Research on P2P resources search algorithm based on cloud computing

GE Junwei[1, a]   WANG Yanfeng[2,b]   FANG Yiqiu[3,c]

[1, 2, 3]College of Computer Science and Technology Chongqing University of Posts and Telecommunications 400065, Chongqing,China

[a]gejw@cqupt.edu.cn   [b]wangyanjieyou@yahoo.com.cn [c] fangyq@cqupt.edu.cn

Abstract. This paper puts forward a model combined cloud computing and P2P, and then achieves chord resource search algorithm in the model. Because chord algorithm one hand, does not consider the problem of node heterogeneity, on the other hand, its routing table has large redundant information. So chord algorithm is improved from the two areas, and last MRC-chord is proposed. The algorithm first defines that the nodes of the same geographical form a ring, and in each ring, the node with strongest overall performance is selected to be super cloud node, all of the super cloud nodes form the master ring , then improves routing table. Experimental results show that: the improved algorithm can reduce average routing hops and average network delay effectively, so improves the efficiency of resource search.

Introduction

In now days, "Cloud computing" is an emerging distributed computing model, but the technology is not mature, the cost of cloud services is high. In view of this, the old P2P technology shows its unique advantages in this new technology stage, and provides a new vitality for the development of cloud computing.[3] The effective combination of cloud computing and P2P technology is the development trend，and the corresponding network resources search is the core of the research. [4]

Now,in P2P network resource search algorithm, because of the most classic and famous, distributed, load balancing, the effectiveness and flexibility of naming, Chord is the best choice for P2P resources search based on cloud computing. [6]Chord is a structured P2P model. In Chord, nodes and all keywords are mapped to an m-bit identifier with the same hash algorithm, each node is connected in accordance with its identifier order. Finally, all nodes form a Chord ring. [2]

In the Chord ring, each node has a pointer to the precursor node, a pointer to the successor node and a routing table. In the routing table of Node n, its j-th($1 \leq j \leq m$),saving the first node that the distance to the current node is 2^{j-1} In the clockwise direction in the Chord ring, denoted by finger[j] , finger [j] $=n+2^{j-1} (mod 2^m, 1 \leq j \leq m)$. [1]

However, Chord does not consider the problem of node heterogeneity and its routing table has large redundant information. So, in order to improve the efficiency of resource search effectively, this paper presents an effective network topology model and improve the structure of the routing table.[5]

Related Work

Design ideas

We can look at the cloud computing network from a macro point of view. Cloud computing network is composed by many cloud server, so one cloud server can be seen as a node in the network, and as the basic elements of the network topology, we call the cloud server as the cloud node. There are also some differences that exist in the data processing capacity, storage space, on-line time and bandwidth between the cloud nodes.

Original Chord algorithm does not consider the heterogeneity between each node in the network, therefore, a model of Chord algorithm considering the heterogeneity of nodes is proposed, this

model transforms the original, single Chord ring into a master-slave multi-ring structure with the help of the network number grouping mechanism of the cloud node IP address and the grading mechanism of the cloud node performance, and takes full account of the heterogeneity of the cloud nodes. In order to reduce the routing table redundancy and to improve the resource lookup performance, this paper puts forward a way that improves routing table by modifying the formula of the cloud node routing table . The model based on the ideals above is called as MRC-chord.

Basic concepts and the construction of the model

Definition 1 A m-bit node identifier that is composed by two parts: HID + LID, wherein HID is m1 bits, LID is (m-m1) bits.

Definition 2 Cloud node HID: The value that is hashed by the network number of the cloud node IP address.

Definition 3 Cloud node ID: The value that is hashed by the cloud node MAC address.

Definition 4 Slave Chord ring: Chord ring which are formed by the cloud nodes with the same HID.

Definition 5 Master Chord ring: Chord ring which are formed by all of the super cloud nodes.

Definition 6 Super cloud node: The cloud node with the best performance in one slave Chord ring.

Definition 7 Backup cloud node: The cloud node whose performance is worse than Super cloud node but better than any other cloud nodes in one slave Chord ring.

Definition 8 The performance of the cloud node: Each node in the system will be evaluated when the network is initialized. The assessment indicators include about four factors：CPU performance, memory capacity, bandwidth, as well as on-line time.
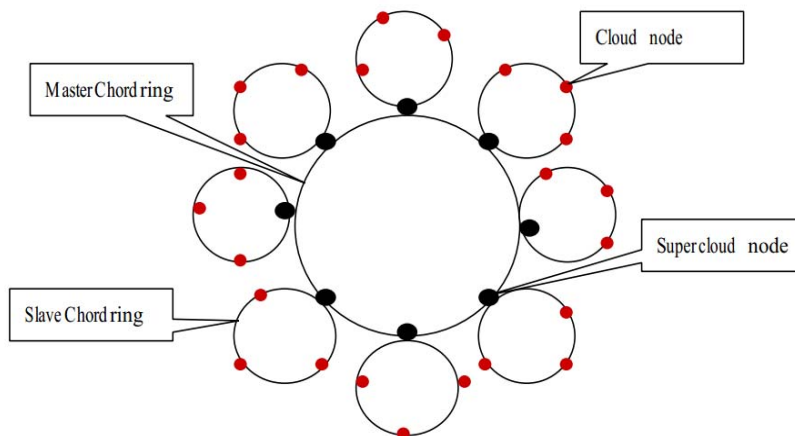
The specific structure of the model is shown in Figure 1:

Figure 1 .MRC-Chord model

In this model, all ordinary cloud nodes can store (key, value), query and publish the data, can join or leave the network. All the super cloud nodes not only store data (key, value), but also are responsible for responding to the requests that the ordinary cloud nodes query, publish the data, join or leave the network. Due to the dynamic nature of the P2P network, super cloud node may also failure or leave the network. Once super cloud node failure, it will cause a great influence in the located Chord ring, as well as the entire network. Therefore, in order to reduce this impact, this paper puts forward backup cloud node. The backup cloud node stores all the information belongs to the super cloud node. Once the super cloud node in the ring failure, the backup cloud node will take over the super cloud node immediately.

