Research of Hierarchical P2P Network based on Chord

Haibo Ma, Deguang Wang, Jiamin Zhang, Li Shi
Dalian, Liaoning Province, China
Dalian Jiaotong University,
e-mail mhb101@126.com

Abstract: Structured P2P model with Chord method can be used to quickly search the location of resources, which is an effective P2P routing algorithm. This article combined the unstructured and structured P2P model, saying that Hierarchical P2P Network based on Chord as the model of the HPNC. The method is composing the nodes not far from them into a group with unstructured P2P model, the group constitutes the lower network; the central node from lower network abstracts into super-node, becomes the upper network node using structured Chord model, the network improves the efficiency of resources search. Building P2P network with HPNC is fast and simple, resources search fast.

Keywords: P2P; chord; DHT; HPNC

1 Introduction

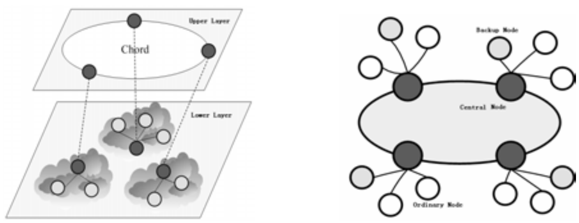
In the P2P network, Computing Model calculates from center to dispersion, from central to the edge, making full use of the processing capacity of the terminal device, while each node actively joins in the network to share resource[1]. P2P network has many kinds of model [2], like Napster, Gnutella, Chord, CAN, etc. Among which, the Chord[3] model is representative. Chord, as a typical and successful routing algorithm of the structured P2P, has become the hotspot in the present research. The Chord model uses the ring topology, and its protocol can map the appointed key words to the corresponding node. From the aspect of algorithm, Chord is a variant of the consensus Hash algorithm[4]. However, the Chord topology has a shortcoming that logical overlay network is completely separated from the topology of the actual underlying Physical layer. In addition, it has neglected the isomerism between the nodes. In order to enhance the running and maintenance efficiency of the P2P network, this paper has improved the P2P model, and putting forward the HPNC (Hierarchical P2P Network based on Chord).

2 HPNC Network Model Overview

2.1 HPNC topology

HPNC combines unstructured and structured features of P2P networks. It has two layers, a structured Chord network in upper layer, and a centralized unstructured network in lower layer, and then the integral structure constitutes a mixed 2 layers P2P networks.

Nodes are organized in a group according to the MAC similar first principle. Group is a small unstructured network, in which topology similar to the centralization of Napster is used, composed of a central node and many ordinary nodes, while the central node administrates the ordinary nodes and provides resource index function for the ordinary nodes in the group. All groups' central nodes constitute the upper Chord network in the form of super nodes, and realize accurate and effective Message routing mechanism based on the DHT[5]. The structure of HPNC is shown in Figure 1.



(a) Topology of HPNC network model

(b) Nodes classification of lowerlayer

Figure 1. HPNC network model

2.2 Nodes Structure of HPNC

The super node is a great important module of the HPNC model, charged by the central node in the lower layer, playing an important part in supporting the route and data transmission in the entire network. The super node is the link for the correspondence between the upper and lower layers, and the interaction between layers is realized by the super node. The super node, as the node of the upper

978-1-935068-01-3 © 2009 SciRes.

layer Chord network, needs to maintain three forms: pointer gauge, key value table, and predecessor and successor table. The pointer gauge is mainly used between super nodes in the upper layer for the news routing, and it will change only when the super nodes add or quit. The key value table stores all resource information saved on this super node, and it will vary with the resource information being added or deleted. Predecessor and successor table saves information of the predecessor and successor nodes of this super node.

Each ordinary node stores the basic information about the central node in order to communicate with it in time, so the table is called the central node table. In addition, each ordinary node has its own shared resources in the network, which are saved in the resource list.

The central node is a special ordinary node, therefore firstly it inherits all attributes of the ordinary node. Moreover, it also needs to maintain two tables for the group management, respectively the Node_table and the MessageIndex_table. The former is used to save basic information of nodes in the group, the latter is used to save all the resource information issued by the nodes in the group.

