AURORA: Autonomous System relationships-aware overlay routing architecture in P2P CDNs

Article · January 2009

2 authors:

Hirochika Asai
The University of Tokyo

Hiroshi Esaki
The University of Tokyo

AURORA: Autonomous System Relationships-aware Overlay Routing Architecture in P2P CDNs

Hirochika Asai
University of Tokyo
panda@hongo.wide.ad.jp

Hiroshi Esaki
University of Tokyo
hiroshi@wide.ad.jp

ABSTRACT

Content Delivery Networks (CDNs) utilizing peer-to-peer (P2P) technology have been developed and deployed to achieve high-quality content delivery over the Internet. Since the P2P CDNs do not take into account the routing policy and the economics among Autonomous Systems (ASes), P2P CDNs consume a larger amount of network resources. In this paper, we propose two types of overlay routing metrics to control inter-domain traffic based on AS relationships and the economics: 1) *cost-based* and 2) *magnitude-based* metric. We show that the peer selection algorithm with the proposed metrics can reduce inter-domain traffic, especially transit traffic by computer simulation.

Categories and Subject Descriptors

C.2.1 [**Computer Communication Networks**]: Network Architecture and Design—*Network communications, Network topology*

General Terms

Algorithms, Design, Economics

Keywords

content delivery, overlay routing, autonomous system relationships, economics

1. INTRODUCTION

Content Delivery Networks (CDNs) are widely employed to distribute large content files such as music files, movie files and software images. Various CDNs utilizing peer-to-peer (P2P) technology have been developed and deployed to avoid excessive server load and to achieve robustness and high-quality content delivery over the Internet[2]; these P2P CDNs constitute so-called overlay networks. Since the topology of P2P CDNs is generally different from the layer 3 network topology and it is hard to take into account the inter-domain routing policy and the economics, these P2P CDNs frequently consume a larger amount of network resources and cost more from the layer 3 operators' viewpoint[7]; that is, the path with higher cost links can be selected for the content delivery despite the existence of paths with lower cost links in P2P CDNs.

In previous studies [4, 8, 10], Round Trip Time (RTT) and hop count have been used in their peer selection algorithms to achieve high-quality (i.e., low delay and high throughput) content delivery. In a recent study [11], *Xie et al.* have proposed an architecture in achieving efficient and fair utilization of network resources in the Internet using both intra- and inter-domain topology information as an overlay routing metric. These studies, however, have never taken into account the relationships among autonomous systems (ASes) and inter-domain economics. Consequently, these studies cannot utilize inter-domain resources efficiently from the layer 3 operators' and the economical viewpoint in considering that customer ASes purchase the Internet access from provider ASes with payment for their actual bandwidth usage. Hence, it is required to reduce inter-domain traffic, especially traffic over transit links for customer ASes.

In this paper, we propose two types of overlay routing metrics to control inter-domain traffic based on AS relationships and the economics: 1) *cost-based* and 2) *magnitude-based* metric. We evaluate the peer selection algorithm adopting these metrics by computer simulation. Two main findings of this paper are: 1) the peer selection algorithm adopting the proposed metrics can adequately reduce inter-domain traffic, especially transit traffic, and 2) the peer selection algorithm adopting the magnitude-based metric can reduce transit traffic without non-disclosure AS relationships information itself but only the neighboring information.

2. AUTONOMOUS SYSTEMS, INTER-DOMAIN ROUTING AND ECONOMICS

The Internet consists of thousands of ASes operated by many different administrative domains such as Internet service providers (ISPs), companies and universities. The routing among ASes is determined by the inter-domain routing protocol such as Border Gateway Protocol (BGP) [5].

The relationships between neighboring ASes can be categorized into three groups [6, 9];

1. *Transit*: An AS purchases the Internet access from another AS by paying some amount of money according to the bandwidth usage. This is also called customer-to-provider (c2p) relationship or provider-to-customer (p2c) relationship.

Table 1: inter-AS link cost

relationship	s2s	p2p	p2c	c2p
cost	1.00000	1.00001	1.00032	1.01024

2. *Peering*: A pair of neighboring ASes can exchange traffic directly and the traffic exchanged between two peering ASes is free of charge. However, when the traffic becomes highly asymmetric, one party *will* start charging the other according to the bandwidth usage [7]. This is also called peer-to-peer (p2p [1]) relationship.

3. *Sibling*: Multiple ASes can belong to the same organization. Even though each AS might be managed separately from the perspective of network administration, traffic can be exchanged among them *without* any payment.

The routing policy among ASes results in *valley-free* paths [5] in the inter-domain routing. A *valley-free* path means that each path between two arbitrary ASes traverses first uphill (c2p) links, and goes across one peering link at most, and then traverses downhill (p2c) links when we ignore sibling links. In this paper, we assume the inter-domain routing obeys this *valley-free* path model.

