Research and Improvement Based On Chord Protocol

MA Xiaojun, WANG Yujian*, FAN Lili
Institute of Information Technology
Beijing Union University
Beijing, China

LIU Chen, WANG Zhanmin
College of Computer Science & Technology
Beijing University of Posts and Telecommunications
Beijing, China

***Abstract*—Distributed hash tables resource locating arithmetic in peer to peer network will lead the overlay topology is not congruent with the physical topology to a large extent. In this paper, an improved method taking advantage of structured node ID is promoted based on the serious research of Chord protocol. The experiments show the improved method can make the neighbouring nodes in physical topology approach as far as possible.**

Keywords-peer to peer network; distributed hash tables; Chord; overlay network

I. INTRODUCTION

Peer to peer (P2P) networks are the base of whole architecture on the Internet. There are no client and server concepts in the most basic TCP/IP protocol. During communication process, all the equipments are of equal importance. The distributional retrieval and routing algorithms based on distributed hash tables (DHT) owing to have advantages of determinability, simplicity and distribution in searching, are becoming the international hotspot in the research and application of structured P2P networks. Distributed hash tables can exhibit good robustness in the face of node failure, suffering from attack and sudden high load, which have good expansibility; can acquire large scale of the system with low system overhead; can be self-configuring and merge the new nodes into the system automatically without manual intervention; can provide simple and flexible interface for multiple P2P applications at the same time[1]. With regards to this, many research teams in the design of the extended mechanism in searching have done a lot of research work in recent years, and put forward Chord, Pastry, CAN and Tapestry etc for constructing structured P2P Distributed Hash Table systems[2][3].

II. CHORD PROTOCOL INTRODUCTION

A. Chord Protocol Brief Description

Chord Protocol is a kind of P2P computing protocol, and it provides a new method to solve the problem of positioning resources effectively. As a rule, Chord Protocol implements such an operation: given a keyword key (usually filename, etc.), it can map the key into some node in the system. Data positioning can be achieved by associating key with data items. The identifier of keyword is derived by hashing keyword, using a hash function. The identifier of a node is derived by hashing its IP address, using a hash function[4]. If each data item is assigned a key in peer-to-peer network application, the data lookup problem can be solved easily by using Chord in peer-to-peer network. Chord protocol does not require every node knows all the other nodes, which promotes the expansibility of consistent Hash operation. Each node only needs to store little routing information about other nodes. Every time when nodes join or leave for the system, Chord needs to update routing information.

B. Consistent Hash

In Chord protocol each node and keyword is assigned a m-bit (binary system) identifier by using SHA-1 as consistent hash function[5]. The identifier of a node is derived by hashing its IP address, and the identifier of keyword is derived by hashing keyword (usually filenames). Consistent hash assigns the identifier for each node as explained below. The nodes and keywords are arranged in the identifier ring mod 2^m that is called the Chord ring. The range of identifiers in Chord ring is 0 to 2^m-1.If k is the identifier of a key, it will be assigned to the node whose identifier is equal to k or the first node after the k. The node is becoming subsequent node of k, expressed as successor(k).

If the identifier is expressed as a ring that is from 0 to 2^m-1, the successor(k) is the first node from k in a clockwise. For the identifier of keywords "Happy Birthday" is 24, it is kept in the node N32, as is shown in Figure 1.

