# Proxima : Towards Lightweight and Flexible Anycast Service

Guodong Wang[1], Yang Chen[1], Lei Shi[2], Eng Keong Lua[3], Beixing Deng[1] and Xing Li[1]
[1] Department of Electronic Engineering, Tsinghua University, 100084, China
[2] Department of Computer Science, Tsinghua University, 100084, China
[3] Carnegie Mellon University, Pittsburgh, PA15213, U.S.A.
E-mail: {wgd08,chenyang04}@mails.tsinghua.edu.cn

***Abstract*—Anycast paradigm has been widely adopted by Internet application to find nearby resources. The current IP anycast implementation suffers from poor scalability. To overcome that, this paper proposes Proxima, a network coordinates (NC) based application layer infrastructure, which is capable to provide lightweight and flexible anycast service. Proxima accurately chooses the best receiver from a group of candidates according to the Round Trip Time (RTT) for certain application. Our experimental results have demonstrated the excellent performance of Proxima in real sever selection scenario.**

## I. Introduction

IP anycast, which allows a networking node to send a message to a near member within a group, was proposed in RFC 1546 [1]. It provides an efficient way to find nearby resources. However, till now, few global services are using IP anycast, mainly due to its poor scalability - anycast IPs can't be aggregated in the routing table. One potential solution to get rid of this bottleneck is to use application layer anycast [5]. Application layer anycast system can be deployed globally, while allowing the node to select the best receiver according to different metrics such as end-to-end latency, bandwidth and load of servers. It has been used to improve the performance of many content distribution Internet services.

OASIS [2] is an implementation of application layer anycast system. However, it depends on a centralized database to store the (IP prefix, GeoLocation) mapping, which will lead to high maintain overhead due to the IP address re-assignment. In practice, it's extremely hard to keep this database up-to-date. Moreover, OASIS uses geographic coordinates to select servers, but the geographical closer distance does not trivially imply the smaller RTT due to the complex Internet routing policy. In order to avoid the complexity of maintaining large-size centralized database and to reference to the RTT when delivering the anycast service, we propose Proxima as our solution for application layer anycast infrastructure.

Network coordinate (NC) system [4] is a scalable method to accurately estimate RTT between any two hosts in the Internet with small overhead. In this paper, we propose a lightweight and flexible NC based application layer infrastructure Proxima, which provides reliable anycast service. Proxima can be deployed globally and help the host to select the proper destination in any application. Proxima is a lightweight system since it leverages NC to estimate RTT between any two hosts in order to eliminate the expensive on-demand probing process. Proxima is also a flexible system as it makes decision over the up-to-date RTT information: upon each anycast request, Proxima estimates RTTs between all servers and client (here client stands for the sender and servers stand for a group of potential receivers in the anycast service scenario), and selects the server with the smallest RTT to the client.

Our experiments over PlanetLab testbed show that the performance of Proxima is much better than naive Round-Robin (RR) selection strategy that is the simplest method to map each request to different servers. Another advantage of Proxima is that it doesn't need to maintain the complicated (IP prefixes,GeoLocation) mapping as in OASIS.

The rest of the paper is organized as follows. In section 2 we describe the detailed design and implementation of Proxima. Section 3 gives performance evaluation results. Finally we conclude this paper in section 4.

## II. Proxima Design

### *A. System Architecture*

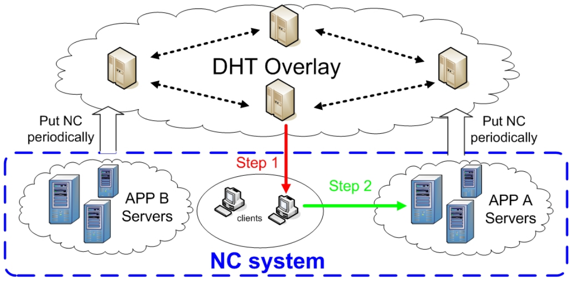

Fig. 1. Proxima architecture

Figure 1 shows the high level architecture of Proxima. In contrast to OASIS, Proxima is a fully decentralized system. It contains two basic decentralized components.

*1) NC System (Pharos):* Each host in the Internet can get a synthetic coordinate, with which RTT between two hosts can be estimated accurately, as long as it joins the NC system. All the servers and clients participate in the NC system and compute their own virtual coordinates. Every node updates

978-1-4244-3968-3/09/$25.00 ©2009

its coordinates periodically in order to keep the prediction accurate.

In [2], a Vivaldi based anycast is evaluated and the performance is bad, sometimes even worse than Round-Robin scheme. But Vivaldi with two dimensional Euclidean coordinate, which is used in [2], is not accurate enough. Besides in their evaluation, they didn't use the latency and update filters [3] to improve the prediction accuracy and stability of NC in real Internet. In Proxima, Pharos [4] which is more accurate than Vivaldi is used as NC system. There are two layer NCs-global NC and local NC in Pharos. The coordinate in each layer is two level four dimensional Euclidean coordinate with a height. Moreover, latency and update filters [3] are used to guarantee NCs are stable. Benefiting from NC, Proxima uses estimated RTT as a metric to select server, so that the overhead is low and the server selection response latency is small.