As described above, customer ASes pay some amount of money in exchanging the traffic with their provider ASes. Therefore, customer ASes are willing to reduce their traffic exchanged with their provider ASes from the economical viewpoint.

3. OVERLAY ROUTING

3.1 Overview

We define *underlay distance* as an overlay routing metric on the basis of AS relationships. We use the symbol d_{ij} and the vector $\boldsymbol{p}_{ij}$ as the underlay distance from AS i to AS j and the AS path from AS i to AS j, respectively, throughout this paper. We adopt two approaches to define the underlay distance from AS i to AS j: 1) cost-based and 2) magnitude-based approach. In the cost-based approach, we directly transform AS relationships into cost as the cost-based underlay distance (CBUD). On the other hand, in the magnitude-based approach, we first calculate ASes' magnitude based on the traffic flow by using eigenvector analysis [3] and use the difference of the magnitude between two neighboring ASes as the magnitude-based underlay distance (MBUD).

3.2 Routing Metric

3.2.1 CBUD: Cost-Based Underlay Distance

We use the information on AS relationships as a routing metric in the cost-based approach. We define the underlay distance in summation of all the inter-AS link's cost through the AS path between two arbitrary ASes. Equation (1) is the definition of the accumulated CBUD from AS i to AS j.

$$d_{ij} = \sum_{link \in \boldsymbol{p}_{ij}} cost_{link} \tag{1}$$

In Equation (1), $cost_{link}$ denotes the cost of the inter-AS link ($link \in \boldsymbol{p}_{ij}$) given in Table 1. We adopt the cost table listed in Table 1 in this paper. Within 31 AS hops, the minimal CBUD adopting this cost table is equivalent to the following; the evaluation order is, 1) minimal AS hops, 2) minimal c2p links, 3) minimal p2c links, 4) minimal p2p links, 5) minimal s2s links, and 6) random.

This approach is very naive and it is easy to calculate the underlay distance. However, there is a difficulty in disclosing the real cost of inter-AS links since AS relationships are generally non-disclosure information for the most of commercial ISPs.

3.2.2 MBUD: Magnitude-Based Underlay Distance

The problem of non-disclosure AS relationships can be solved by adopting our magnitude-based inter-AS distance estimation method [3]; we have shown that we can estimate the AS relationships form an AS adjacency matrix (i.e., neighboring information). We adopt the difference of the magnitude for the MBUD. The accumulated MBUD between AS i to AS j is defined in Equation (2).

$$d_{ij} = \sum_{link \in \boldsymbol{p}_{ij}} \left(1 + \frac{\Delta\Psi^{(n)}_{link}}{max(\boldsymbol{\Psi}^{(n)})} \right) \tag{2}$$

In Equation (2), $\Delta\Psi_{link}$ denotes the difference of the magnitude between two edge ASes of the link ($link \in \boldsymbol{p}_{ij}$), and $max(\boldsymbol{\Psi})$ denotes the maximum value in the estimated magnitude for each AS.

4. EVALUATION

We adopted the proposed overlay routing metrics for the peer selection algorithm in P2P CDNs to evaluate them. The most appropriate node which minimizes the underlay distance is selected to communicate with.

We evaluated the algorithm adopting the proposed metrics through the computer simulation with a certain communication model and "The CAIDA AS relationships dataset (09/06/2008)" [1]. We measured inter-domain traffic by AS relationships in four peer selection methods; 1) uniformly random selection, 2) minimal AS hops selection, 3) minimal CBUD selection and 4) minimal MBUD selection (for $n \in \{0, 2\}$; n denotes the recursion times in estimating the magnitude). The uniformly random selection is the most naive algorithm, and the minimal AS hops selection is the alternative to the minimal hops selection since we do not deal with the intra-domain topology in the simulation. The minimal CBUD selection and the minimal MBUD selection are the algorithms adopting the proposed metrics.

We show that both the minimal CBUD selection and the minimal MBUD selection algorithms reduce inter-domain traffic by 45.5% in comparison with the random selection algorithm. Moreover, the minimal CBUD selection algorithm reduces the traffic through transit links by 16.4% and the minimal MBUD selection algorithm for $n = 2$ reduces the traffic through transit links by 11.1% in comparison with the minimal AS hops selection algorithm (Figure 1).

[1]We use lower-case letters for peer-to-peer relationship in AS relationships, and upper-case P2P denotes peer-to-peer technology in this paper.

