

Parallel Routing Method in Churn Tolerated Resource Discovery

Zhuang Miao¹

College of Command Information Systems, PLA Univ. of Sci. & Technology, Nanjing, 210014, China
E-mail: emiao_beyond@163.com

Xiancai Zhang

College of Command Information Systems, PLA Univ. of Sci. & Technology, Nanjing, 210014, China
E-mail: zhangxiancai12@sina.com

Peiyi Yu

College of Command Information Systems, PLA Univ. of Sci. & Technology, Nanjing, 210014, China
E-mail: ypy02784@163.com

Jiabao Wang

College of Command Information Systems, PLA Univ. of Sci. & Technology, Nanjing, 210014, China
E-mail: jiabao_1108@163.com

Qianqian Zhang

College of Command Information Systems, PLA Univ. of Sci. & Technology, Nanjing, 210014, China
E-mail: miipl606@163.com

Structured P2P network has good scalability, robustness and routing algorithm with high performance and low cost. Thus, in theory, structured P2P network is a good choice for network resource discovery. However, because of churn, it has bad performance and can't be used extensively in real world. To overcome this problem, a parallel routing method is presented. Network resource information is organized by the node group. The query is moved in node group by Random Walk, and among different groups by the routing algorithm of FreePastry. The parallel algorithm is proposed and its complexity is analyzed. The experiment results verify that the parallel routing algorithm cost is the same as FreePastry and the traffic load is a little bigger than that of FreePastry, but smaller than that of the other ones. The parallel method is tolerated to churn and is helpful for resource discovery of structured P2P network with churn.

CENet2017
23-24 July 2017
Shanghai, China

¹Speaker

1.Introduction

Theoretically, structured P2P is suitable for building a resource discovery [1]. Structured P2P based resource discovery protocol is the best in respect of robustness and scalability. To achieve high performance and keep low cost, routing algorithm is the key point. However, because of churn, it has bad performance and can't be used extensively in real world. In structured P2P network, high burstiness of information transmission will cause inappropriate big amount of the bandwidth. The P2P network has to transfer mass data in the period of churn and can't deal with the users' requests, which debases the efficiency of resource discovery. Efficient routing algorithm is the key point of this problem.

To solve the problems of churn in structured P2P for resource discovery, a churn tolerated resource discovery protocol (CTRDP) is presented in this paper. The protocol has two main parts, parallel routing algorithm and corresponding resource information organization [2]. Compared with traditional routing algorithms, the presented one can improve the success rate of routing and remain the high routing efficiency with low cost.

2.Related Works

Li proposed a strategy of node selecting [3] and Tian proposed a strategy of data redundancy[4]. Their strategies are helpful to reduce the transportation load. To accelerate the data dissemination under churn, Li proposed a churn-resilient protocol. It can assure alternating path and data proximity and it is helpful to construct a P2P content delivery system [5]. Zhang summarized several key steps to solve Churn problem [6]. On modeled communication-oriented structured P2P system, a real-life prototype is implemented to evaluate the performance [7]. Trifa thought the reorganization of the affected objects' replicas set results in performance degradation and objects loss. To reduce churn effects, he proposed a new replication technique [8]. Applying the OverSim framework and OMNeT++ simulator, Rocamora analyzed hierarchical and flat DHTs comparatively with or without churn [9].

Besides, several complex networks works are meaningful for resource discovery. Hong modeled many coupled infrastructure networks as interdependent networks and studied a failure cascade process considering their effects on failure propagation and proposed a costless defense strategy [10]. He proposed capacity allocation and load redistribution model respectively, and investigated an in-process restoration strategy [11, 12]. Furthermore, he analyzed and modeled epidemic spreading model of complex dynamical network, which is helpful for the designing of effective control strategies [13].

3.Parallel routing Algorithm

The routing algorithm in structured P2P has low success rate under churn, which limits the widely application of resource discovery. The main reasons are the routing table failure and node leaving before routing. Routing table failure is that when some nodes leaving the overlay, the corresponding table entries cannot be renewed in time. The redundancy table can solve this problem and the cost of table renewed is decreased.

