2010 Second International Conference on Communication Software and Networks

A Directed-multicast Routing Approach with Path Replication in Content Addressable Network

Wenbo Shen, Weizhe Zhang, Hongli Zhang, Hui He
School of Computer Science and Technology
Harbin Institute of Technology
Harbin, China
shenwenbosmile@gmail.com

Abstract—**In order to improve the efficiency of routing and resource location under large scale failure of Content Addressable Network, a directed-multicast routing approach is proposed to solve the backtrack problem of the traditional greedy approach. Possessing the advantages of directed routing approach and multicast, this new approach also has an extending coefficient to increase its feasibility. Meanwhile, we explore a new replication approach that we term Path Replication. Combining with Path Replication, the directed-multicast routing approach is more efficient in routing process. The efficiency of directed-multicast routing was proved in the simulations on PlanetSim.**

Keywords-Content Addressable Network; routing algorithm; directed multicast; Path Replication

I. INTRODUCTION

Peer-to-Peer networks are robust, low cost, scalable, which can make a good use of Internet resources such as computing, storage, and network bandwidth, so they have been widely used in information retrieval, resources sharing, distributed computing, large scales storage management systems and other areas. Content Addressable Network [1](CAN) is based on the multi-dimensional spatial structure of Peer-to-Peer network, using a distributed hash table which maps data and nodes into zones to complete data storage and query in the multi-dimensional Cartesian space. Compared to the structure of ring-structured-based Chord network [2], the XOR distance metric based Kademlia [3] and hop table based Skipnet [4], the multi-dimensional space based CAN can combine the information of network measurement and geography to solve the unmatched problem existed in Peer-to-Peer overlay network topology [5].

Routing approaches of CANs are the current researching focus. An efficient routing approach should locate resources quickly, shorten the routing path, thus to response rapidly and reduce commutation overhead when resources join and leave the network. The traditional routing approaches are divided into two categories: 1 broadcast routing approach. In this approach, nodes forward query messages which they receives to all their neighbors. So a large number of useless message would be generated and forwarded in network, consuming network bandwidth and increasing communication overhead [6]-[7]. The routing efficiency of broadcast approach can be very low. 2 greedy routing approach. Greedy routing approach, which is used in Content Addressable network, is a kind of directed-unicast approach. By selecting next hops according to neighbors' coordinates, messages can get closer to their destination gradually in the directed routing approach [8]. However, this approach does not work well when nodes on the selected path fail. When a neighbor on the selected path fails, the current node has to find a new path by detecting other neighbors or even backtracks to the node who forwarded the message to it when it finds that all its other neighbors fail. Greedy routing approach is impossible to route and response rapidly because redetection and backtracking are highly time-consuming. In addition, Wu and others proposed a theory that increases routing table entries to improve routing efficiency: each node's routing table would contain not only its neighbors' information, but also contain information of its neighbor's neighbors [9]. This approach will increase nodes' burden undoubtedly because the routing tables which nodes have to maintain are much larger than these in traditional approaches. And the maintaining of routing tables will also increase the communication overhead of the networks. Therefore, in this paper, we explore an alternative routing approach which we call directed-multicast routing approach, and combine it with a new replication approach which replicates data on path, to improve the efficiency of routing and resource location of Content Addressable Network.

First of all, we give a brief introduction of basic principles of directed-multicast routing approach by discussing the shortcomings of broadcast routing approach and greedy routing approach, focusing on the dynamic character of Content Addressable Network and large scale failure problem; Then, we will bring in an coefficient to increase directed-multicast routing approach's feasibility, improving its efficiency under large scale failures; Meanwhile, we discussed a new replication approach that we term Path Replication, where data are replicated equally likely when they are passing through nodes of different paths. Then, the directed-multicast routing approach is combined with Path Replication to improve its retrieval and query efficiency and make system more fault-tolerant and reliable. Finally, we give experiments on PlanetSim to evaluate the performance of extended directed-multicast routing approach and Path Replication, comparing with the traditional greedy routing approach.

