ISP-Aided Neighbor Selection in P2P Systems

Vinay Aggarwal
Deutsche Telekom Labs/TU Berlin
vinay@net.t-labs.tu-berlin.de

Anja Feldmann
Deutsche Telekom Labs/TU Berlin
anja@net.t-labs.tu-berlin.de

I. Motivation

P2P systems are so popular that they contribute more than 50% to the overall network traffic in the Internet. The widespread use of such P2P systems has put ISPs in a dilemma! On the one hand, P2P system applications have resulted in an increase in revenue for ISPs, as they are one of the major reasons cited by Internet users for upgrading their Internet access to broadband. On the other hand, ISPs find that P2P traffic poses a significant traffic engineering challenge. P2P traffic often starves other applications like Web traffic of bandwidth, and swamps the ISP network. This is because most P2P systems rely on application layer routing based on an overlay topology on top of the Internet, which is largely independent of the Internet routing and topology.

To construct an overlay topology, unstructured P2P networks usually employ an arbitrary neighbour selection procedure. This often results in a situation where a node in Frankfurt downloads a large content file from a node in Sydney, while the same information may be available at a node in Berlin. It has been shown that P2P traffic often crosses network boundaries multiple times. This is not necessarily optimal as most network bottlenecks in the Internet are assumed to be either in the access network or on the links between ISPs, but not in the backbones of the ISPs. Besides, studies have shown that the desired content is often available “in the proximity” of interested users. This is due to content language and geographical regions of interest. Since a P2P user is primarily interested in finding his desired content quickly with good performance, we believe that increasing the locality of P2P traffic will benefit both ISPs and P2P users.

P2P systems form overlays at the application layer, which are virtual networks formed on top of the underlying Internet routing infrastructure. As such, the logical paths and links of an overlay lie on top of the physical paths set up by intra-domain (e.g., OSPF, MPLS, IS-IS) and inter-domain (e.g., BGP) routing protocols running at the Internet underlay. Hence, when the overlay nodes cooperate with each other to route packets on behalf of any pair of communicating nodes, the traffic is still carried through the physical Internet routing paths.

It has been shown that overlay routing can enable users access to paths with potentially better performance than those made available by the Internet. However, ISPs use traffic engineering (TE) to provide better routing performance to their customers. This leads to the situation that P2P systems are reinventing and reimplementing a routing system whose dynamics interact with the dynamics of the Internet routing system. The goals of overlay routing and ISP’s traffic engineering are not aligned. An overlay tries to find optimal routing paths between its own peers, while the ISP has to keep in mind the whole network performance, which includes all the underlay as well as the overlay users. This not only leads to duplication of routing functionality, but also to inefficient routing path oscillations and triangle inequalities.

In summary, we identify the following drawbacks:

- The ISP has limited ability to manage its traffic and therefore incurs potentially increased costs for its interdomain traffic, as well as for its inability to do traffic engineering on its internal network.
- The P2P system has limited ability to pick an optimal overlay topology and therefore provide optimal performance to its users, as it has no prior knowledge of the underlying Internet topology. It therefore has to either disregard or reverse engineer it.
- Different P2P systems have to measure the path performance independently.

While we do not know of a P2P network that tries to reverse-engineer the Internet topology, there are some proposals that suggest that P2P networks should bias their overlay topology by choosing neighbours that are close in the sense of high throughput or low latency, e.g., [1], [2], [3], [4] or that are within the same AS, e.g., [5], [6]. Others such as the Brocade [7] system propose to build an overlay on top of a structured DHT-based P2P system that exploits knowledge of the underlying network characteristics. Yet another system [8] proposes to use caching to relieve the tension between ISPs and P2P systems. A recent proposal [9] uses iTrackers as portals of network providers to enable ISP-P2P collaboration.

II. An Oracle Service

We propose a simpler solution where ISPs and P2P systems collaborate so that both benefit. Instead of the P2P node choosing neighbours independently, the ISP can offer a service, which we call the *oracle*, that ranks the potential neighbours according to certain metrics. This ranking can be seen as the ISP expressing preference for certain P2P neighbours. Possible coarse-grained distance metrics are:

- inside/outside of the AS
- number of AS hops according to the BGP path
- distance to the edge of the AS according to the IGP metric

For P2P nodes within the AS the oracle may further rank the nodes according to:

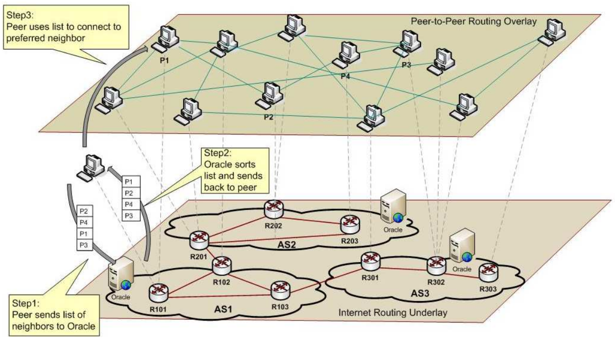

Fig. 1. How the Oracle works

- connection information such as: last-hop bandwidth
- geographical information such as: same point of presence (PoP), same city
- performance information such as: expected delay, available bandwidth
- link congestion (traffic engineering)

This ranking can then be used by the P2P node to select a closeby neighbour although there is no obligation. Figure 1 summarizes the operation of the oracle.

