УДК 004.75, 004.724.2

к.т.н., с.н.с. Порєв Г.В.

Національний технічний університет України «КПІ»

LOCALITY METRIC IN DISTRIBUTED PEER-TO-PEER NETWORKS

Abstract

Gennadiy V. Poryev, research fellow, PhD
Locality metric in distributed peer-to-peer networks

This paper concerns with proposed traffic-less and provider-independent locality metric in the distributed peer-to-peer networks. The metric is based on the locally pre-computed topological model of the national Internet segments and locality classes inferred from it. Such metric can be used for various purposes, including automatic node domain clustering, selective QOS, spanning tree construction etc.

Introduction

Since the beginning of 21st century the usage, scale and diversity of peer-to-peer (P2P) networks widened significantly, and the application scope of P2P systems has been notably extended.

The initial driving motivation behind P2P systems was to relieve loading stress form centralized server farms. However, the intrinsic asymmetry of end-user broadband data links has caused Internet Service Providers (ISP) to increase maintenance and upgrade cost of the "last mile" hardware in order to keep quality of service steady.

For this reason, researches in the area of P2P systems are aiming to optimize P2P traffic and consider the inherently clustered nature of the Internet as a potential leverage mechanism. The general idea is to maximize network throughput inside the particular network clusters while minimizing the traffic usage between such clusters. The scope of the cluster is not defined clearly, and there is usually more than one clustering layer.

Locality problem and its existing solutions

Usenet was introduced 30 years ago as one of the first P2P networks. However, only at the end of the 1990s P2P applications have achieved a breakthrough and become very popular because of the widespread use of file sharing platforms like Napster. Nowadays, there is a wide variety of P2P file sharing networks. Among them are Gnutella, eDonkey2000 (ED2K) and BitTorrent [1]. Various surveys suppose that 30% to 50% of today's end-user-generated traffic is caused by P2P applications. In [2] the authors claim that most P2P systems use application level routing based on the overlay topology and completely neglect the topology of the underlying transport network.

The authors of [3] have recently described the design, deployment and evaluation of an approach minimizing the expensive cross-ISP traffic. The authors show that the application of their approach significantly reduces the latency delays.

We believe that the consideration of a node's topological locality is the key to efficient communication in P2P systems. It improves performance and increases availability, since the probability of transmission failures increases with the distance and depends also on bandwidth conditions.

Modern network modeling environments that deal with network topology rarely take locality into account. We deem the ping and route tracing method as generally unreliable as it heavily depends on link speeds and bandwidth conditions. As shown in [5] standard routing trace methods may also be unreliable and affected by bandwidth conditions or indicating non-existent links due to traffic switch-overs.

A number of researches have proposed schemes that involve building the external (in relation to the P2P overlay) infrastructures dedicated to keep track of conditions of intra-network and inter-network links, remembering explicitly measured routing paths and delivering path prediction on their basis. Such schemes, for instance, include P4P [4] and iPlane. Other proposals are concerned with an active intervention into the P2P exchange protocols to augment traffic usage patterns in accordance with ISP policies.

Locality metric

In network based applications we often need global knowledge of all network nodes and distances between these nodes. By having this knowledge, nodes in a network are able to construct complex infrastructures and achieve efficient communication at the application as well as at the network layer.

Contrary to the methods employing external infrastructures for distance estimation, we propose an approach based on the preloaded and pre-computed topological structure of the Internet and running locally on client machines.

The core of our proposal is the combined metric (LM) which is calculated locally on each node independently of the others and is only meaningful within the scope of this node. Metric is calculated given the remote IP address of the peer and all information than can be inferred from it, including previously stored.

Life-cycle of LM-capable node in P2P network may start with initiating startup sequence upon achieving connectivity to the Internet.

LM implementation works by initially preloading structural information from publicly accessible services called Regional Internet Registries (RIRs) and converting it into an internal graph-like data structure.