978-0-7695-3961-4/10 $26.00 © 2010 IEEE
DOI 10.1109/ICCSN.2010.8

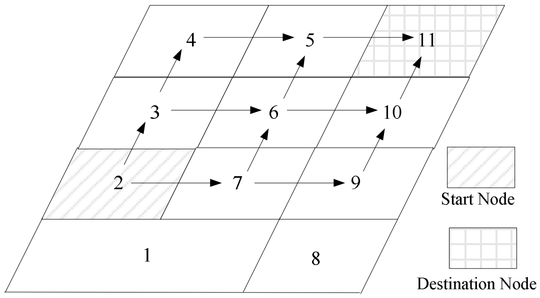

Figure 1. Message flows of directed-multicast approach under two-dimensional logical space

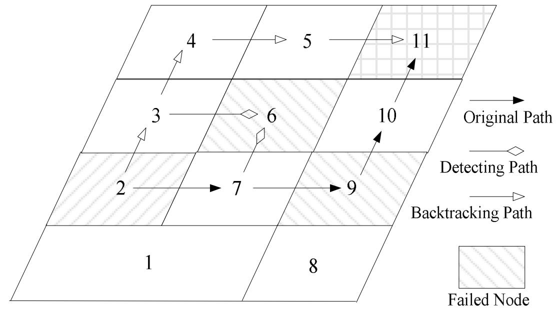

Figure 2. Routing process of greedy approach when nodes fail

II. Approach Of Directed-Multicast Routing

Derived from Content Addressable Network, directed-multicast routing approach is also based on multi-dimensional space logical structure. If the dimension of multi-dimensional space is d, thus entry number of the routing table of each node in the logical space is at least $2d$ (except for nodes on boundary). In the d-dimensional space, there is a partial order between two nodes of their locations in the d-dimensional coordinate system. The basic idea of directed-multicast routing approach is that, in the process of forwarding messages, if the ith-dimensional coordinate of the destination node is greater than that of current node, the current node will forward messages to all its neighbors on positive direction of the ith-dimension instead of to just one neighbor on the positive direction or all its neighbors in all directions. Fig. 1 illustrates message flows of directed-multicast routing approach under two-dimensional space. The start node is node 2; the destination node is node 11. The arrows represent the flows of routing messages. It can be seen that in directed-multicast routing approach, the routing messages flow through all nodes between node 2 and node 11, forming a rectangular routing area. Since the start node and destination node are distributed in the multi-dimensional space randomly, so the average distance between the start node and destination node is close to half of the length of diagonal of the logical space. The average area between start node and destination node is close to $\sqrt[d]{Size}$ (d represents space dimension, *Size* represents the whole logical space size). We suppose that there are N nodes in the logical space, thus the average number of nodes which are included in the rectangular routing area is $\sqrt[d]{N}$. Therefore, the communication overhead of directed-multicast routing approach should be O ($\sqrt[d]{N}$) level. Comparing to the broadcast routing approach's O (N), the directed multicast routing approach has a significant improvement in performance. When the number of nodes is huge (N value is large) and the logic spatial dimension is high (d value is large), the directed-multicast routing approach has a much lower communication overhead than broadcast routing approach.

Due to its feature of multicast, directed-multicast routing approach selects several paths to forward it messages, which enable it tackle node failure efficiently in routing process. However, greedy approach cannot handle large scale failure problem efficiently. For example, in Fig. 1, the best path selected by greedy approach may be the path 2-7-9-10-11. Suppose that node 6 and node 9 are invalid. Firstly, receiving message from node 2, node 7 may forward the message to node 9 in routing process. As node 9 has crashed, node 7 has to redetect, and the message will be forwarded from node 7 to node 6. However, node 6 fails either. The routing path may backtrack to the start - node 2, which has to send message to another neighbor - node 3. Node 3 receives the message, and the next hop which is selected may be the failed node 6, so node 3 has to redetect again and finally forwards this message to node 4. The final routing path is 2-3-4-5-11. Fig. 2 shows the redetection and backtrack process. In the above process, it has to redetect three times and backtrack one time. These four processes slow down the system response speed. However, in the directed-multicast routing approach, if the start is node 2, the destination is node 11. Due to that the x, y coordinates of node 11 are both greater than those of node 2, node 2 will forward messages to the positive direction of both x and y axes, sending messages to node 3 and node 7. And each node on routing path will also forward messages to next hops in this coordinate-comparing way. The message flows in Fig. 1 will be formed as nodes forward their messages using directed multicast routing approach. Even node 6 and node 9 fail, which only interrupts two paths of the directed-multicast approach, routing messages through the path 2-3-4-5-11 can also arrive at the destination node. Therefore, node failures do not affect the efficiency of directed-multicast routing approach severely.