The design of the routing table

For each finger table of one node, in order to improve the efficiency of the resource search, this paper presents a new algorithm that modifies the original Chord finger table formula.

The formula is as follows:

$$\text{Finger[j]}=\begin{cases}(n+d)\bmod 2^{m}...................1\leq i\leq\left\lceil\log_2^d\right\rceil, j=1\\(n+2^{i}-2^{i-2})\bmod 2^{m}.........\left\lceil\log_2^d\right\rceil<i\leq m, j=i-\left\lfloor\log_2^d\right\rfloor\end{cases}\qquad(4)$$

According to the analysis of the results calculated by the original Chord finger table formula, when i is within the range of $[1, \left\lceil\log_2^d\right\rceil]$,the nodes pointed by the successor pointer are the same one . With the help of the first expression in the formula (4), the repetitive portion of the finger table is eliminated.

Table 1.N1's finger table calculated by the original Chord formula

Finger[j]	successor
N1+1	N8
N1+2	N8
N1+4	N8
N1+8	N14
N1+16	N21
N1+32	N38

Table 2.N1's finger table calculated by the formula (4)

Finger[j]	successor
N1+6	N8
N1+12	N14
N1+24	N32
N1+48	N51

Compare Table 1 and Table 2, there is no redundant information in N1's finger table calculated by the formula (4), and the length of the finger table is also shortened greatly. so the efficiency of resource search is improved.

Simulation

This experiment mainly simulates two models of Chord and MRC-Chord in the same environment, the simulator is Peersim.

The experiment of average routing hops: The result is shown in the Figure 2, when the number of the nodes is 1000, The average routing hops in MRC-Chord is less than in Chord. And with the increase of the number of nodes, the curves of MRC-chord is smoother than chord. That is, the more the number of the cloud nodes, the more obviously the query efficiency is improved in MRC-chord.
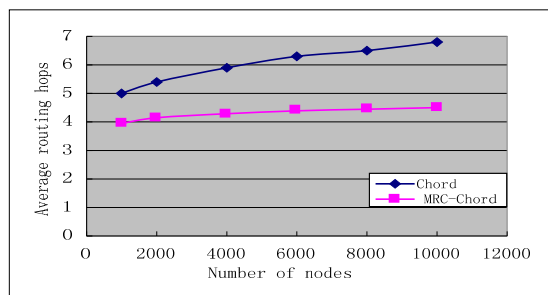

Figure 2. Average routing hops in Chord and MRC-Chord

The experiment of average routing delay: Figure 3 shows routing delay comparison between MRC-Chord and Chord, MRC-Chord routing delay is shorter than Chord, and with the increase of the number of nodes, the reduction in the average delay is more obvious. So MRC-Chord routing performance is better than Chord.
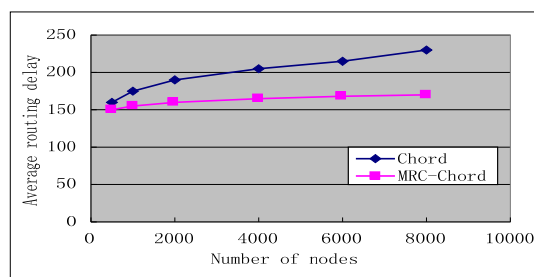
Figure 3. Average routing delay in Chord and MRC-Chord

Summary

This paper combines cloud computing and P2P, and puts forward MRC-Chord which is formed by the cloud nodes, these cloud nodes are all acted by the cloud servers. MRC-Chord on one hand solves the problem that chord algorithm does not consider the node heterogeneity by the way of Constructing super cloud nodes, on the other hand reduces the routing table redundant information by the way of improving the routing table algorithm. Experimental results show that: the improved algorithm can reduce average routing hops and average network delay effectively, so improves the efficiency of resource search.

References

[1]Gupta A,Liskov B,Rodrigues R. Eficient Routing for Peer-to-Peer   Overlays[C]//Proceedings of the 1st Symposium on Networked Systems Design and Implementation. San Francisco, CA, USA:[S, n.],2004.

[2] Jiang Shouxu, Han Xi first, Li Jianzhong. Chord system based on super nodes[J]. Mini-Micro Systems, 2007,28(2):266-270.

[3] Zhang Zhen. Resource search algorithm in Peer-to-peer network[J]. Computer Engineering and Applications, 2006,31(6):147-152.

[4] Stoica I,Morris R, Liben-Nowell D, et a1. Chord: A scalable peer- to-peer lookup protocol for interact applications[C]. IEEE/ACM Transactions on Networking, 2004,12(2):205-218.

[5] Yu S, Yu J, Kamil K, et a1. DR-Chord-F an efficient double-ring chord protocol [C]. Ummuqi, China:Proc 7th 1EEE Int Conf Grid and Coop Comput,2007: 16. 18.

[6] Ozalp Babaoglu, Moreno Marzolla, Michele Tamburini, et a1. Design and Implementation of a P2P Cloud System[C].Technical Report UBLCS-2011-10.