Node leaving before routing is that the node doesn't route the request before it is off-line. There is only one query routed in the overlay, so the success rate of routing is low. We can reach a conclusion that increasing the number of query will improve the success rate. Based on the theory, a parallel routing algorithm which combines Random Walk [14] and traditional routing algorithm is designed.

3.1 Routing Process

According to parallel routing algorithm, the query is moved in node group by Random Walk, and among different groups by the routing algorithm of FreePastry (<http://www.free-pastry.org/FreePastry>). There are multiple queries routed concurrently in overlay. Figure 1 shows the process of parallel routing algorithm.

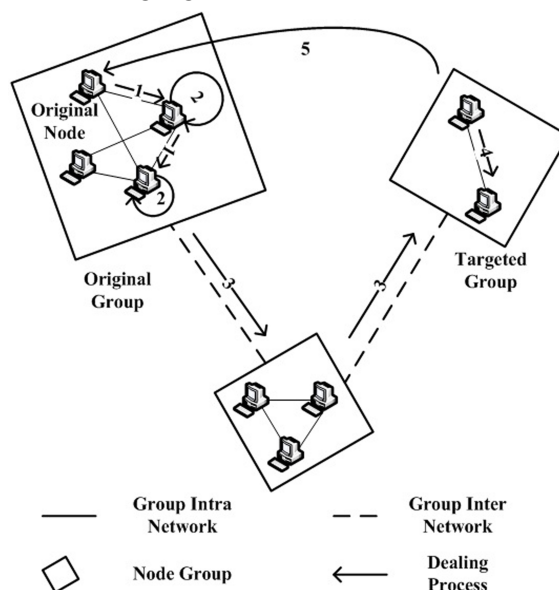


Figure 1: Process of Parallel Routing Algorithm

Step 1: Select the node group whose code has the longest prefix with the query code as the original node group, and move query by Random Walk in the original group. If TTL (Time to Live) is less than 7, go to step 2;

Step 2: Check the node which received the query in the original group whether there is information can meet the query firstly. If so, go to step 4, otherwise, go to step 3;

Step 3: Receive all nodes in original group route queries independently according to FreePastry. When the node is in other group receiving query, check whether there is appropriate resource information. If so, go to step 4, otherwise, return to step 3;

Step 4: Send the query to other nodes of the group by Random Walk when it arrives at the targeted node group, and go to step 5;

Step 5: Return all nodes that accept the query with corresponding resource information to the original node.

This algorithm will improve the routing success rate.

3.2 Complexity Analysis

There are multiple queries routing in the overlay according to the parallel routing algorithm. So long that one query arrives at the targeted node group, the routing is success. Compared with traditional algorithm, the routing success rate of the new algorithm is higher.

Assuming the routing success rate of one query is p , and there are n queries routed independently in overlay, the success rate of parallel routing algorithm can be described as: $P = 1 - (1-p)^n$. It can be proved that $P \geq p$.

Proof:

$n = 1, P_1 = 1 - (1-p) \geq p$, so $P \geq p$;

Amusing $n = k, P_k = 1 - (1-p)^k, P_k \geq p$;

when $n = k + 1, P_{k+1} = 1 - (1-p)^{k+1}$;

$1-p < 1$, so $(1-p)^{k+1} < (1-p)^k$;

$P_{k+1} > P_k \geq p$;

So $P \geq p$

When n is bigger than 1, the success rate of routing is higher than the traditional one. The query can arrive the targeted node group in $O(\log N)$ hops. Every query is routed by FreePastry which can route query in $O(\log N)$ hops. And it only takes $O(1)$ hops to move query in node groups.

The cost of one query routing in the group inter network is very low, and the TTL of Random Walk in the original and targeted group is small enough to take the low cost. The traffic load of our method is bigger than that of FreePastry, but less than Random Walk and Flooding. The experiments of parallel routing algorithm will be given as follows.

4. Experiments

To verify the theoretical results of our method, we experimentalize to measure the performance. The experiment is based on an implementation in Java. PeerSim (<http://peersim.sourceforge.net/>) is a simulator and it is easy to extend. Our algorithm is simulated by PeerSim. In the experiment, the churn is modeled as exponential distribution (ED) and the time is the cycle of simulation. By changing the nodes' mean online time in PeerSim, the level of churn is adjusted constantly. Long online time means low level of the churn and they are in accordance with positive relation. We focus on the performance of success rate under churn, routing time and traffic load that produces it.