The oracle acts like an abstract routing underlay to the overlay network. But as it is a service offered by the ISP, it has direct access to the relevant information and does not have to infer or measure it. For example, an ISP knows whether a customer has a DSL broadband or a modem connection, its link delay, etc.

The ISP benefits in multiple ways:

- they can now influence the P2P routing decisions via the oracle and thus regain their ability to perform traffic engineering (control the traffic flow)
- by influencing the neighbourhood selection process of the P2P network, they can keep a significant portion of their network traffic localized within their internal network, and hence gain cost advantages by reducing costs for traffic that leaves their network boundary
- the P2P measurement traffic to infer network distances is omitted
- due to the ability to better manage their traffic flow, they can provide better service to their customers and ensure fairness for other applications like Web traffic, etc., especially at times of peak demand

The benefit to P2P nodes of all overlays is also multifold:

- they do not have to measure the path performance themselves
- they can take advantage of the knowledge of the ISP
- they can expect improved performance in the sense of low latency and high throughput as bottlenecks can be avoided.

As the ability to control/manage its traffic is crucial to the operating costs of every ISP, we expect that the benefit accruing from this ability will outweigh the potential risks of providing an oracle, namely that the oracle exposes some information about the ISP topology and the network performance. As the oracle server only needs to roughly rank the IP nodes, it does not need to reveal more information about its network than can anyhow be inferred by reverse-engineering the ISP network via measurements. Moreover, the ISP does not reveal the exact details of the criteria used in sorting the list of neighbours.

The oracle is available to *all* overlay networks. One does neither need nor want to use a separate oracle for each P2P network. Furthermore, as an open service, it can be queried by any application and is not limited to file-sharing systems. The oracle can be used by any application where the users have a choice of more than one destination to connect to. Possible examples are content distribution networks (CDN), mirror websites, etc. The larger is the content to be exchanged among peers, the more will be the benefits of using the oracle, both to the applications as well as the ISPs.

This also implies that querying the oracle does not necessarily imply participation in file sharing systems. The oracle acts as a *peer mapping service*, which helps users of an application to select "good" neighbours. This should limit the desirability of the oracle logs to, e.g., the music industry. Moreover, to protect their identity, the P2P users could permute, e.g., the last byte of the IP addresses it is interested in or use an anonymization service for querying the oracle.

The oracle service can be realized as a single server or a set of replicated servers within each ISP, that can be queried using a UDP-based protocol or run as a Web service. It can rely on a semi-static database with the ISP's prefix and topology information. Updating this information should not impose any major overhead on the ISP.

We have also proposed a scheme whereby oracles from multiple ISPs can collaborate to build a global coordinate system [10]. It can be used by P2P and other applications to get estimates of path properties to potential neighbours/servers within and outside their ISPs. The coordinate system is built through ISP-P2P collaboration on the one hand, and collaboration betweem multiple ISPs (to exchange summaries of network connectivity information between them) on the other hand.

III. How the Oracle works

With the help of an example, we show how the P2P users can use the oracle service. Consider the example network shown in Figure 2. It shows the simplified internal topology of a hypothetical ISP, with various users A, B, C, D, and E having different connection bandwidths at the connection edge. The oracle service runs on a publicly known IP address, and has a map of its entire network topology, containing information like link bandwidths, router topology, etc.

When user A wishes to connect to another peer for bootstrapping to a P2P network, we assume that it finds B and E as possible candidates through a P2P bootstrapping mechanism, e.g., a Web Cache or previously stored list of active users. Now, A queries the oracle server for path properties of B and E. The oracle server knows that B has a last hop bandwidth of 16 Mbit, which is much larger than the 4 Mbit bandwidth of E. Hence, it recommends A to connect to E. The oracle can either

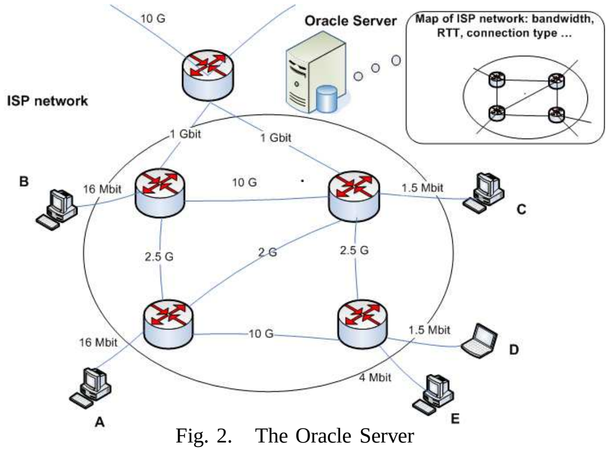

Fig. 2. The Oracle Server

rank B ahead of E, or can return a bandwidth classification of B as high, and E as medium. This enables A to connect to a user having a better bandwidth.