From the practical standpoint, the result of the first layer of LM is a class of target node in relation to originator node. Proposed flavors are:

- **Subrange** identifies the presence of the remote node's IP address in the same IPv4 subrange specified in the database file dealing with administrative IP subranges. This flavor identifies the presence of the remote node most likely within the scope of operation of a single router or local area network.
- **Range** identifies the presence of the remote node's IP address in the same IPv4 range specified in the *delegated-* file or the *ripe.db.route.gz* WHOIS excerpts dealing with ASNs and IPv4 delegations. This flavor means that both nodes most likely reside within the scope of the same department and that the traffic between these nodes is unlikely to travel outside.
- **AS** identifies the presence of the remote node's IP address within the address space allocated to the same AS as defined in the *ripe.db.route.gz* file. This flavor suggests that the traffic between two nodes is handled by the ISP internally.
- **ASSET/IX** states that both nodes belong to different ASes announced by the same ASSET and therefore is able to implicitly communicate through it.
- **ASSET-link** indicates that the node addresses belong to different ASes, which belong to different ASSETs, and at least one ASSET includes the other ASSET.
- **Backbone** indicates that the node addresses belong to different ASes, which belong to different ASSETs, and both ASSETs are declared within the scope of a third ASSET.
- **Distant** identifies that all previous affinity tests had failed to yield a positive match and the relative locality of the originator and target node cannot be reliably estimated. Therefore they are assumed to be located topologically far away.

Experiments

Certainly, the true impact of the LM implementation on the network efficiency cannot be reliably measured until this functionality makes its way into existing popular clients of P2P networks. However, estimations could be done, assuming that nodes with specific locality flavors will provide for significant speed bursts.

The test surveys were conducted on the swarms formed around content units published on the private Russian BitTorrent tracker RuTracker.Org. The survey consists of accumulating 6000 IP addresses per 20 highly-active swarms and calculating LM flavor of local node against each of them.

The results suggest that given traffic shaping relevant to the presence within UA-IX coverage, if non-distant peers were queried prior to distant ones, the originator node might have had no need to contact distant peers at all, having exhausted its inbound bandwidth capacity with such "filtered out" high-speed peers. It is also interesting to note that the number of active peers under UA-IX coverage is significantly less than

Conclusion and future works

In our future work we would like to address the additional layers of LM calculated by direct measurements involving additional traffic. These layers may be expressed as weighted scores by which all peer priorities are then fine-tuned within the boundaries of their respective first-layer flavors. It should be noted, that implementation of an additional LM layers will require modifications to the existing software and creating extensions to existing protocols in order to have any impact on the performance.

We are currently developing the software library implementing LM method under LGPL license to assist software engineers wishing to optimize performance of their P2P applications.

References

1. *B. Cohen*, "Incentives build robustness in BitTorrent" // First Workshop on the Economics of Peer-to-Peer Systems, 2003.
2. *A. Feldmann and V. Aggarwal*, "Isp-aided neighbor selection in p2p systems," in IETF P2P Infrastructure Workshop (P2Pi), Berlin, Germany, May 2008.
3. *D. R. Choffnes and F. E. Bustamante*, "Taming the torrent: a practical approach to reducing cross-isp traffic in peer-to-peer systems," SIGCOMM Comput. Commun. Rev., vol. 38, no. 4, pp. 363–374, 2008.
4. *H. Xie, A. Krishnamurthy, A. Silberschatz, and R. Y. Yang*, "P4P: Explicit Communications for Cooperative Control Between P2P and Network Providers," DCIA P2P MARKET CONFERENCE, 2008.
5. *B. Augustin, X. Cuvellier, B. Orgogozo, F. Viger, T. Friedman, M. Latapy, C. Magnien, and R. Teixeira*, "Avoiding traceroute anomalies with Paris traceroute" in IMC '06: Proceedings of the 6th ACM SIGCOMM conference on Internet measurement. New York, NY, USA: ACM, 2006, pp.153–158.