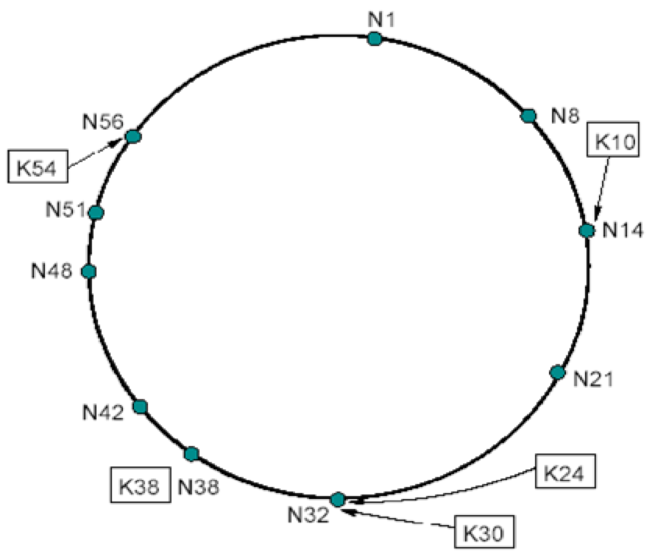

Figure 1. Chord ring with 10 nodes

*Corresponding author: xxtyujian@buu.edu.cn
This work is supported by the digital technology research team (#PHR200907120).

Figure 1 shows a Chord ring whose m is 6. This circle has 10 nodes and stores five keywords. Successor node of Keyword K10 is node N14, so resource object that K10 represents will be stored in the node N14. Likewise, resource object that keyword K24 and K30 represent will be stored in the node N32, resource object that keyword K38 represents will be stored in the node K38, resource object that keyword K54 represents will be stored in the node N56.

C. Routingtable Finger Table

Suppose keyword/node identifier space is m-bit (binary system). Each node n maintains a routing table which mostly contains m records, the routing table is call Finger Table.

In the Finger Table of node n, No. i record contains the identifier of node s. The identifier position of node s is at least 2^i-1 after node n in identifier ring, namely:

$$s = \text{successor}(n+2^i-1) \qquad (1 < i < m)$$

All the calculations will have $(n+2^i-1) \bmod 2^m$. So node s after 2^m -1 is called the routing table (finger) of node n, expressed as *n.finger[i].node*. Every routing table item in Finger Table contains the Chord identifier, IP address and port number of correlative node.

It is an easy observation that the first finger of node n is the direct subsequent node of node n in Chord ring, for convenience, and the direct subsequent node is called successor node and is not called the first routing finger again. The process of looking up key in node k is as follows: if key falls over between the k and successor(k), successor (k) node should have some information about key, and can send inquiry requests to successor (k)->IP node. If key is not between k and successor(k), the routing table item of finger table need to be searched. Find the one that start domain is the closest to key and start is less than the key, and then transmit inquiry requests to IP address of the routing table item and asked it to go on finishing search process. It can be seen in this strategy that every inquiry makes the distance of getting to destination node reduce by over half. Therefore, the average number of searching is $\log_2 N$, and search efficiency is very high.

D. Join and Stabilization of Nodes

Chord structure is able to process the situation that node changes primly. When a new node r wants to join the Chord network, it must connect with a known Chord node, and request to look up the IP address of successor (r). Then, the new node requests successor(r) to look up its precursor node. Finally, the two nodes that the new node request to find update its record respectively, which will insert the new node between them. In this way r enters into the ring of Chord. Each node after joining network does not only initialize its pointer of Finger Table, but also informs all the related nodes to update their Finger table. Obviously, at some point, the finger table of node will follow outdated information. But this outdated information still remains the complete structure, and information that does not update timely only reduce search efficiency, and will not cause obstacles for the overall operation of network.

When a node leaves normally, it will transfer the hash table that is stored in it to its successor node, and notify the precursor node that it is going to leave. In this way, the precursor node can update the successor node, which makes it point to the successor of node that is about to leave. When a node collapses, it is unable to complete these steps during the normal exit, which will cause precursor node to point to successor node effectively no longer. When inquiry request is passing through this node, error will happen. In order to relieve this problem, every node does not only contact direct successor node, but also contact all direct successor nodes in x steps. So even if x-1 consecutive nodes behind it fail simultaneously, it also can skip them and contact No. x successor node, and go on to keep Chord ring structure.

III. RESEARCH ON STRUCTURED NODE ID

A. Reason of Using Structured Node ID

In the existing P2P routing algorithms, node chooses next routing jump based on node ID, node ID is generated by hash algorithms randomly[6]. The path of P2P routing is composed of application-level routing jump between initial node and destination node, rather than IP routing jump. Actual P2P routing efficiency is measured by end-to-end routing delay. Due to the peer-to-peer node is distributive in geography, some application level routing jumps traversing different autonomous system, and other routing jumps are from one node to another node in the same LAN[7]. Because P2P routing algorithms ignore the delay between nodes, often choose the routing jump with long delay.

