A Distributed Route Structure Based on IPv6 in P2P System

YU Jian
School of Electronic and Information Engineering,
School of Computer Science and Technology,
Tianjin University,
Tianjin, 300072, China
yujian@tju.edu.cn

YU Mei
School of Computer Science and Technology,
Tianjin University,
Tianjin, 300072, China
yumei@tju.edu.cn

XUE Yifeng
School of Software,
Beijing University of Aeronautics &Astronautics,
Beijing，100083，China
xef@163.com

DAI Jufeng
School of Electronic and Information Engineering,
Tianjin University,
Tianjin, 300072, China
jfd@tju.edu.cn

Abstract:—**A hierarchy distributed route structure based on IPv6 address in Peer to Peer Systems is proposed in order to solve the problems that physical topology and logical network does not match so as to inefficient search in Distributed Hash Table at present. By building node identifier using IPv6 address and hashing different level clustering identifier to set up hierarchy route, near nodes in physical network are closer in overlay. Smaller hops across domain, lower search latency, and higher efficiency are proved.**

Keywords-Peer to Peer(P2P); Distributed Hash Table(DHT); Chord; IPv6; Hierarchy Route

I. INTRODUCTION

In a typical P2P network, the data resource distributing in each independent nodes, how to efficiently index, search, locate and access them is an important issue. The latest result is the distributed search and routing algorithm based on DHT. In the application layer, DHT organized all P2P nodes into a structured overlay network which file indexes in it and enquiry messages will route through it. By distributed hash function DHT only maps the entered keywords into a certain nodes in the overlap, and then connect with the node through certain routing algorithm. This type of P2P network topology models typically are Chord[1], Pastry[2], CAN[3], and Tapestry[4].

At present, the problem in DHT is overlooking the vicinity of nodes in the physical network, that is, the adjacent nodes in overlap maybe far apart in the actual physical network, leading to the disjoint between the logical network and the physical network and resulting in the lower search efficiency [5-9].

To solve the question of the overlay not matching the physical network in the existing structured P2P system, using the clustering of IPv6 address, it is proposed that the hierarchy node identifier is constructed using IPv6 address in Chord, so that the nodes are adjacent in both the physical and logical network, and the topology and logic network are effectively anastomosed, which the lookup delay can be reduced well.

II. REVIEW OF CHORD

Chord is a distributed search algorithm which is proposed by UC Berkeley and MIT and for searching data in P2P network. Given a keyword, Chord can effectively map the keyword to a certain node in the network. Therefore, so long as in the P2P networks each data item V is given a keyword K, the (K,V) is stored or accessed on the node which the keyword mapped on by Chord. Each node and keyword in Chord has a m-bit identifiers. The keyword identifier K is given by hashing the keyword itself, and node identifier N by IP address. Hash functions can be SHA-1. According to its identifiers (modulus 2^m) ascending, all nodes are arranged clockwise on a logical ring which is called Chord Ring. Mapping rules in Chord: the (K, V) is stored on the node which the node identifier is equal to K or is the next node on Chord ring; this node is known as successor (K). In Chord, each node need store m other nodes information, which is called Finger Table. In the table the nodes is not directly adjacent nodes, which the distance between them will be 2^i (i is the subscript of item in the finger table).

In the course of enquiries the requests are sent to the node of the nearest key by enquiry node. If the node received requests find the information which enquired, it can response to enquiry node; else transmit the request to the node with nearest keys in the finger table. This process has not continued until to find the corresponding nodes. With N nodes in the P2P networks, the hops need to be $O(\log_2 N)$ and each node only store ($\log_2 N$) other nodes of information.

Chord failed to take full advantage of the underlying physical network topology information. DHT may be to map a local data to the node with farther distance from the local nodes in physical network. In key lookup on overlap the neighbor nodes in the physical network may also be cross domain. As the main factors for search delay is the cross-domain time and time in domain is relatively small, the actual lookup is inefficiency, and the cost of resources locating is relatively large.

In the design of IPv6, the routing aggregation can be done by designing the hierarchical address structure and address distribution [10-12]. Network prefix distribution in IPv6 address reflects the strong level attribute. The same prefix length of the nodes is in the same sub-domain. Routing constructed by IPv6 addresses can make the nodes in the same sub-domain or be close neighbors near each other in the overlap, so that the inter-domain enquiries can be reduced and the lookup efficiency can be improved.

