending 2 set. However, this can
   cause a higher traffic on the peering links of the actual ISP, as all
   peers are first tested (this includes exchanging data) and only the
   good peers.

3. Security Considerations

   This memo does not introduce any new security considerations, other
   than the ones stated in [I-D.ietf-alto-protocol]

[4](#). **Conclusion**

   This memo aims at giving guidance to P2P application designers and
   implementors to reduce the query load to ALTO servers. The proposed
   mechanism does not need any modification of the ALTO server and can
   be easily implemented in the ALTO client or in the P2P application.

   The figures for [Section 2](#) are still missing.

[5](#). References

[5.1](#). Normative References

   [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate
              Requirement Levels", [BCP 14](#), [RFC 2119](#), March 1997.

[5.2](#). Informative References

   [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.

   [I-D.akonjang-alto-proxidor]
              Akonjang, O., Feldmann, A., Previdi, S., Davie, B., and D.
              Saucez, "The PROXIDOR Service",
              [draft-akonjang-alto-proxidor-00](#) (work in progress),
              March 2009.

   [I-D.ietf-alto-protocol]
              Alimi, R., Penno, R., and Y. Yang, "ALTO Protocol",
              [draft-ietf-alto-protocol-01](#) (work in progress),
              December 2009.

   [I-D.kiesel-alto-3pdisc]
              Kiesel, S. and M. Tomsu, "Third-party ALTO server
              discovery", [draft-kiesel-alto-3pdisc-01](#) (work in
              progress), October 2009.

   [I-D.kiesel-alto-h12]
              Kiesel, S. and M. Stiemerling, "ALTO H12",
              [draft-kiesel-alto-h12-01](#) (work in progress), March 2010.

   [I-D.kiesel-alto-reqs]
              Kiesel, S., Popkin, L., Previdi, S., Woundy, R., and Y.
              Yang, "Application-Layer Traffic Optimization (ALTO)
              Requirements", [draft-kiesel-alto-reqs-02](#) (work in
              progress), March 2009.

   [I-D.penno-alto-protocol]
              Penno, R. and Y. Yang, "ALTO Protocol",
              [draft-penno-alto-protocol-04](#) (work in progress),
              October 2009.

   [I-D.shalunov-alto-infoexport]
              Shalunov, S., Penno, R., and R. Woundy, "ALTO Information

                 Export Service", draft-shalunov-alto-infoexport-00 (work
                 in progress), October 2008.

   [I-D.stiemerling-alto-h1h2-protocol]
                 Stiemerling, M. and S. Kiesel, "ALTO H1/H2 Protocol",
                 draft-stiemerling-alto-h1h2-protocol-00 (work in
                 progress), March 2009.

   [RFC5693]   Seedorf, J. and E. Burger, "Application-Layer Traffic
                 Optimization (ALTO) Problem Statement", RFC 5693,
                 October 2009.

Authors' Addresses

    Martin Stiemerling
    NEC Laboratories Europe/University of Goettingen
    Kurfuerstenanlage 36
    Heidelberg 69115
    Germany

    Phone: +49 6221 4342 113
    Fax: +49 6221 4342 155
    Email: martin.stiemerling@neclab.eu
    URI: http://www.nw.neclab.eu/

    Sebastian Kiesel

    Email: ietf-alto@skiesel.de