B. Design Proposal of Structured Node ID

Using the method of structural node ID can effectively embed physical location and network distance into the identifier space, and make logical topology and physical topology consistent, so can contribute to implement local routing. Structured ID space is constructed by using hierarchical and location-based node ID assignment proposal. The node ID based on the location consists of two parts: first, the prefix that is assigned to a region and used to mark; second, the suffix that is made up of bit sequences generated randomly. Assigning the prefix by layering, can embed network topology information into the logical structure of DHT, and can be used in the routing process. Proximity relation of different region in the identifier ring reflects their neighboring in actual geographical position, similarly, the distance of different region in logic overlay network reflects the distance in physical network. When routing is running along node identifier space, the same process will also happen in physical space. Logic path overhead can be limited in the multiples of small constant factor of network path overhead, so it can get O(1) complexity.

Design proposal of structured node ID is as follows:

(1) From the point of view of administrative districts, design proposal of structured node ID using layering and prefix. As is shown in Figure 2(a).

(2)From the point of view of the geographical spatial division directly, design proposal of structured node ID using layering and prefix. As is shown in Figure 2(b).

Province mark	City mark	District and county mark	Random mark

(a)

Area mark	Child area mark	Leaf area mark	Random mark

(b)

Figure 2. Proposal of structured Node ID Using Layering and Prefix

Structured node ID also has four parts: the first three parts is prefixes that represents area that node locates. The last part that is random bit string is suffixes, is used to maintain load balance. While node ID that basic Chord protocol proposes only has random identifier part in Figure 2. Assignment on prefix mark, can divide according to the geographical position relations: using a rectangular to cover a planar region, the rectangular is divided into four small rectangles by quaternary tree, then these small rectangles is divided circularly until all the small divided rectangles can include the nodes in the same subnet. This design proposal mainly takes consideration to neighborhood in geography position, and divides an area using Hilbert Curves. The Sample is shown in Figure 3.

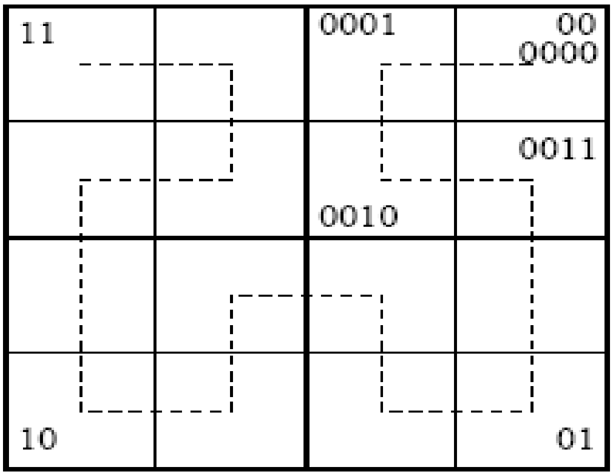

Figure 3. Hilbert Curves(dashed lines)

The logical distance between nodes is reflected by numerical difference between nodes ID. At this time, prefix part of nodes ID need to be matched firstly in the routing process. Secondly, the identifier part of node ID (node ID low part) need to be matched .Prefix perch part of two node ID matches more bits , and two node ID is closer. For basic Chord protocol, the process of routing is going forward in a clockwise in the identifier ring. But for the routing taking prefix as center, the process of routing is matching prefix firstly, and secondly comparing node identifier.

Using a Hilbert Curve across a planar area, which would divide this area into several smaller area, and assigning continuous area marks to small neighboring area, the area marks can consist of a one-dimensional numerical ring. The spatial order of small area will be continuously kept in digital space, and the distance between the area marks reflects the physical distance.

IV. EXPERIMENTAL RESULTS ANALYSIS

In order to test what effects on Chord network in using design proposal of structured ID, a lot of experiments have been done in our research, using SChord simulator and iterative Chord simulator[8]. Using both under the circumstances of same starting condition, the experimental data was recorded separately.