The basic function of the backup node has no difference from the ordinary node. However, since it also needs to be the substitute for the central node, besides the ordinary node attributes, it needs to have group internal node table and predecessor and successor table as the same as the super node. The backup node needs to periodically exchange the information of these two tables with the super node, in order to guarantee the accuracy and timeliness of the data.

2.3 Features of HPNC

Introduced the nodes grouping thought based on the physical location, the model will group the nodes according to the geographical position in order to group nodes close in physical space to be in a group. It fully uses the characteristics of the geographic proximity that the connection and access speed of nodes are fast, the resources inquiring latency is small, and resources downloading speed is quick. It has overcome the detour problem of the structured P2P system that universally exists, and increased the hit rate of the resources localization, reduced the expenses of the resources localization, and raised the efficiency of the resources localization.

The model has introduced the concept of the layered hybrid model. According to different degrees of ability, nodes are divided into ordinary nodes, central nodes, backup nodes, and super nodes. The upper layer takes the central node of the lower layer as the super node to construct Chord structured virtual network, and the lower layer consists of ordinary nodes and backup nodes, separately centralized managed by each group's central node. This not only fully uses the node computing power and memory power, but also alleviates huge network fluctuation caused by nodes adding or quitting on the Chord.

By the heartbeat induction mechanism, we can catch the node invalidation in time and process it rapidly, thus the fluctuation which the node sudden invalidation brought to the network will be effectively avoided.

With the new management mechanism and the key value issue mechanism, and by introducing special node to manage the network and check the key value issue, the manageability of the P2P network is strengthened, and the commercial value of the network is enhanced.

3 Network Maintenance of HPNC

3.1 Node ID Conversation

In the network, each node has its own ID, and different ID has different function. The node’s behaviors like adding, quitting, or performance updating etc. can make the node ID converse. Its conversing process is shown in Figure 2. If the dynamic variation of the nodes cannot be handled in time, the network fluctuation or even collapse will be very likely caused. Therefore, effective maintaining nodes' dynamic change, and real-time dealing with ID conversion between the nodes are necessary means to guarantee normal operation of the P2P network.

1. The ID conversing process is composed of some basic behaviors, like the ordinary node adds, exits, fails, and usurps, the backup node abdicates, fails, and usurps, and the central node abdicates, fails etc.

2. The ordinary node and the backup node can converse ID, and the backup node and the central node can switch ID. But the ordinary node and the central node cannot converse ID directly, the conversation between them can be completed through transforming to the backup node.

3. The concept of abdicate and exit must be distinguished: Abdicate refers to lose the existing status,

namely the status degrades, but this node is still in the network. Such as the backup node abdicates to the ordinary node, although this node no longer has the function of a backup node, but still exists in the group as an ordinary node. But exit is that the node must quit from the whole network. The ordinary node can exit directly, but the backup node must abdicate to an ordinary node first, then exit the network. In the same way, if the central node wants to exit, it must abdicates to the backup node first, then the ordinary node, and then exit.

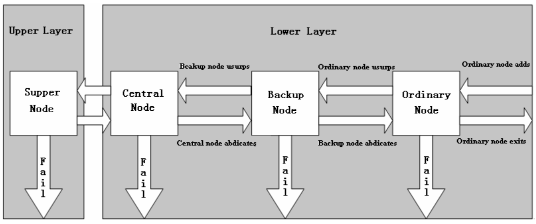


Figure 2. Nodes ID Conversation

3.2 Group members adjustment

New nodes added to the network may have high performance, so when they join, the central node will call the group member to adjust function module in order to arrange the group member's performance. The node performance is not irrevocable, in different times or certain circumstances, the node performance will also change greatly. Therefore, when the node performance parameter changes, it will inform the central node to update, then the central node will call the group member to adjust the module in order to adjust the member ID. When the central node finds that an ordinary node's performance is higher than the backup one, it will replace the backup node with the ordinary one. For the same reason, if the backup node's performance is higher than the central node, the two will exchange. Certainly, a new ordinary node's performance may extremely superior, so it will be replaced with a backup node first, then with the central node. Flow chart is shown in Figure 3.