However, if node 5 and node 10 are invalid in the two-dimension space, the directed-multicast routing approach is not feasible either. In this case, we extend directed-multicast routing approach to solve this problem. Comparing coordinates in the routing process, if one node's ith-dimensional coordinate is greater than that of the destination node, we still forward messages to this node, but these messages need to be identified, allowing a maximum of K times forwarding beyond the above rectangular area, and the value K varies as networks change. This extended approach forms a routing message coverage which covers the original directed-multicast routing approach's coverage. And the extended directed-multicast routing approach's message coverage will also cover all neighbors of the destination.

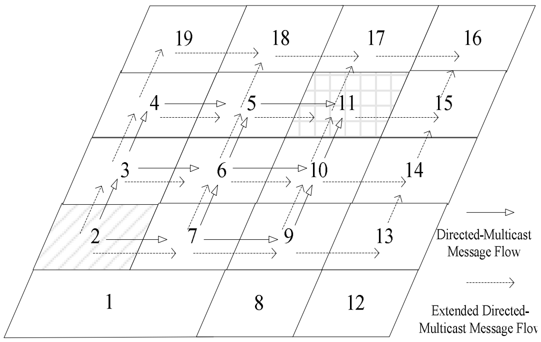

Figure 3. Message flows under two-dimensional logical space

```
Let C.Ci, S.Ci, D.Ci be the ith dimensional coordinate of the Current node, the Start
node and the Destination node. Let Message.k be the extending coefficient.
Let nextHopList be the next hop list of the current node.
Begin
    For each dimensional i
        if (D.Ci > C.Ci> S.Ci)
            nextHopList.add (all neighbors in i dimensional positive direction);
        else if (S.Ci > C.Ci> D.Ci)
            nextHopList.add(all neighbors in i dimensional negative direction);
        else
            if (Message.k>0)
                if (D.Ci > C.Ci)
                    nextHopList.add(all neighbors in i dimensional positive direction);
                else
                    nextHopList.add(all neighbors in i dimensional negative direction);
                Message.k --;
End
```

Figure 4. Extended directed-multicast routing algorithm

Thus, the routing message will be able to arrive at the destination node unless all neighbors of destination node have failed. In Fig. 3, solid arrows represent the original multicast routing message flows and the dotted arrows represent the messages flows of extended directed-multicast routing approach with K = 1.

In Fig. 3, the coverage area of original directed-multicast routing approach is covered by that of extended directed-multicast routing approach. And we can see that node 11 is beyond accessible unless all its neighbors failed. Obviously, in this case, it is unavoidable that node 11 is inaccessible.

As we have stated before, when a node receives a message, it will check its coordinates first, to find out whether it is located between the start node and the destination node. If so, it will forward this message in the coordinate-comparing way of the directed-multicast routing approach; if not, that is to say this message has been forwarded out of the rectangular area between the start node and the destination node. Therefore, the current node will decrease the extended coefficient held in this message and forward it towards the destination direction. The extended directed-multicast routing algorithm is summarized in Fig. 4.

From Fig. 4, we can see that directed-multicast routing algorithm is one special case of the extended directed multicast routing algorithm when $K = 0$.

III. Combined Directed-Multicast Routing With Path Replication

Data replication in the network can improve the efficiency of data retrieval. Traditional replication approach is biasing - more replications for data that are accessed by more people, and fewer replications for other data which may become popular in the future. And the uneven distribution of data copies will also reduce the efficiency of data replication. To solve this problem, we propose a new replication approach which we call Path Replication, where all data are replicated equally likely on each node when they pass through nodes on a path. And to improve its performance, we combine it with directed-multicast routing approach. Therefore, we have proposed four principles to describe the combination of these two approaches.