III. CHORD BASED ON IPV6

A. *Aggregation in IPv6 Address*

Unicast address identifies an independent IPv6 interface. A node may have multiple IPv6 network interface. Each interface must have an associated unicast address. Unicast address may be considered to contain some information which included in the 128-bit field. This address can completely definite a specific interface. In addition, the address contains a number of part information. When all the information is put together, 128-bit addresses identifying a node interface will be constituted.

1) *Unicast Address in IPv6*

Every IPv6 unicast address can be seen as two fields, one field is used to identify the network, and the other fields are used to identify node interface on the network, that is interface identifier and subnet prefix, As shown in Figure 1. Interface identifier length depends on the subnet prefix length. The length of them may change.

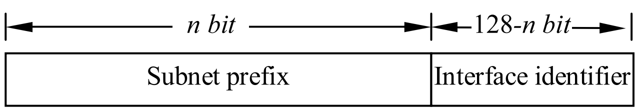

Figure 1. Unicast address format in IPv6

2) *Aggregatable Global Unicast Address in IPv6*

Aggregatable global unicast address is a type of clustering, which is independent of the ISP. Aggregatable address based on the supplier must be changed with vendors, however the address based on exchange station is located by IPv6 switch entities directly. The address provided by the exchange station, and users and providers signed a contract for network access. This kind of access to a network is provided directly from the provider or indirectly from the exchange offer, but routing through the exchange station. It enables users change ISP without re-site. At the same time users are also allowed to use multiple ISP to deal with the monolithic network address.

Aggregatable global unicast address includes address format which the initial three bits are 001 (This format can be used for other undistributive unicast prefix in the future). Address format is shown in Figure 2.

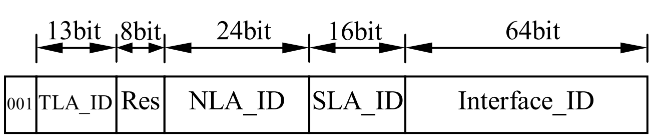

Figure 2. Aggregatable global unicast address format in IPv6

FP field: format prefix of an IPv6 address, three-bit length, used to identify the address type. At present the field is set as "001", identifies aggregatable global unicast address.

TLA_ID field: Top aggregation identifiers, including the most high-level address routing information. This refers to the largest routing information of network interconnection. At present, the field is 13-bit, it will be the largest of the 8,192 different top routing.

RES field: 8-bit length, reserved for use in future. Eventually it may be used to extend top or next level aggregation identifier field.

NLA_ID field: Next-level aggregation identifier, 24-bit length. The identifier is used to control the top aggregation to arrange address space by institutions. These institutions (which may include large-scale ISP and other agencies providing public access to network) can address the 24-bit field in accordance with their own hierarchical structure.

SLA_ID fields: Site-level aggregation identifiers, used for internal network structure by some agencies. Each agency can create their own internal hierarchical network structure using the same method with IPv4. If all 16-bit field is used for flat address space, there will be up to 65,535 different subnets. If the former 8-bit is used for the more high-level routing within the organization, there will be 255 high-level subnets, and each high-level subnet can be as many as 255 sub-subnet.

Interface identifier field: 64-bit length, including the value of the 64-bit interface identifier in IEEE EUI-64.

B. *Chord Based on IPv6 Address*

1) *Node Identifier Structure*

In Chord the node identifier is built by hashing its IPv4 addresses, however the randomness of hash function makes physically adjacent nodes in the far space. IPv6 addresses can be used as a node addresses, and the IPv6 network address prefixes and the interface identifier are respectively hashed to construct the node identifiers by these two hash values. In the network prefix for hashing, the top aggregation identifier (including FP, TLA_ID and Res), next-level aggregation identifier NLA_ID and site-level aggregation identifiers SLA_ID are hashed respectively, these three parts of the hash value forms a network prefix. It can make the same prefix network nodes mapped to the adjacent space in logical network, but also make the same level aggregation node identifier more near, so as to achieve logical network and physical topology of the space effectively match.