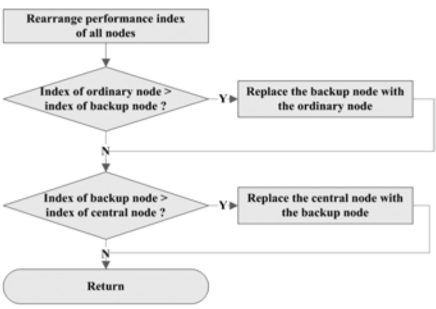


Figure 3. Adjustments of Group Members

4 Service of HPNC

4.1 Issue Service of Resources Information

The resource information is an information pair which is composed of the resource key words and IP address of the issue source.

Any node in network has the resource information issue function. The issue of resource information is composed of introgroup issue and extragroup issue. The introgroup issue is that the central node where the issue source is adds the resource information to MessageIndex_table for the group member s' inquiring. Extragroup issue refers to that the central node converses the resource information to the identifier (K, N) through the Hash Algorithm, then issues to the super node named N=Successor(K) according to the Chord principle. For example, a node whose IP address is 198.10.10.1 issues resource information that the key word is ring, shown in Figure 4:

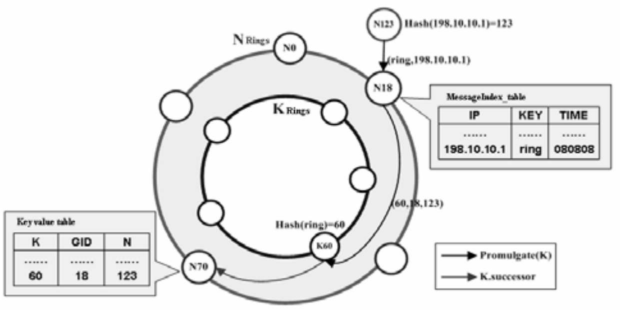


Figure 4. Issue service of Resource Information

The node whose IP address is 198.10.10.1 will be mapped as node N123 through the Hash Algorithm, if it acts as an ordinary node in a group which takes node N18 as the central node. If N123 want to issue resource information that the key word is ring, firstly N123 transmits information pair (ring, 198.10.10.1) to N18, then in the group N18 inserts this information and the issue time into the MessageIndex_table, like this the resource information introgroup issue is completed. Then N18 turns the information pair to identifier (60, 18, 123) through Hash Algorithm, while Hash (ring) = 60, 18 is the group mark where the issue source is, and 123 is the node mark of the issue source. The K60 successor node is N70, so N18 transmits identifier (60, 18, 123) to N70, and N70 inserts this data into its own key value table, thus the resources information extragroup issue is completed.

978-1-935068-01-3 © 2009 SciRes.

4.2 Deletion Service of Resource Information

Any node in network has the resource information deletion function. The resource information deletion is composed of introgroup deletion and extragroup deletion. According to the node deletion request, introgroup issue will delete the corresponding information in MessageIndex_table. Extragroup deletion is that the central node informs the super node which issues that this will delete the resource information, and lets it delete this resource information in the key value table.

When some node quits or fails, the central node will batch delete all information which this node once issued.

4.3 Inquiry Service of Resource Information

Any node that needs to inquire resource must send request to the central node, and the inquiry request is consisted of key words which the node searches for. The inquiry process is mainly divided into three steps. First is to inquire in MessageIndex_table. If there is no result, Hash the key words to get the key words' identifier, then inquire it in the key value table. If still not found, search it in the upper layer through the pointer gauge.

When inquiring in the upper layer, if the key word inquired is some K[i].start in the pointer gauge, then return to its K[i].successor directly. However, the number of routing information that the pointer gauge saves is only m, in other words, the key words that the node needs to inquire may not in the K[i].start item in the pointer gauge. In this case, we need to observe which K[i].interval sector the key words identifier has fallen into in the pointer gauge. Then forward the inquiry message to the corresponding K[i].successor in this sector, and continue to search by the K[i].successor node through the pointer gauge.