1) Data join routing principle: When new data are joining in the network; they are forwarded through paths formed by the directed-multicast approach which is discussed in Session II.

2) Data query routing principle: Data query messages are also forwarded along the paths formed by directed-multicast approach.

3) Data replication on path principle: Each node has a probability p to replicate the new data when the are passing through it.

4) Response principle: Every node who has the query data can respond to the query node and stop forwarding the query message.

Replicated according to the above principles, copies of original data would exist in the d-dimensional rectangular area which begins with the start node of data joining and end with the destination node of data joining. Fig. 5 presents a case of data joining that the start node is node 7 and the destination node is node 16. According to Principle 1, the data will flow through the area which is marked by arrows. When the data are passing by, each node on the routing path determines whether to replicate the data according to the probability p. Finally, there would be nodes (for example, the node 6, 13, 17) which have copies of the data (Areas in shadow) distributed in rectangular area.

When other node such as node 3 wants to access the data on node 16, according to Principle 2, query messages are forwarded in directed-multicast routing approach. Node 3's query messages will cover the rectangular area between node 3 and node 16. And node 6 and node 17 will respond to node 3 according to Principle 4. The query process is shown in Fig. 6. Arrows represent the flows of query messages. Since the closest node that has the query data is node 6 which is a neighbor of node 3, so just one message is enough to reach node 6 to get the data on node 16. As the normal query path without data replication may be 3-6-7-11-17-16. This query process requires 5 hops which are 5 times of hops in the Path

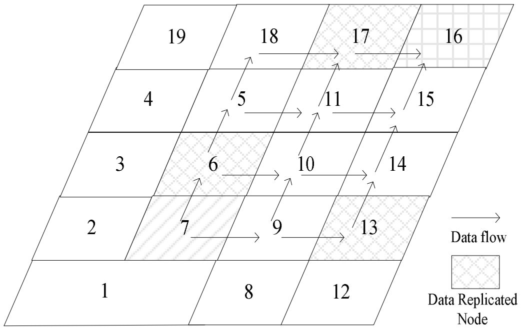

Figure 5. Data flows and data replicated on path in directed multicast routing approach

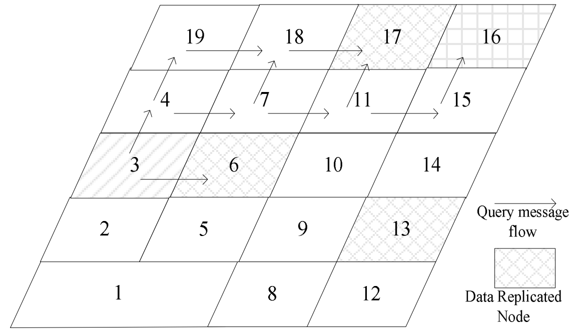

Figure 6. Query message flows under data replication case

Replication case. Therefore, Path Replication can shorten the data retrieval path greatly, thus can save much data retrieval time. And since nodes which have the query data, will not forward the query message, Path Replication also reduces communication overhead.

IV. Experiments And The Analysis Of Result

In this section, we present experimental results quantifying the performance of two new approaches by using PlanetSim [10] as a simulation platform. The first experiment reflects the performance difference between directed-multicast routing approach and the traditional greedy approach proposed by Sylvia [11] when large-scale failures happen in network. Then, we will give resources statistics within certain hops to illustrate the efficiency of the combination of directed-multicast routing approach and Path Replication.

