Multi-path Routing Policy for Distributed Caching System

Lei Yang*

College of Communication Engineering, Chongqing University, Chongqing, 400044, China

Received: March 23, 2016 Revised: July 16, 2016 Accepted: July 18, 2016

Abstract: The massive volume of content traffic imposes serious challenges to today’s Internet. Content distribution is widely considered as a useful method of efficiently and successfully processing content traffic. To this effect, effective cache technologies are urgently required for both Internet service providers (ISPs) and Internet users. However, transmission efficiency of major methods for content routing is too low to meet Internet’s needs. This paper proposes an ant multi-path routing algorithm in which network state information is with an excepted heuristic factor function to promote the selected probability of the best possible multiple paths to improve routing efficiency. The state information of links works in the proposed algorithm to update the pheromone, allowing multi-paths to be selected with optimized proportions to distribute content appropriately. Simulation results demonstrate that the proposed method reduces access latency by 10%, thus improving the user’s quality of service (QoS) by a considerable margin. Experimental results further demonstrated the superior network transmission capacity of the proposed optimization framework over other existing systems.

Keywords: Content routing, Distributed caching system, Multi-path routing, Routing policy.

1. INTRODUCTION

In recent years, the number of Internet user has increased dramatically not only due to the boom in mobile Internet and computer networks, but also due to the massive growth of user generated content (UGC) [1, 2]. Though IP networking has seen limited success in managing Internet use, content-based networks, such as content delivery networks (CDNs), have flourished despite the enormous volume of content [3]. The dominant trend in development involves networks that no longer facilitate simple data spreading, but rather more sophisticated content access and dissemination.

User and network operator requirements for content distribution have increasing demands as well. Though they are often utilized, intellectualized network optimization technologies, content routing nodes, or routing policy management systems may not be the best choice to accommodate traffic [4].

QoS routing, which represents a critical design issue for both fixed and mobile networks, has been extensively studied to date. Brandt et al. [5], for example, proposed a routing control platform (RCP) that not only provides configuration at the local network scale but also collects information about external destinations and internal topology to select the border gateway protocol (BGP) routes for each router in an autonomous system (AS). In effort to manage resources in the cloud data center appropriately, previous researchers developed the VXLAN controller to manage IP multicasts in overlay networks using openFlow as opposed to dynamic registration protocol [6, 7]. Barazzutti et al. [8] explored an all pair shortest path algorithm to demonstrate the manner in which scalable, lightweight, dynamic graph query mechanisms can be implemented to truncate computation time in the presence of network dynamism.

Content routing policy must be integrated with the distributed policy decision points (PDPs) [9 - 11] to allow information sharing. As a result, distributed PDPs separate the management from the underlying network to better utilize the routing and content resources in each local area.

* Address correspondence to this author at the College of Communication Engineering, Chongqing University, Chongqing, 400044, China; Tel: +8618701240375; E-mails: yanglei0375@163.com, yanglei@hpnl.ac.cn
Researchers around the world are currently searching for methods that optimize content routing quality and efficiency, but are limited by the following persistent disadvantages:

- Default best-effort Internet routing results in the absence of end-to-end QoS.
- Existing routing algorithms primarily focus on router and link factors on a single path and thus do not effectively use the cyber source.
- Routing policy primarily focuses on local features in an individual AS, lacking a network-wide view of topology or traffic.

In effort to remedy these limitations, some researchers have proposed using routers and caches to secure the content that must be distributed for efficient content routing. Jin \textit{et al.}\[12\], for example, established the content-delivery-as-a-service (CoDaaS) technique for distributing UGC and reducing delay to enhance the user experience. Ni \textit{et al.}\[13\] analyzed the corresponding CDN performance and routing overhead between multi-clusters. Miura \textit{et al.}\[14\] proposed Active Anycast, a request routing technique for applied active network technology, to improve either network delay or server processing delay.

This study was focused on problems with multi-path content routing within distributed routing policy management systems. We established an ant-routing algorithm as an optimal multi-path transmission that considers global network state and possible states of traffic in a load-balanced manner. Our primary goals were:

- To build a distributed policy management architecture that collects subnet information and deploys the routing table policy effectively.
- To create a multi-source content routing policy that routes traffic from the content server and cache server to significantly improve client QoS.
- To ensure the PDP uses the most effective policy to obtain maximum profit by judging and weighing routing costs appropriately.

Section II describes the detailed design and implementation of the distributed routing policy management system. In Section III, the ant multi-path routing policy (AMPR) is described in addition to the content policy. Section IV presents our simulation scenarios and results, and Section V concludes the paper and points out future research directions.

2. SYSTEM DESIGN

This section describes the design and implementation of the PDP, which can be used to collect information about end users and topology in an AS. The routing policies generated at the each AS must be highly fine-tuned in order to accurately reflect user equipment, user performance, content stream characteristics, and other important factors.

IP networks are ripe for the introduction of distributed routing policy management systems, where cache servers and routers perform basic content delivery and sharing functions at the behest of routing policies generated by the AS PDP. As shown in Fig. (1), the PDP communicates with each of the routers and end users in the AS; these sessions allow the PDP to not only learn the router configuration and user properties, but also to detect content stream characteristics. The PDP sends different routing policies for each router/switch or each content request from end users. To avoid creating a single point failure, PDP modules must be deployed in each individual AS. Over the entire network, the PDP may exchange inter-domain routing policy information with a neighboring AS. Utilizing the PDP to exchange information across domains enables the entire network’s routing optimization to evolve.