The target of our experiment is as below. First, we compare the traffic load of our method with FreePastry. It is to prove that under churn, the traffic load and burstiness of our method are both reduced. Furthermore, the simulator of CTRDP is presented. It is to verify that our parallel routing algorithm has better routing success rate and is the same as FreePastry in efficiency and cost.

The total number of nodes in the network is 1000. If one node leaves, there must be a new node to join immediately. The information of each node consists of two types, resource and routing. The amount of each information is 10 and the length of code is 128. Number of resource type and bandwidth of node are in line with the uniform distribution.

Success Rate under Churn. The comparison of success rate between CTRDP and FreePastry is given in Figure 2. The degrees of churn keep the same value as that described in previous section. The result shows that CTRDP always has a higher success rate than FreePastry. The reason is that multiple queries are routed in the overlay independently. Once a query arrives at the target, the routing is successful. In the redundancy table, each entry of table may be pointed to several nodes. When one part of the nodes leaves the overlay, the table can renew immediately with others. With the churn level growing, the number of node which left the overlay becomes bigger and bigger. So the renew time is growing with the increase of churn level.

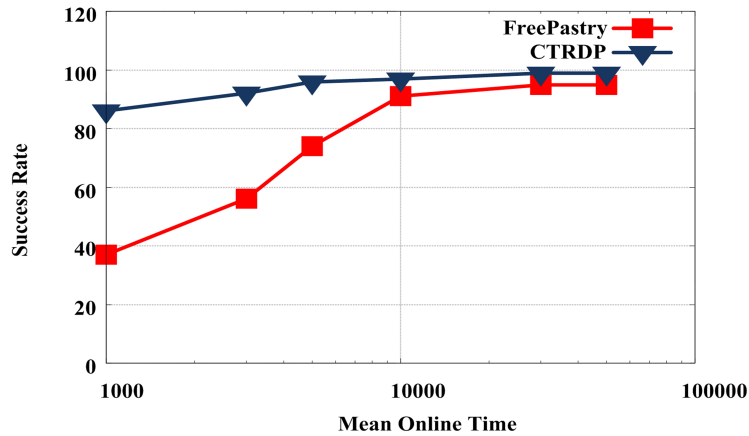


Figure 2: Success Rate of Routing under Churn

Routing Time. The analysis in Section 3.2 shows that the cost of parallel routing algorithm should be the same as FreePastry. The experiment verifies that the new routing algorithm take $O(\log N)$ hops for routing query. The result is demonstrated in Figure 3. The “Static Network” means that nodes are static in the overlay. The “Dynamic Network” means that it works with churn. The churn model is ED and the mean online time is 5000. The “Theoretical Value” is the theoretical value of FreePastry’s routing algorithm, $\log_{16} N$; because $b = 4$. We can get it that parallel routing algorithm can route query to target in $O(\log N)$ easily like FreePastry easily. Our traffic load of is little bigger than FreePastry's, but less than Random Walk's and Flooding's.

