P3ON: Proximity-based Peer-to-Peer Overlay Networks

Kunwoo Park, Sangheon Pack, and Yanghee Choi
Advisor: Taekyoung Kwon
School of Computer Science and Engineering
Seoul National University, Seoul, Korea
Email: {kwpark, shpack}@mmlab.snu.ac.kr and {yhchoi, tkkwon}@snu.ac.kr

***Abstract*— Peer-to-peer(P2P) overlay networks provide robust, scalable file sharing services. The hop distance between two virtual nodes in P2P overlay networks is irrelevant to the real distance between two nodes. Therefore, many lookup operations may be executed, without considering the real distance, even though a node wants to find a file object located in another node within the same subnet. To overcome this drawback, we propose a proximity-based peer-to-peer overlay networks (P3ON). In P3ON, the proximity between two nodes is exploited to achieve a fast lookup operation. In other words, two nodes within the same subnet are mapped to the adjacent overlay nodes, so that the lookup time between them can be reduced significantly. In addition, two phase lookup algorithm and caching scheme are suggested to maximize the locality in P3ON. In our future work, we will extend P3ON to support large-scale P2P services as well as hierarchical mobile P2P services.**

I. Introduction

In structured peer-to-peer (P2P) overlay networks, a number of distributed hash table (DHT) based lookup algorithms have been extensively investigated, e.g., Chord [1], CAN [2], Pastry [3], and Tapestry [4]. Most of them provide the lookup time that is bounded approximately to $log(N)$, where N is the number of nodes. However, the hop distance between two virtual nodes in the overlay networks is not related to the real distance between two nodes. For example, two adjacent overlay nodes may be located in different continents. In this case, the number of lookup operations in P2P overlay networks is just one, while the real latency for lookup operations may be quite long. Especially in Chord that uses a ring topology as an overlay network, this problem becomes more serious because Chord provides only a uni-directional finger table for faster lookup operations [1].

In this work, we propose a proximity-based P2P overlay network (P3ON), which exploits two factors: proximity and locality, in order to allow a faster P2P service in large-scale networks. Recently, the proximity in P2P overlay networks has been investigated in the literature [5], [6]. However, how to assign an identifier (ID) to a node or item depending on the proximity has not extensively conducted in these works. We design a simple proximity-based ID assignment algorithm that is well consistent with the current Internet architecture. In our scheme, two nodes within the same subnet are mapped to the adjacent overlay nodes, so that the lookup time between them can be reduced significantly.

Of course, there is no evidence that there are more references between two neighbor nodes in the same subnet. Therefore, we propose a two-tier overlay ring architecture consisting high and low tier overlay rings. In the two-tier overlay ring, two phase lookup algorithm is used based on the locality and caching scheme. In this architecture, even though all nodes and items are mapped to the high-tier overlay ring, each subnet maintains another overlay ring (i.e. low-tier overlay ring). By using the two phase lookup algorithm, a node first tries to resolve its query within the low-tier overlay ring because each node in the same low-tier overlay ring guarantees a low latency lookup time. If the query cannot be resolved in the low-tier overlay ring, the node refers to the high-tier overlay ring. At the same time, the referenced item is cached at the low-tier overlay ring. Consequently, as the number of references to the same item increases, the performance of the two phase lookup algorithm is also improved dramatically.

II. Proximity-based P2P Overlay Networks (P3ON)

If the hop distance between overlay nodes is based on the proximity (i.e. geographical distance) between two nodes and there is a locality in the lookup procedures, the lookup time in large scale P2P overlay networks can be substantially lowered. To accomplish this, we design two algorithms: *proximity-based ID assignment* and *two phase lookup algorithm.*

A. Proximity-based ID Assignment Algorithm

In P3ON, an ID is assigned to each overlay node by considering the proximity. To achieve the proximity in overlay networks, we use the hierarchy characteristics of IP address. A node's ID is generated using its IP address and a collision-resistant hashing function such as SHA-1 [7] as follows.

$$\text{Node}'\text{s ID} = f(\text{node}'\text{s network prefix}) || f(\text{node}'\text{s host id})$$

where f is the SHA-1 function.

For example, let's assume that a node's IP address is 147.46.216.56 and the network prefix is 147.46. If $f(147.46)$ = 327 and $f(216.56) = 731$, then the node's ID will be 327731. Using this algorithm, we can place two nodes in the same subnet to adjacent overlay nodes in P2P networks.

On the other hand, an item's ID is constructed as follows:

$$\text{Item}'\text{s ID} = f(\text{item}'\text{s name}) || f(\text{item}'\text{s name})$$

Therefore, if f(item's name) is equal to 200, the item's ID will be 200200. Using this ID, we place the item in the P3ON overlay ring.

Figure 1 shows the basic P3ON architecture. All nodes and items are mapped to one point in the *high-tier overlay ring (main ring)*. At the same time, nodes located in the same subnet make another virtual overlay ring, called *low-tier overlay ring (subnet ring)*. Based on these two tier overlay rings, P3ON achieves a faster lookup time, which will be elaborated in the next section.