The stretch rate of overlay network path is the ratio of path length of logic overlay network to the shortest network distance between source node and destination node. Under the two circumstances of using random node ID and structured node ID separately, the comparison of stretch rate is shown in Figure 4.

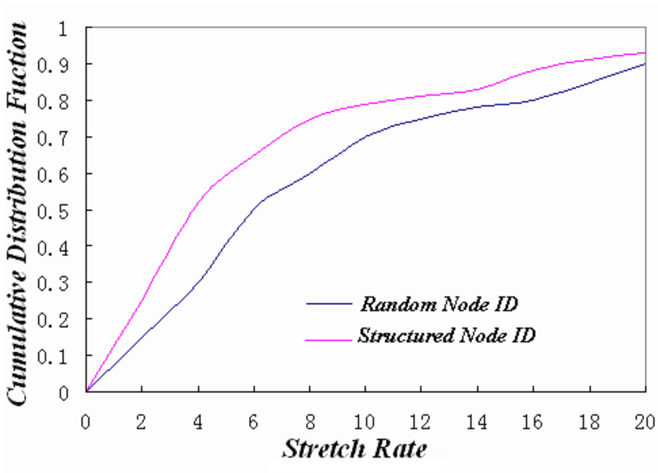

Figure 4. Two Methods in Comparison of Stretch Rate

As can be seen from Figure 4, for the same stretch rate, such as 4, the probability of using structured node ID is greater. Under the circumstances of same probability, such as 0.3, the stretch rate of using structured node ID is smaller. At this time, the path length of overlay network approaches the path length of physical network more.

V. CONCLUSION

We studies on the sizeable inconsistent phenomenon between logic overlay network based on Chord protocol and lower physical network, and puts forward the improved proposal using structured node ID. Experimental results show that the improved proposal using structured node ID can get smaller stretch rate, i.e. the ratio of lookup length of logic overlay network to the shortest network distance between source node and destination node is smaller.

ACKNOWLEDGMENT

The authors would like to thank Education Committee of Beijing of China. This work is supported by the science-technology development plan of Education Committee of

Beijing (#KM200811417001), digital technology research team (#PHR200907120), and the project of backbone of middle-aged and young teachers.

REFERENCES

[1] RISSON J, MOORS T. Survey of research towards robust peer-to-peer networks: search methods[J]. Computer Networks: The International Journal of Computer and Telecommunications Networking, 2006, 50(17): 3485-3521

[2] RAMASWAMY L, GEDIK B, LIU L. A distributed approach to node clustering in decentralized Per-to-Peer networks, parallel and distributed systems[J]. IEEE Transactions, 2005, 16(9):814-829

[3] Cohen E, Fiat A, Kaplan H. Associative search in peer to peer networks: harnessing latent semantics [J]. Computer Networks, 2007, 51(8): 1861-1881

[4] Wang Jing, Yang Shoubao, Guo Leitao. A Bidirectional Query Chord System Based on Latency-sensitivity[C] // Proc of the 5th International Conference on Grid and Cooperative Computing. Changsha, China: IEEE Press, 2006

[5] Yu S, Yu J, and Kamil K, et al. DR-Chord-FAn efficient double-ring Chord protocol [C]. 6th IEEE International Conference on Grid and Cooperative Computing Proceedings, Urumchi, Xinjiang, China, 2007: 16-18

[6] Huang C, Hsu T, and Hsu M. Network-aware P2P file sharing[J]. IEEE Journal on Selected Areas in Communications, 2007, 25(1): 204-210

[7] Wang J Z, Bhulawala V. Design and implementation of a P2P cooperative proxy cache system [J]. Journal of lnterconnection Networks, 2007, 8(2): 147-162

[8] Hong F, Li M, Yu J, et a1. PChord: improvement on chord to achieve better routing efficiency by exploiting proximity[C]. In Proceedings of the 25th IEEE International Conference on Distributed Computing Systems Workshops (ICDCSW05), June 2005