A. *The efficiency of directed-multicast routing approach and greedy approach under large scale failure*

The query routing path's length, which is used to metric the system's overhead, also reflects the retrieval efficiency of routing approaches, especially when large scale failure happens in the network. When the dimension of the network is *d*, each node's routing table in Content Addressable Network contains neighbors in *2*d* directions, that is to say, each node in a *d*-dimension network has *2*d* neighbors. Our experiment uses two-dimensional space and 100 added nodes. Firstly, we make nodes in the network fail randomly. We term the ratio of failed nodes number to total nodes number as failure ratio. Then, active nodes are randomly selected to query resources and successful routing messages will be recorded to calculate the average hops number of greedy approach and extended directed-multicast routing approach under different failure ratios. When the failure ratio is too high(greater than 30%), the network may break into some small cut sets, the message data will not be persuasive as messages are routed within certain cut set, rather than in the whole network, therefore, we only collect data when the failure ratio is smaller than 30%.

Fig. 7 plots that the average hops number of resource queries of the two approaches under different failure ratios. From the Fig. 7, we can see that when the failure ratio is 0, the traditional greedy approach works as well as the directed-multicast routing approach. But as the failure ratio increases, the query resource overhead of greedy approach increases quickly, whereas the overhead of directed-multicast approach increases slowly. When the failure ratio is very high (30%), the directed-multicast routing approach has a much better performance than the traditional greedy approach. Therefore, directed-multicast routing approach is more efficient than traditional greedy approach under large scale failure.

B. *The comparison between directed-multicast routing approach with Path Replication and greedy approach*

The number of node which is added into the system is 100, and resources number is 2000. The dimension of Content Addressable Network space is assumed as two whereas its scope ranges from 0 to 2 ^ 20. We randomly select nodes to query resources in this network, and like the former experiment, successful routing messages are recorded to collect data of the resources number within certain hops. And we give data as Fig. 8.

Fig. 8 illustrates that as the probability of replication increases, hops which are needed to reach resources decrease. From Fig. 8, we can see that resources which can be reached within 0, 1, and 2 hops increase rapidly whereas resources which need many hops to reach decrease sharply. The number of hops needed to reach majority resources concentrates in the 3-8 without data replication. When we replicate data with a probability of 0.8, this number reduces to 0-5, which shows that a proper replicate of data can improve query efficiency. As fewer hops are needed to reach resources with the Path Replication, the communication overhead also reduces. By comparing with the non-replication traditional approach, we can see that the combination of the directed multicast routing approach and Path Replication approach has a better performance in resource location and query than the traditional greedy approach.

V. Conclusion

The CAN-based directed-multicast routing approach proposed by this paper works well under large scale failure, and makes the network more fault-tolerant and reliable.

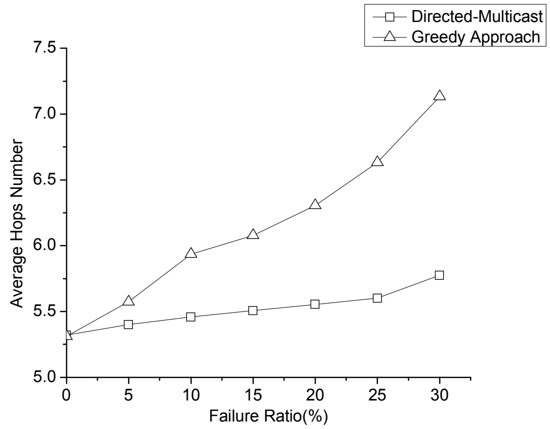

Figure 7. Average routing hops of resource queries under different failure ratios

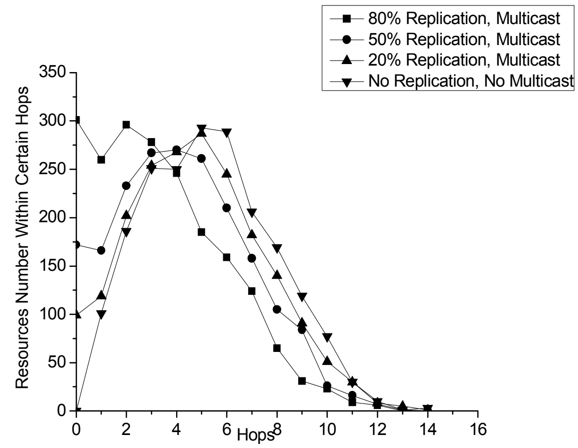

Figure 8. Comparison of query overhead between directed multicast routing with the traditional approach

Along with the Path Replication, this new routing approach can also improve the routing efficiency and reduce the communication overhead in traditional Peer-to-Peer system.