In Chord identifier space size is 2^{160}, that is 160-bit node identifier length. n1-bit can be used in the hash value of top aggregation identifiers (including FP, TLA_ID and Res, 24 bits), n2-bit is used to identify next-level aggregation identifier (NLA_ID, 24 bits), and n3-bit does site-level aggregation identifiers (SLA_ID, 16 bits), i.e. n-bit (n = n1 + n2 + n3)used on the 64-bit network prefix in IPv6 address, with 160 - n bit identifies interface identifier. Under the uniqueness of the IPv6 address and the the nature of hash function, the uniqueness of the node identifier can be ensured. the node identifier format in hashing IPv6 address is shown in Figure 3.

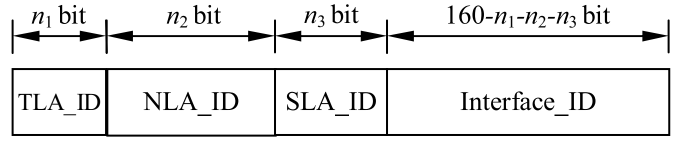

Figure 3. Node identifier structure in IPv6

Figure 4 is Chord ring by hashing IPv6 address. Obviously, the nodes with the same prefix network were mapped to the close distance position in logical network, and this kind of proximity reflected in different sub-domain. The nodes with the same pattern are in the same sub-domain.

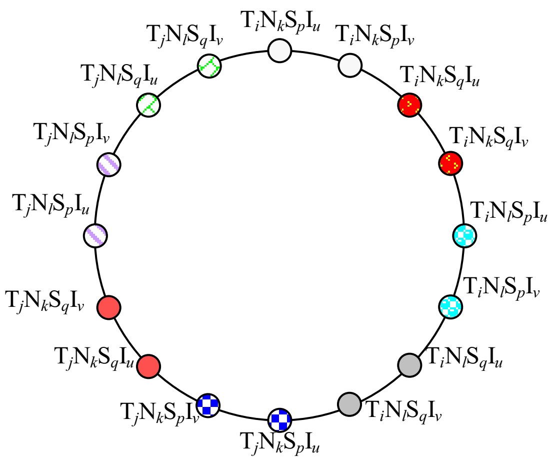

Figure 4. Chord based on IPv6 address

2) Hierarchy in Chord Based on IPv6

Figure 4 shows that the nodes with the same network prefix are in the near distance in Chord ring. Figure 3 also shows that the node identifiers prefix is divided into three parts, which are three levels of aggregation identifiers. After the three parts are respectively hashed, the node region can be divided into three different levels sub-regions, namely, the same TLA_ID nodes will be gathered together, the different TLA_ID nodes forms into different region TLA. In the same TLA region the nodes with the same NLA_ID will together form of the sub-region NLA. Similarly in the same NLA region the nodes with the same SLA_ID will together form the different sub-region SLA. In a large area would be divided into the small area at different levels, making a level of aggregation and more refined to better nodes aggregation. Figure 5 describes the levels of nodes aggregation. Nodes in the lowest level have the same network prefix.

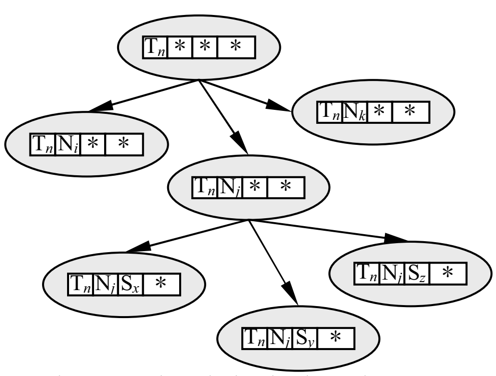

Figure 5. Hierarchy in Chord Based on IPv6

3) Routing Structure in Chord Based on IPv6

a) finger table

In order to achieve the Chord ring based on IPv6 address with hierarchical structure, and make full use of location information in the node identifier, and to the extent possible to reduce cross-domain operations when searching, for lower latency, finger table in Chord need to be established, table structure as shown in TABLE 1.

In the finger table of Chord ring based on IPv6 address, original Chord routing is remained. Original routing changes to be the local routing in the SLA sub-region which node is in, that is, the nodes with the same network prefix constitute a local Chord ring. In new finger table the precursor and successor nodes in three different levels has been added, so that when the searched node is not in the local sub-region routing will hop to another to continue.