*2) Distributed Hash Table (DHT) Overlay:* A group of dedicated nodes are deployed as the DHT overlay, where the coordinates of all the servers is stored according to the IPs, regardless of service. DHT overlay can provide robust decentralized database. Since the virtual coordinates of the hosts change according to the fluctuation of network RTTs, a server calls DHT API-put(SHA-1(IP),coordinates) to store coordinates as long as its own coordinates are updated.

Proxima can utilize any implementation of DHT system including the public DHT services such as OpenDHT [6]. Furthermore, it can guide the anycast using the up-to-date latency information and select nearest destination in any service effectively.

### B. *Work flow of Proxima*

Compared with OASIS, Proxima uses RTT to select server, it is more reasonable. Since NCs are updated periodically, Proxima uses the up-to-date RTT information from Internet to direct anycast.

Figure 1 also shows the work flow of Proxima when a client C proposes an anycast request. At first, client C gets a group of candidate server IPs according to the application. Different applications will have different methods to give a group of potential destinations, such as AND resolver [5], tracker in P2P system etc. Then Proxima is answering the most important question "which is the best server". As *step 1* in figure 1, C calls DHT API-get(SHA-1(IP)) to get server coordinates according to the server IPs. Then in *step 2*, C calculates the up-to-date network distances (RTT) from himself to all the candidate servers with the virtual coordinates so that to select the nearest server which has the smallest RTT to C.

## III. Performance Evaluation

From Dec 29, 2008 to Jan 4, 2009, we carry out experiments on 101 PlanetLab hosts around the world. We choose 15 nodes from the 101 nodes as the candidate servers, while the rest of the nodes will be treated as clients. We use OpenDHT [6] to implement DHT overlay. In Pharos implementation, all the nodes are divided into three clusters-American cluster, European cluster and Asian cluster according to their network distances to the anchors which are famous web sites in different regions.

As in [2], we compare the performance of Proxima with that of Round-Robin scheme. Two consecutive server selection experiments are run at each client host. The experiments use Proxima and Round-Robin strategies to select server respectively. We measure RTT from client to selected server with ICMP ping and measure the bandwidth with iperf-2.0.4 (the TCP window size is 64K). All the data from two experiments are included in the following Cumulative Distribution Function (CDF) graphs.

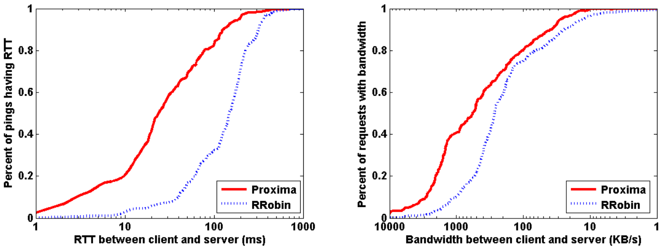

Fig. 2. Client-server RTTs and TCP bandwidth CDFs

Figure 2 shows the CDFs of the RTTs measured by ICMP ping between client and selected server and the CDFs of bandwidth between client and select server. We can see that Proxima significantly outperforms anycast using round robin, both considering RTTs and bandwidths. And Proxima avoids maintaining the centralized database.

## IV. Conclusions

This paper proposes Proxima, a lightweight and flexible anycast service system using Network Coordinate. The sparking point of Proxima is to use accurate NC system to support the scalable selection of the best server with up-to-date latency information. Our experimental results on PlanetLab demonstrate that Proxima is an effective infrastructure to support global anycast in a scalable way.

## Acknowledgments

This work is supported by the National Basic Research Program of China (No.2007CB310806) and the National Science Foundation of China (No.60473087, No.60703052, No. 60850003).

## References

[1] C. Partridge, T. Mendez, and W. Milliken. "RFC 1546 - Host Anycasting Service". Nov. 1993.

[2] J. Freedman, K. Lakshminarayanan, and D. Mazieres. "Oasis: Anycast for any servic". *Proc. of Usenix NSDI*, May 2006.

[3] J. Ledlie, P. Pietzuch, and M. Seltzer. "Stable and Accurate Network Coordinates". *Proc. of ICDCS*, July 2006.

[4] Y. Chen, Y. Xiong, X. Shi, et al. "Pharos: Accurate and Decentralised Network Coordinate System". *To appear in IET Communications*.

[5] E. Zegura, M. Ammar, Z. Fei et al. "Application-Layer Anycasting: A Server Selection Architecture and Use in a Replicated Web Service". *IEEE/ACM TRANSACTIONS ON NETWORKING*, Aug. 2000

[6] S. Rhea, B. Godfrey, B. Karp et al. "OpenDHT: A Public DHT Service and Its Uses". *Proc. of ACM SIGCOMM*, Aug. 2005.