A Method of Content Discovery in CDN

Xinhua E$^1$ and Binjie Zhu$^2$

$^1$Faculty of Information Technology, Beijing University of Technology, Beijing, 100124, China
$^2$China Mobile Group Beijing Company Limited, Beijing, 100007, China

Elson1900@163.com

Abstract. The CDN management system distributes content objects to the edge of the Internet to achieve the user's near access. The content discovery is an important issue in content delivery network. The content discovery algorithm is used to find the location of the content in the content distribution network. In this paper, we design a content resource discovery algorithm Global-hop, which optimizes the propagation process of topology change event messages. The events that cause the topology change are divided into two categories: one is the increase of nodes, and the other is the reduction of nodes. The reduction of nodes is divided into node withdrawal and node failure. Different from One-hop, the broadcast path of topology change event message in the Global-hop algorithm is tree-like broadcast. The broadcast topological information in the original algorithm divided nodes into several Slice, and each Slice was divided into several Unit. Each Slice has a Slice Leader, and each Unit also has a Unit Leader with other nodes being ordinary nodes. The message propagation delay of Global-hop algorithm is less than that of One-hop algorithm.

1. Introduction

Nowadays services such as Video-on-Demand and live TV programming available over Internet. Users want to have access to high quality video at any times. Content Delivery Network (CDN) is a solution that has been used to delivery traditional Web page content and Video content. Both content discovery and content routing belong to content request location routing, and the process of directing the user’s request to the server is called content request routing. Content request routing in content distribution networks is divided into directional routing and location routing: the so-called directional routing directs content request into the server near the user, and when there is no requested content in the local server, the location of the requested content is found by the location routing system.

2. Related researches

In the early P2P network, the function nodes of the network are all peer to peer, and content distribution is stored in network nodes, but the centralized index is applied. When a node needs to search for contents of other nodes, it should send a look-up request to the index server, which is a centralized method of P2P content indexing [1-7]. Whereas the centralized index server is not applied in the distributed P2P content index. Content index distributes in network nodes which cooperate with each other to complete the content search process.

3. Content resource discovery

This paper designed a content resource discovery algorithm Global-hop, which adopted the structured
DHT to allocate and query the index. ID number N.ID can be achieved by the Hash value of IP information of each node in the network, N.ID=HASH(IP). The ID number of these nodes are arranged in sequence in the order of the ID space to form a node ring. The size of ID space is determined by the number of digits m of the Hash value, T=2^m. The T is the maximum number of ID represented by ID space, and m is the number of bits of the HASH value. Therefore, the node ring composed of all nodes ID in the network is a subset of the ID space, and the number of nodes in the node ring is equal to the number of host nodes. It is a node ring with 5 nodes. M is 3, so the maximum number of nodes that the ID space can represent is 8. The IP hash values of the 5 nodes are 0, 1, 2, 5, 7, and they form a node ring. The node followed closely by node n on the ring is called the descendant node of node n, and node n is called the precursor node of this node.

Content index tables are distributed in the nodes of the global resource discovery layer. The Hash function value of the resource name is used as the ID number of the resource, R.ID. The Hash value of the host in which the resource is located is used as the node ID number. The resource ID and node ID constitute the index entry <R.ID, N.ID>, and index entries of all resources make up the resource index table. According to the resource ID, the index item is placed in the corresponding node. The resource ID is used as the key of the index, and the index information is stored in the descendant node of the key. As shown in Figure 1, indexes with keyword 2 are stored in nodes with NID 2, indexes with keyword 4 are stored in nodes with NID 5, and indexes with keyword 6 are nodes with NID 7.

Index information tables are needed in index query. The information of all the nodes are stored in index information tables, so the node in which corresponding information is located can be routed. After the arrival of resource request, it searches the routing information table according to the ID of the resource index to get the ID of its storage node, and then requests the corresponding node according to the node information. As shown in Figure 2, the node 7 is searching for an index with the keyword 3. By looking up the routing table, the descendant node with keyword 3 is 5, and then the request is routed to a node with a NID of 5. Find the descendant node of the object, find the index to the descendant node, and find the node of the resource according to the index.