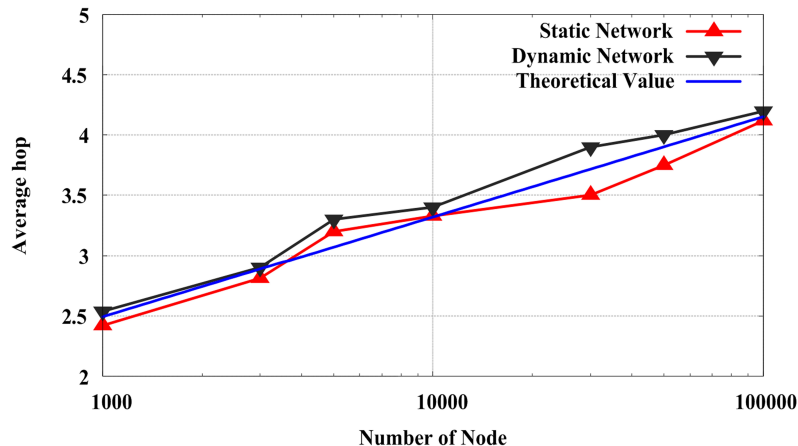


Figure 3: Routing Time of Parallel Routing Algorithm

5. Conclusion

In this paper, a churn tolerated parallel routing method is presented for resource discovery. Parallel routing algorithm is combining Random Walk with the traditional routing algorithm. There is more than one query to be routed independently. However, once one arrives at target, the routing will be success. So the success rate of our algorithm is higher than traditional ones. With the redundancy table, the success rate of routing becomes higher. And then, the analysis of parallel routing algorithm is presented. The conclusion is that the cost of parallel routing algorithm is the same as FreePastry. The traffic load is little bigger than FreePastry, but smaller than other ones. Finally, the experiments are presented to verify all inferences and conclusions.

References

- [1] R. Ranjan, A. Harwood, R. Buyya. *Peer-to-Peer based resource discovery in global grids: A Tutorial*[J]. IEEE Communications Surveys and Tutorials, 10(2): 6-33(2008)
- [2] Z. Miao, Q. Q. Zhang, S. Q. Wang, Y. Li, W. G. Xu, J. B. Wang. *A resource information organization method based on node encoding for resource discovering*[J]. Lecture Notes in Electrical Engineering, 277: 1263-1270(2014)
- [3] J. Li, J. Stribling, R. Morris. *Bandwidth efficient management of DHT routing tables*[C]. Proceedings of 2nd Conference on Symposium on Networked Systems Design & Implementation, USENIX, Berkeley, CA, pp, 99-114(2005)
- [4] J. Tian, Y.F. Dai. *Study on durable peer-to-peer storage techniques*[J]. Journal of Software, 18(4): 1379-1399(2007)
- [5] Z. Li, G. Xie, H. Kai, Z. Li. *Churn-Resilient Protocol for Massive Data Dissemination in P2P Networks*[J]. IEEE Transactions on Parallel & Distributed System, 10(1): 45-65(2017)
- [6] Y. X. Zhang, D. Yang, H. K. Zhang. *Research on Churn Problem in P2P Networks: Research on Churn Problem in P2P Networks*[J]. Journal of Software, 20(5): 1362-1376(2009)
- [7] Z. Ou, E. Harjula, O. Kassinen, M. Ylianttila. *Performance evaluation of a Kademlia-based communication-oriented P2P system under churn*[J]. Computer Networks, 54(5): 689-705(2010)
- [8] Z. Trifa, M. Khemakhem. *A novel replication technique to attenuate churn effects*[J]. Peer-to-Peer Networking and Applications, 9(2): 1-12(2016)
- [9] J. Rocamora, J. R. Pedrasa. *Evaluation of hierarchical DHTs to mitigate churn effects in mobile networks*[J]. Computer Communications, 85: 41-57(2016)
- [10] S. Hong, B. Wang, M. Xiao, J. Wang, T. Zhao. *Failure cascade in interdependent network with traffic loads*[J]. Journal of Physics A: Mathematical and Theoretical, 48(48): 485101(2015)
- [11] S. Hong, X. Zhang, J. Zhu, T. Zhao, B. Wang. *Suppressing failure cascades in interconnected networks: Considering capacity allocation pattern and load redistribution*[J]. Modern Physics Letters B, 30(05): 1650049(2016)
- [12] S. Hong, C. Lv, T. Zhao, B. Wang, J. Wang. *Cascading failure analysis and restoration strategy in an interdependent network*[J]. Journal of Physics A: Mathematical and Theoretical, 49(19): 195101(2016)

- [13] S. Hong, H. Yang, T. Zhao, X. Ma. *Epidemic spreading model of complex dynamical network*[J]. International Journal of Systems Science, 47(11): 2745-2752(2016)
- [14] Q. Lv, P. Can, E. Cohen, K. Li. *Search and replication in unstructured peer-to-peer networks*[C]. Proceedings of 16th ACM International Conference on Supercomputing, Acm, NewYork, NY, pp, 132-143(2002)

POS(CENet2017)073