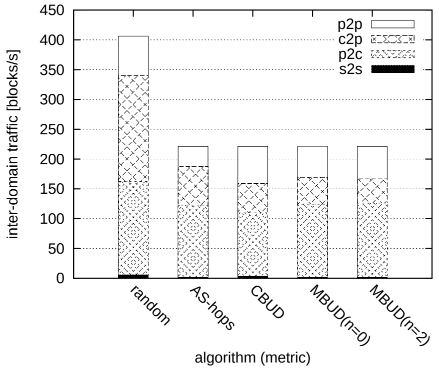

Figure 1: Simulation result: The breakdown of the average of inter-domain traffic (total traffic of all inter-AS links) by AS relationships. The minimal CBUD selection algorithm reduces transit traffic by 16.4% **and the minimal MBUD selection algorithm for** $n = 2$ **reduces transit traffic by** 11.1% **in comparison with the minimal AS hops selection algorithm.**

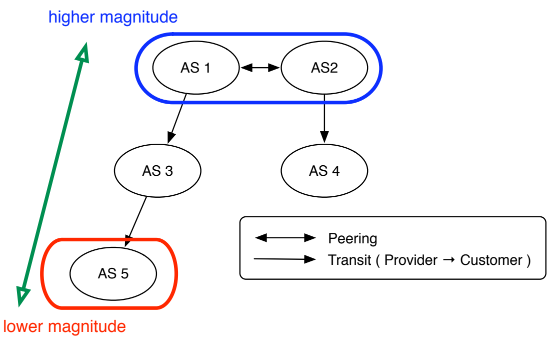

Figure 2: Valley-free path topology and AS magnitude. The magnitude of each AS represents the position in the hierarchical valley-free path.

5. DISCUSSION

The cost-based peer selection algorithm is very naive solution to control inter-domain traffic, but it is essential for the inter-domain traffic control to achieve consensus on the cost table among all ASes. We used the cost translation table listed in Table 1 and we imposed to use this table on all ASes in this paper.

We have shown the proposed metrics are applicable for peer selection. We consider the proposed metrics can be applicable for overlay routing, but we should solve how to construct the overlay network topology (e.g., neighbor selection) over the layer 3 clouds before applying them for overlay routing. Our magnitude estimation method [3] can quantify each AS in magnitude, so we will adopt it for the overlay network topology construction process such as neighbor selection (Figure 2).

6. CONCLUSION

We proposed cost-based and magnitude-based underlay distance as novel overlay routing metrics. The proposed overlay routing metrics deliver the following contributions; 1) the peer selection algorithm adopting the proposed metrics can adequately reduce inter-domain traffic, especially high-cost transit traffic, and 2) the peer selection algorithm adopting the magnitude-based metric can reduce transit traffic without the non-disclosure AS relationships information itself but only the neighboring information.

As the future work, we will discuss and evaluate an AS relationships-aware overlay routing architecture including the overlay network construction method. Then, we will implement the architecture and evaluate it in the real Internet.

7. REFERENCES

[1] The CAIDA AS Relationships Dataset (09/06/2008). http://www.caida.org/data/active/as-relationships/.

[2] *Peer-to-Peer Systems and Applications* (2005), vol. 3485 of *Lecture Notes in Computer Science*, Springer Berlin / Heidelberg.

[3] Asai, H., and Esaki, H. A methodology for magnitude-based inter-as distance estimation (to appear). In *The Tenth Workshop on Internet Technology – WIT2009* (June 2009).

[4] Dabek, F., Cox, R., Kaashoek, F., and Morris, R. Vivaldi: a decentralized network coordinate system. In *SIGCOMM '04: Proceedings of the 2004 conference on Applications, technologies, architectures, and protocols for computer communications* (New York, NY, USA, 2004), ACM, pp. 15–26.

[5] Erlebach, T., Hall, A., Panconesi, A., and Vukadinovi, D. Cuts and disjoint paths in the valley-free path model of internet BGP routing. In *Combinatorial and Algorithmic Aspects of Networking* (2005), vol. 3405/2005 of *Lecture Notes in Computer Science*, Springer Berlin / Heidelberg, pp. 49–62.

[6] Gao, L. On inferring autonomous system relationships in the internet. *Networking, IEEE/ACM Transactions on 9*, 6 (Dec 2001), 733–745.

[7] Huang, C., Li, J., and Ross, K. W. Can internet video-on-demand be profitable? In *SIGCOMM '07: Proceedings of the 2007 conference on Applications, technologies, architectures, and protocols for computer communications* (New York, NY, USA, 2007), ACM, pp. 133–144.

[8] Li, W., Chen, S., and Yu, T. Utaps: An underlying topology-aware peer selection algorithm in bittorrent. *Advanced Information Networking and Applications, 2008. AINA 2008. 22nd International Conference on* (March 2008), 539–545.

[9] Shakkottai, S., and Srikant, R. Economics of network pricing with multiple isps. *IEEE/ACM Trans. Netw. 14*, 6 (2006), 1233–1245.

[10] Waldvogel, M., and Rinaldi, R. Efficient topology-aware overlay network. *SIGCOMM Comput. Commun. Rev. 33*, 1 (2003), 101–106.

[11] Xie, H., Yang, Y. R., Krishnamurthy, A., Liu, Y. G., and Silberschatz, A. P4p: provider portal for applications. In *SIGCOMM '08: Proceedings of the ACM SIGCOMM 2008 conference on Data communication* (New York, NY, USA, 2008), ACM, pp. 351–362.