![Figure 1](image1.png)  
**Figure 1.** The storage of the index in the node ring

![Figure 2](image2.png)
4. Topology information updating algorithm

The events that cause the topology change are divided into two categories: one is the increase of nodes, and the other is the reduction of nodes. The reduction of nodes is divided into node withdrawal and node failure. The node withdrawal is a node withdrawing from the network because of requirements like maintenance, etc, which is a prepared exit, while the node failure is a node failing, causing the node to be called a failure node, which happens at random.

When a node is added, the new node will get routing table information from an existing node in a network. The hash value of the new node IP is used as the NID number of the node. Insert node information into the routing information table based on the NID number. Broadcast the change information of topology with the descendant node of the new node functioned as the root node of the broadcast tree.

When the node withdraws, the withdraw node sends out the withdraw information to the descendant node, and its descendant node broadcast the change information of the topology functioned as the root node of the broadcast tree.

When the node fails, failure node will not take the initiative to send invalidation information, so nodes need to send Keep-Live information to determine that the neighbor node do not fail. Each node sends Keep-Live information to its predecessor node in a certain period. The predecessor node is regarded as a failure node when it does not reply, and a broadcast tree will be built with the node as the root node.

Divide all the nodes in a node ring into M groups with a Lg node in each group, Group Leader. The minimum NID node in each group is the Lg node of the group, and it consists of a broadcast sub-tree. When topology change information is found in a node, the information will be sent to the root nodes of M broadcast sub-tree, and will be sent to each node through broadcast sub-tree.

When a node finds the topology change, node will send the event information to all nodes in the node ring, and nodes will update the topology change onto nodes in each group. If nodes send information to all the nodes directly, a large number of bandwidth will be consumed. Due to bandwidth constraints, each node is allowed to send information to M nodes. The discovery node is only responsible to send information to M Lg nodes. Lg nodes will send information to M nodes in its group, and these M nodes will send information to another M nodes, and a broadcast sub-tree will be formed with Lg functioned as root node. A update system based on trees is formed hereby.

In each broadcast sub-tree, topology information is broadcast from root node in the group to each node. Firstly, root node send information to M descendant node. After receiving the time information, M descendant node will keep broadcasting.

Pseudo code of establishing message broadcasting path tree within a group is shown in Figure 3. The GL node in each group is the root node. N (i,j) is used to represent relevant position id in the tree. I
represents the floor of the sub-node, and \( j \) represents the number of the leaf node in the floor.

When a node receives a message to build a tree, it should firstly get the tree coordinates of the node, and then calculate the coordinate \( N_n \) of \( M \) sub-node according to the coordination of the node, as shown in Formula.

\[
N_n(i,j) = (i+1, (j-1) M + n), (1 \leq n \leq M)
\]  

When receiving coordinates of sub-node, calculate \( S_n \) the ordinal number of nodes by coordinate number, as shown in Formula 4-2.

\[
S_n = \sum_{n=1}^{i-1} M^n + j + S_r
\]

\( S_r \) is the ordinal number of root nodes. If the ordinal number is more than \( S_t \), then the serial number is \( S_c = S_n - S_t \), and \( S_t \) is the total serial number, otherwise \( S_c = S_n \).