TABLE 1. FINGER TABLE OF NODE N

Symbol	Definition
Finger[i]	successor（（$n+2^{i-1}$）mod 2^m），$1\leq i\leq m$，pointer to the first successor of node $(n+2^{i-1})\text{mod } 2^m$
successor	Finger[1].node，the successor of node n
predecessor	the predecessor of node n
SLA_successor	the first node of next SLA_ID sub-domain
SLA_predecessor	the last node of prior SLA_ID sub-domain
NLA_successor	the first node of next NLA_ID sub-domain
NLA_predecessor	the last node of prior NLA_ID sub-domain
TLA_successor	the first node of next TLA_ID sub-domain
TLA_predecessor	the last node of prior TLA_ID sub-domain

b) Routing

When node n will search keyword K, search process as follow:

(1) if n=K, then return node n and end.

(2) if n<K<=n.Finger[1].node, then hop to n.Finger[1].node and return.

(3) if n.predecessor<=K<n, then return node n and end.

(4) if SLA_predecessor<K<SLA_successor, then hop to the nearest node to node n according to Chord algorithm, search will go on in local region.

(5) if K< TLA_predecessor or K>TLA_successor, then hop to TLA_predecessor or TLA_successor, search will go on in another TLA sub-region and turn to (1).

(6) if K<NLA_predecessor or K>NLA_successor, then hop to NLA_predecessor or NLA_successor, search will go on in another NLA sub-region and turn to (1).

(7) if K<SLA_predecessor or K>SLA_successor, then hop to SLA_predecessor or SLA_successor, search will go on in another NLA sub-region and turn to (1).

Search steps: if the searched node is in the local SLA sub-region, then search will achieve in the local region in term of Chord routing algorithm. Otherwise if the keyword K and node n are in different TLA sub-regions, NLA sub-regions, SLA sub-regions, then hop to the corresponding sub-region to search. This kind of enquiries process not only can reduce the sub-region hops, but also shorten the distance between inter-region hops because of adjacent sub-region also near in logical network.

c) node join and depart

When new nodes from the node n insert, first, hashing different parts of its IPv6 address to form node identifier, and then, searching the routing of the known node n in order to determine the location of the new node, that is, the TLA sub-region, the NLA sub-region and SLA sub-region. In accordance with its precursor and successor node routing information, new node routing table will be established. If the new node is the first node, the last node, or the only node in sub-region, neighbor sub-region routing information need to update. When the node leaving, its precursor and successor node routing need to update. If the leaving node is the first node, the last node, or the only node in sub-region, neighbor sub-region routing information also need to update.

IV. Performance Analysis

By simulation experiment performances of Chord based on IPv4 and IPv6 are compared on the mean search hops and the mean search latency.

A. mean hops

Under the situation of different nodes, randomly generated the network of different sub-domain, each node meanly have 100 documents, use both address, the mean hops of each network are shown in Figure 6.

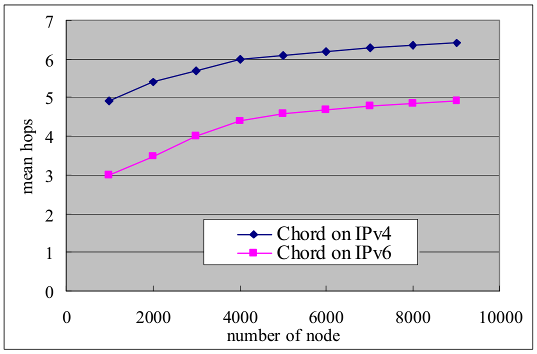

Figure 6. comparison on mean hops

Figure 6 shows that the mean hops of Chord based on IPv6 is lower than which based on IPv4. It is because that Chord uses IPv6 address to build node identifier with the hierarchical structure and considers the actual network structure, so that the physical and logical network are more matched.

B. mean latency

The mean end-to-end routing delay is a main performance parameter which measures actual locating in P2P system. The routing Delays in the network contain inner-domain routing delay and inter-domain routing delay, and the former is far smaller than the latter. Suppose the mean latency between two nodes in the same network is 20 ms, it is 80 ms in the different network. Figure 7 compares the mean latency on Chord based on IPv4 address and based on IPv6 address.