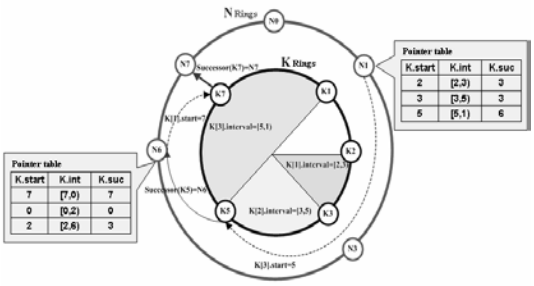


Figure 5. Inquiry service of Resources Information

For example, some sub-node Ni of the super node N1 wants to inquire resource that K is 7, but there is no result in MessageIndex_table and the key value table. Therefore, N1 searches in its own pointer gauge, and has discovered that K7 is in K[3].interval=[5,1) sector, so it forwards the inquiry request to K[3].successor, namely N6 node. N6 searches in its pointer gauge and has discovered K[1].start=7, the key word that has been searched for, so it returns the information of N7 which is the successor node of K7 to Ni. Ni will know the information of the source issue node of the resource by visiting N7.

5 Performance Assessment of HPNC

5.1 Memory Price of Data

In system memory price mainly includes two aspects: pointer gauge and the number of information backup. Suppose that Chord and HPNC have the same number of nodes, expressing the node number in system by Psystem=2N, and group number by Pnumber=2N-M in HPNC. Suppose that in all situations each group's node number is the same as Gnumber=2M. Therefore, in the Chord system the pointer gauge's size which each node needs to maintain is N, while in the HPNC system, each super node pointer gauge size is only N-M, and other nodes do not need to maintain the pointer gauge, obviously N-M<N. Since HPNC uses the group centralized management method, each central node needs to maintain two more tables, Node_table and MessageIndex_table. Although the system's memory price is increased, more stable performance and more reliable management are purchased.

5.2 Searching Efficiency

A key parameter to evaluate the overlay network is how to locate to the nearest node. We all know that the biggest problem of the Chord overlay network localization is: In the logical space, the node relations cannot match the actual network, namely in the overlay network the neighboring node is possibly very far in the under layer physical network. But in HPNC, by the form of MAC grouping, the information resource is issued inside and outside the group, the efficiency of searching in physical local is greatly raised, therefore the problem is well solved.

Figure 6 is a comparison between HPNC and Chord in jumping number aspect, while the group number is 5, and localization success ratio of the resource in the group is 50%.

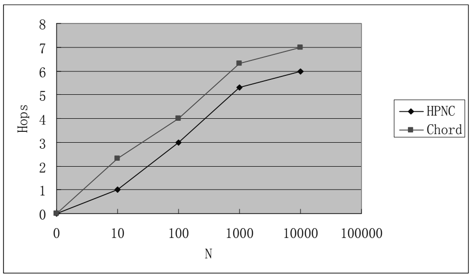


Figure 6. Comparison of searching efficiency performance

5.3 Scalability of System

When the number of nodes in the Chord system increases, time and space complexity of the system will exponentially increase with the increasing number of node. When the number of nodes increases from 210 to 230, data storage capacity of each node will be 3 times. However, in HPNC, as long as the number of group is invariable, the data storage capacity will not change a lot. The increasing number of nodes in the group only raises the data storage capacity of the central node, but will not cause big fluctuation in the whole network, so it has better extendibility.

References

[1] Zhou Wenli, Wu Xiaofei. Survey of P2P technologies [J]. Computer Engineering and Design.Vol.27, No.1,76-79,2006.

[2] Lu Xiangchen, P2P Technology and Its Application [J]. Computer World, Vol.12,No.9,13-14,2002.

[3] Shao Ying, Liu Ye. Analysis and Implementation on Chord of Structured P2P Routing Protocol[J]. Computer Engineering. Vol.33, No.19,122-124,2007.

[4] Karger D, Lehman E, Leighton T, et al. Consistent Hashing and Random Trees: Distributed Caching Protocols for Relieving Hot Spots on the World Wide Web[C]//Proceedings of the 29th Annual ACM Symposium on Theory of Computing. 1997-05.

[5] ROWSTRON A,DRUSCHEL P. Pastry: Scalable, distributed object location and routing for large-scale peer-to-peer systems [A]. Proceedings of the 18th IFIP /ACM International Conference on Distributed Systems Platforms (Middleware 2001) [C]. 2001.