Consider another instance, when E is already connected to both C and D in a P2P network. When E wishes to download a large multimedia file, it queries the oracle about its connected neighbours. The oracle knows that even though both D and C have similar last-hop bandwidth, the node D is topologically and/or geographically closer to E than C. If E downloads the large multimedia file from D instead of C, it will use up lesser network resources and cause lesser network congestion. Hence, the oracle recommends D over C to the querying node E.

IV. Evaluation

We have analyzed this scheme on various models of P2P systems, with both application layer and network layer simulation frameworks. We have also conducted Testbed experiments and Planetlab deployment. Our results show that P2P users, on consulting the oracle, are able to keep most of their peerings within the ISP boundaries, without any adverse effects on the overlay graph structural properties like small node degree, small path length and connectedness. Not only does the scalability of P2P systems improve considerably, we also show that there is no adverse effect on the query search phase of P2P networks. The P2P users are still able to locate all available content, and the content download times decrease considerably. Congestion analysis shows that the traffic distribution using the oracle is close to the theoretical optimum. The ISPs are able to save costs by keeping large amount of traffic within their network, peform better traffic engineering, and provide better service to customers. We have demonstrated that the scheme maintains its benefits across various models of P2P user behaviour, namely churn, free-riding, and query patterns, as well as different ISP and P2P topologies. These results have been published in [11], [12], [13].

With the Internet transforming from a client-server model to a user-generated-content model, where different nodes generate, search, seek, and download content at the same time, and where the content ranges from low negotiation traffic to heavy multimedia content, we believe that collaboration between ISPs and P2P systems can contribute significantly to both ISPs as well as Internet users.

References

[1] S. Ratnasamy, M. Handley, R. Karp, and S. Shenker, “Topologically aware overlay construction and server selection,” in *INFOCOM*, 2002.

[2] K. Shanahan and M. Freedman, “Locality Prediction for Oblivious Clients,” in *IPTPS*, 2005.

[3] M. Adler, R. Kumar, K. Ross, D. Rubenstein, T. Suel, and D. Yao, “Optimal Selection of Peers for P2P Downloading and Streaming,” in *INFOCOM*, 2005.

[4] A. Su, D. Choffnes, F. Bustamante, and A. Kuzmanovic, “Relative Network Positioning via CDN Redirections,” in *ICDCS*, 2008.

[5] T. Karagiannis, P. Rodriguez, and K. Papagiannaki, “Should ISPs fear Peer-Assisted Content Distribution?” in *IMC*, 2005.

[6] R. Bindal, P. Cao, W. Chan, J. Medved, G. Suwala, T. Bates, and A. Zhang, “Improving Traffic Locality in BitTorrent via Biased Neighbor Selection,” in *IEEE ICDCS*, 2006.

[7] B. Zhao, Y. Duan, L. Huang, A. Joseph, and J. Kubiatowicz, “Brocade: Landmark Routing on Overlay Networks,” in *IPTPS*, 2002.

[8] G. Shen, Y. Wang, Y. Xiong, B. Zhao, and Z. Zhang, “HPTP: Relieving the Tension between ISPs and P2P,” in *IPTPS*, 2007.

[9] H. Xie, Y. Yang, A. Krishnamurthy, and A. Silberschatz, “P4P: Provider Portal for P2P Applications,” in *SIGCOMM*, 2008.

[10] V. Aggarwal, A. Feldmann, and R. Karrer, “An Internet Coordinate System to Enable Collaboration between ISPs and P2P Systems,” in *ICIN*, 2007.

[11] V. Aggarwal, A. Feldmann, and C. Scheideler, “Can ISPs and P2P Systems Cooperate for Improved Performance?” in *ACM SIGCOMM Computer Communication Review, 37(3)*, 2007.

[12] V. Aggarwal and A. Feldmann, “Locality-Aware P2P Query Search with ISP Collaboration,” in *Networks and Heterogeneous Media, 3(2)*, 2008.

[13] V. Aggarwal, O. Akonjang, and A. Feldmann, “Improving User and ISP Experience through ISP-aided P2P Locality,” in *11th IEEE Global Internet Symposium*, 2008.