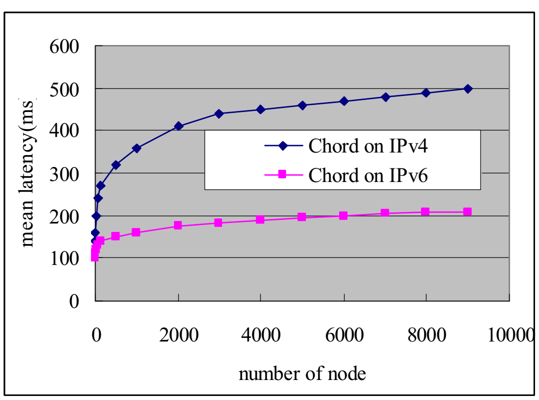

Figure 7. comparison on mean delay

It is showed in Figure 7 that as the Chord based on the IPv4 did not consider the physical and logical network structure matching on node identifier building, when lookup domain crossed may be more, routing latency will be far longer. For Chord based on IPv6 address considers the close to physical location and the logical position, the search distance on the overlap is approximate to physical network, and crossed domain number will be less, routing delay will be reduced.

V. Conclusions

By using the hierarchical clustering of the IPv6 address, the Chord ring with the level of the node identifier is built, which makes the proximity nodes in the physical network are also in the neighboring in logic ring, and this is also reflected in the vicinity of the different areas sub-domain. The search on this kind of Chord can be reduced on the logic hop between sub-domains, so as to greatly decrease time and effectively improve the query efficiency.

Acknowledgments

I am grateful to my mentor for his tutor, and thank the fellows for help. I can not complete the article without them.

References

[1]. Ion Stoica, Robert Morris, David Karger, M. Frans Kaashoek, Hari Balakrishnan, "Chord: A Scalable Peer-to-peer Lookup Service for Internet Applications," SIGCOMM'01, San Diego, California, USA, August 27-31, 2001.

[2]. Antony Rowstron1 and Peter Druschel, "Pastry: Scalable, decentralized object location and routing for large-scale peer-to-peer systems," Proc. of the 18th IFIP/ACM International Conference on Distributed Systems Platforms. Heidelberg, Germany, November 2001.

[3]. Sylvia Ratnasamy, Paul Francis, Mark Handley, Richard Karp, Scott Shenker, "A Scalable Content-Addressable Network," SIGCOMM'01, San Diego, California, USA, August 27-31, 2001.

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

[5]. Jiping Xiong, Youwei Zhang, Peilin Hong, et a1, "Chord6: IPv6 based topology-aware chord," http://www.iaria.org/conferences/ProgI/CASICNS05.html.

[6]. Xu Zhiyong, Min Rui and Hu Yiming, "HIERAS:A DHT-Based Hierarchical Peer-to-Peer Routing Algorithm," In Proceedings of the 2003 International Conference on Parallel Processing (ICPP'03), pp.187-194, Kaosiung, Taiwan, Oct., 2003.

[7]. Xiong Ji-peng, Guo Li-peng, Hong Pei-lin, Li Jin-sheng, "New Routing Performance Enhanced DHT Networks Based on IPv6," MINI-MICRO SYSTEMS, 2006, 27(8):1421-1425.

[8]. Prasanna Ganesan, Krishna Gummadi, Hector garcia-molina, "Canon in G Major: Designing DHTs with Hierarchical Structure," In:ICDCS, 2004.http://www.cs.washington.edu/homes/gummadi/papers/hierarchical-dhts.pdf, 263-272.

[9]. M. J. Freedman, M. VutUkuru, N. Feaxnster, "GeograPhic Locality of IP Prefix," Proc. ACM SIGCOMM Internet Measurement Conference, 2005.

[10]. Hinden R, Deering S, Nordmark E, "IPv6 Global Unicast Address Format," RFC3587, August 2003.

[11]. R. Hinden, S. Deering, "Internet Protocol Version6 (IPv6) Addressing Architecture," RFC3513, April 2003.

[12]. Deering S, Hinden R, "RFC 2460: Internet Protocol, Version 6(IPv6) Specification," http:/ / www. ietf. org / rfc / rfc2460. txt, Dec. 1998.