The next step of research will focus on the following three aspects:

1) Find a proper replication coefficient: Experiments should be conducted to find a proper coefficient for different data amount and different spatial dimensions, to establish a balance between copy number and resource retrieval efficiency.

2) Reducing the communication overhead: Nodes have to generate and forward more messages in directed-multicast routing approach than in traditional greedy routing approach. As the logical distance between query node and destination node can be calculated, we may limit the using of extended directed-multicast routing approach in cases that the distance between query node and destination node is within certain hops. Thus, we can avoid heavy communication overhead which the directed-multicast routing approach may bring to the network. So, we need more experiments to find proper distances under different network scales. In addition, directed controlled flooding can also be used to reduce the communication overhead brought by directed-multicast routing approach.

3) Solving overlay network breaks problem: It is difficult to solve the problem which exists in the current Content Addressable Network, that when large scale node failures occur, the system may crash and the overlay network may break into some cut sets, so more research should be done in this area to solve the dynamic failure problem of Content Addressable Network.

ACKNOWLEDGMENT

This paper was supported in part by the National Natural Science Foundation of China under Grant No.60703014; the National Grand Fundamental Research 973 Program of China under Grant No.G2005CB321806; the Specialized Research Fund for the Doctoral Program of Higher Education, SRFDP No.20070213044; the Development Program for Outstanding Young Teachers in Harbin Institute of Technology No.HITQNJS.2007.034; the China Postdoctoral Science Foundation No.20070410263; the Hei Long Jiang Postdoctoral Foundation No.LBH-Z07108.

REFERENCES

[1] S. Ratnasamy, P. Francis, M. Handley, R.M. Karp, and S. Shenker, "A scalable content-addressable network", in Proc. SIGCOMM, 2001, pp.161-172.

[2] I. Stoica, R. Morris, D.R. Karger, M.F. Kaashoek, and H. Balakrishnan, "Chord: A scalable peer-to-peer lookup service for internet applications", in Proc. SIGCOMM, 2001, pp.149-160.

[3] P. Maymounkov, and D. Mazieres, “Kademlia: A peer-to-peer information system based on the XOR metric”, in Proc. IPTPS, 2002, pp. 53-65.

[4] N.J.A. Harvey, M.B. Jones, S. Saroiu, M. Theimer, and A. Wolman, "SkipNet: A Scalable Overlay Network with Practical Locality Properties", in Proc. USENIX Symposium on Internet Technologies and Systems, 2003.

[5] S. Ren, L. Guo, S. Jiang, and X. Zhang, "SAT-Match: A Self-Adaptive Topology Matching Method to Achieve Low Lookup Latency in Structured P2P Overlay Networks", in Proc. IPDPS, 2004.

[6] S. Ratnasamy, I. Stoica, and S. Shenker, "Routing Algorithms for DHTs: Some Open Questions", in Proc. IPTPS, 2002, pp.45-52.

[7] A.I.T. Rowstron and P. Druschel, "Pastry: Scalable, Decentralized Object Location, and Routing for Large-Scale Peer-to-Peer Systems", in Proc. Middleware, 2001, pp.329-350.

[8] B.Y. Zhao, J. Kubiatowicz, and A.D. Joseph, "Tapestry: a fault-tolerant wide-area application infrastructure", presented at Computer Communication Review, 2002, pp.81-81.

[9] Z.D. Wu, L. Ma, and K. Wang, “Efficient Topology-aware Routing in Peer-to-Peer Network,” in Proc. GCC, 2002, pp.172-185.

[10] J.P. Ahulló, R.M. Andreu, H.T. Navarro, M.S. Artigas, P.G. López, and C.P. Gavaldà.(2009, Oct.) PlanetSim User and developer tutorial. [Online].Available: http://ants.etse.urv.es/planet/PlanetSim/

[11] S. P. Rathasamy, “A scalable content-addressable network,” Ph.D. dissertation, Univ. California, Berkeley, CA, Fall 2002.