---

**Figure 3.** Pseudo code for the process of building a message broadcast path tree

```plaintext
For each node get the tree create information do
  Get_tree_id (information from father node);
  For \( n = 1 \) to \( M \) do
    \( N_n(i,j) = (i+1, (j-1) M + n), (1 \leq n \leq M) \);
    \( S_n = \sum_{n=1}^{i-1} M^n + j + S_r \);
    If \( S_n > S_t \)
      \( S_c = S_n - S_t \);
    Else
      \( S_c = S_n \);
      Mark the children node \( S_c \);
      Send information to children node;
  End For
End For
```

---

5. The performance analysis of the algorithm

Different from One-hop, the broadcast path of topology change event message in the Global-hop algorithm is tree-like broadcast. The broadcast topological information in the original algorithm divided nodes into several Slice, and each Slice was divided into several Unit. Each Slice has a Slice Leader, and each Unit also has a Unit Leader with other nodes being ordinary nodes. The message broadcast path is: SL- other SL- descendant or precursor nodes broadcasting in turn, which is a linear broadcast path essentially. The broadcast path in the method designed in this paper is: GL- sub-nodes in each layer, which is a tree-like broadcast way, and can broadcast faster. In the original method, messages should be sent firstly from the nodes of discovery event to SL nodes, and then from SL nodes to other nodes. The event messages pass from SL to UL nodes. 3 hops are needed in the previous several broadcast. The messages broadcast finally from UL node in a linear way. The node number of a Unit is \( U \), so the hop of the time message broadcast all over a Unit is \( U/2 \), the total hops needed are \( 3 + U/2 \), and progressive function set of hops is \( \Theta(n) \).

In the method designed in this paper, messages are sent from the message discovery node to \( M \) GL node directly, and do not need to be sent to other GL nodes through GL nodes in this group. Each node has same routing message, so it can make full use of the information and send it to other GL nodes.
When arriving GL nodes, event messages are sent to each nodes by message broadcast tree and with GL as root nodes. Therefore, the hops needed by Global-hop to broadcast topology change messages are H, as shown in Formula 4-1. The progressive function set of hops is \( \theta \log_a(n) \).

\[
H = \log_a(N - \frac{N-1}{a})
\]  

(3)

Through calculation, the comparison of the two algorithms of message broadcast is shown in Figure 4. The algorithm 1 is Global-hop, and the algorithm 2 is One-Hop. Moreover, with the increase of the number of nodes in the network, the gap will be larger. In terms of broadcast time, the total broadcast time of the event messages of One-hop algorithm is divided into broadcast time and periodic waiting time, so apart from the broadcast time corresponding to each step, the total message delay also includes the periodic waiting time. While Global-hop does not need the periodic waiting time, so its total message broadcast delay only includes the broadcast time each hop corresponding to. Under the circumstances of the same network, the average broadcast time corresponding to each hop is the same, so the message broadcast delay of Global-hop is smaller than one-hop.

**Figure 4.** comparison of message broadcast hops of Global-hop and One-Hop

**References**

[1] Sanaa Sharafeddine, Karim Jahed, Omar Farhat, et al. Failure recovery in wireless content distribution networks with device-to-device cooperation[J]. Computer Networks, 2017, 128(9): 108-122.

[2] Rami Halloush, Hang Liu, Lijun Dong, et al. Hop-by-hop Content Distribution with Network Coding in Multihop Wireless Networks[J]. Digital Communications and Networks, 2017, 3(1): 47-54.

[3] Xiaoying Zheng, Ye Xia. Optimizing network objectives in collaborative content distribution [J]. Computer Networks, 2015, 91(14): 244-261.

[4] Uttam Mandal, Pulak Chowdhury, et al. Energy-efficient networking for content distribution over telecom network infrastructure[J]. Optical Switching and Networking, 2013, 10(4): 393-405.

[5] Lazaros Gkatzikis, Vasilis Sourlas, Carlo Fischione, et al. Low complexity content replication through clustering in Content-Delivery Networks[J]. Computer Networks, 2017, 121(5): 137-151.

[6] Bernd Klasen. Efficient content distribution in social-aware hybrid networks [J]. Journal of Computational Science, Pages 209-218, 2013, 4(4): 209-218.

[7] Michele Mangili, Jocelyne Elias, Fabio Martignon, et al. Optimal planning of virtual content delivery networks under uncertain traffic demands [J]. Computer Networks, 2016, 106(4): 186-195.