B. Two Phase Lookup Algorithm

In P3ON, an item can be searched by two phase lookup algorithm. In the first phase, a node tries to find the item from the low-tier overlay ring (*Phase I*). To do this, the node uses the item's 160 bits ID that is derived from the item's name and hashing function (i.e. f(item's name)). As mentioned before, an item is initially located only in the high-tier overlay ring. Therefore, a lookup miss will occur for the first lookup searching the item in the low-tier overlay ring. Figure 1 demonstrates the first lookup procedure in the two phase lookup algorithm. The node id 327731 is the concatenation of hashing results of subnet prefix (i.e. 327) and host id (i.e. 731). In this example, the node 327731 tries to find an item 200. The item 200 appears as 200200 over the high-tier overlay ring and it is mapped to node 200198. At first, node 327731 finds the item 200 with the ID 327200 concatenating the hashed value of the node's network prefix ID and item's name. As similar method in Chord, the node 327176 is responsible for the item in the low-tier overlay ring. However, since the node 200198 does not belong to the same low-tier overlay ring, the node 327731 cannot find the item 327200 from the node 327176. In short, a lookup miss occurs at node 327176 (refer to (1)). Then, the node 327731 searches the item 200 in the high-tier overlay ring (*Phase II*) with the ID 200200 (i.e. the original ID of the item).

In the second lookup procedure, the 320 bits item ID (i.e. f(item's name) $\|$ f(item's name)) is used as a search key (refer to (2)). In P3ON, all items are mapped to one of nodes in the high-tier overlay ring, so that there is no lookup miss in the second lookup so far as the item is available over the entire networks (refer to (3)). After the second lookup procedure, the node 327176, where the lookup miss occurred, caches the location of the item 200200 (refer to (4)). Accordingly, the subsequent lookup for the item 327200 can be resolved by the node 327176.

The two phase lookup algorithm allows a faster lookup time, especially in large scale P2P overlay networks. If some nodes located in the same low-tier overlay ring or in the same subnet (e.g. nodes 327731 and 327700) have searched an item, other nodes can reduce the item lookup time for the same item by utilizing the cache information.

III. Conclusion and Future Works

In this work, we propose a proximity-based peer-to-peer overlay networks (P3ON). P3ON combines the proximity-based ID assignment and two phase lookup algorithm to provide fast and efficient lookup procedures. A node can obtain the query result immediately from the low-tier overlay ring if the same query has existed before in the low-tier ring. This is because query results are cached at the low-tier ring. Furthermore, since the low-tier ring simply replicates the information of the high-tier ring, load balancing and fault tolerance supported in the high-tier ring are also achieved in the low-tier ring.

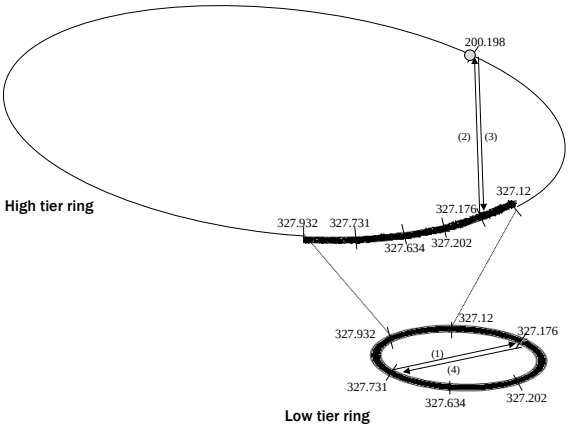

Fig. 1. Two-tier overlay ring architecture in P3ON

In our future work, we will endeavor to address the following issues in P3ON. Also, we will implement the prototype for P3ON and evaluate its performance by extensive measurements and simulations.

- Scalability in P3ON
- Hierarchical mobility support in P3ON
- Optimization of lookup procedures in P3ON

Acknowledgment

This work was supported in part by the Brain Korea 21 project of the Ministry of Education and in part by the National Research Laboratory project of the Ministry of Science and Technology, 2005, Korea.

References

[1] I. Stoica, R. Morris, D. Karger, M. F. Kaashoek, and H. Balakrishnan, "Chord: A scalable peer-to-peer lookup service for internet applications," in *Proc. SIGCOMM 2001*, 2001.

[2] Q. Lv, S. Ratnasamy, and S. Shenker, "Can heterogeneity make Gnutella scalable," in *Proc. IPTPS 2002*, Mar. 2002.

[3] A. Rowstron and P. Druschel, "Pastry: Scalable, distributed object location and routing for large-scale peer-to-peer systems," in *Proc. Middleware 2001*, 2001.

[4] B. Zhao, J. Kubiatowicz, and A. Joseph, "Tapestry: An infrastructure for fault-resilient wide-area location and routing," *Technical Report UCB-CSD-01-1141*, U. C. Berkeley, April 2001.

[5] M. Castro, P. Druschel, Y. C. Hu and A. Rowstron, "Proximity neighbor selection in tree-based structured peer-to-peer overlays," *Technical report MSR-TR-2003-52*, Microsoft, 2003.

[6] E. Lua, J. Crowcroft, and M. Pias, "Highways: Proximity Clustering for Massively Scaleable Peer-to-Peer Network Routing," in *Proc. IEEE P2P 2004*, August 2004.

[7] National Institute of Standards, FIPS 180-1, "Secure hash standard," April 1995.