The PDP, the core component of the system, includes the following modules: The information extraction module (IEM), content list, communication module, and calculation module.

The content list, which is comprised of a content list table (CLT) and content routing table (CRT), is mounted by the PDP. The CLT contains information from the cached content in the local cache server. The CRT contains information from the cached content of neighboring cache servers. When the PDP receives a content request, it first searches CLT for a matched cached copy and directs the cache server to replay the content. If CLT fails to match the request, the PDP searches CRT to find the content information in neighboring autonomous domains. If CRT fails to find the matched content, the PDP will ask the original content server to perform the search.
The information extraction module maintains a path status table (PST). Whenever content is sent or received over the link, the path utilization ratio and available bandwidth of each neighboring path are updated. With PST, the PDP can forward the content through the correct link with minimal delay even if traffic loads or network topologies change.

The calculation module, of course, is content routing decision and generation module. This module is used to communicate with original servers, routers, and other cache servers.

3. POLICY DESIGN AND IMPLEMENTATION

The network described here uses PDP, which employs event content for event matching and routing to generate content policies that can be leveraged to improve user QoS. Simultaneous content analysis lends the PDP the ability to control local nodes on a relatively broader scale.

3.1. Ant Multi-path Routing

After receiving a response message, the PDP must determine which content/cache server the content should be routed from. Routing policy PDP, which builds a cooperative policy between the cache server and content server, is typically implemented for this purpose. Non-optimal routing policies may increase the overall duration of content transmission. A robust and efficient cooperative routing policy means that the clients can download desired content from many sources instead of a single, original server. Multi-path load balancing technology should be used in order to ensure that limited network resources are utilized efficiently. On one hand, traffic flow transfers to less-loaded links, minimizing the congestion caused by unbalanced flow distribution; on the other hand, the maximum utilization rate of the path is minimized while the residual bandwidth of the corresponding link is maximized. To this effect, the network has increased ability to accept future arrival requests without having to re-route existing connections.

We established an ant-routing algorithm that works to find the appropriate paths for content. In an ant colony algorithm, each ant releases a pheromone along a path allowing other ants to choose paths with stronger pheromones; these paths are then further amplified by increased pheromones, forming a positive feedback mechanism. The excepted heuristic factor of the path is the probability that an ant will select this path; the greater the excepted heuristic factor, the more likely the path is to be selected. Pheromones evaporate with time so that unused or unpopular paths slowly fade. The ants find their respective destinations effectually through the exchange of information. In the network, clients request data and receive the corresponding content similarly to ants looking for food and carrying the food back to the colony. The CLT and CRT in the PDP can be interpreted as the ants leaving pheromones along their paths.

In this paper, we use \( G = (V, E, B, D) \) to represent an undirected graph with \( n \) nodes and \( m \) paths, where \( V \) is the set of vertices, \( E \) is the set of edges, \( B \) is the set of link bandwidths, and \( D \) is the set of link delays. The terms “content servers”, “cache servers”, “routers”, and “user” represent a vertex in the graph. A content request can be described as \( C = (\text{Src}, \text{Dst}, \text{BW}, \text{DL}) \), where \( \text{Src} \) is the set of the source nodes, \( \text{Dst} \) is the destination nodes, \( \text{BW} \) is the minimum bandwidth of the service command, and \( \text{DL} \) is the minimum delay of the service command. For the available edge \( p \) connected by nodes \( i \) and \( j \) in \( E \), it is important to note the following:

- \( \text{cost} \ (i, j) \) is the communication cost of edge \( p \).
- \( \text{bandwidth} \ (i, j) \) represents the bandwidth of edge \( p \).
- \( \text{delay} \ (i, j) \) is the delay of edge \( p \).
Based on the above definition, multi-path routing is designed to search the required multiple paths in $G$ to connect $Src$ and $Dst$.

Letters $s$ and $d$ define the source node and destination node, respectively. $p_{s,d}(n,i,j)$ represents the probability that the $n_{th}$ ant at node $i$ will choose any specific path to node $j$.

$$p_{s,d}(n,i,j) = \frac{\tau_{s,i}(n,s,d) \cdot \eta_{i,j}(s,d)^\beta}{\sum_{w \in \text{Options}(i)} \tau_{s,i}(n,s,d) \cdot \eta_{i,j}(s,d)^\beta}$$  \hspace{1cm} (1)

$\text{Options}$ is the set of the next hop options available, $\beta$ represents the weighting of heuristic expectation information, $\tau_{i,j}(s,d)$ and is the pheromone information of the path between nodes $i$ and $j$; as mentioned above, the ant tends to choose the path containing more pheromones. $\eta_{i,j}(s,d) = F_{i,j}$ is the heuristic expectation information (which allows the ant to select new links) of the path between nodes $i$ and $j$.

In this paper, $F$ of the path between nodes $i$ and $j$ is defined as follows:

$$F_{i,j} = F_1 \cdot F_2$$  \hspace{1cm} (2)
$$F_1 = \frac{A}{\cos(t(i,j))}$$  \hspace{1cm} (3)
$$F_2 = B \cdot \text{bandwidth}(i,j) \cdot \frac{C}{\text{delay}(i,j)}$$  \hspace{1cm} (4)

$F_1$ reflects the fact that cost influences the path choice, while $F_2$ is the QoS decision factor of path selection.

Pheromones are volatile chemical substances which decrease in strength over time. Again, as an ant moves along its path, the pheromones on that path increase. Meanwhile, pheromones on all paths decrease for the purpose of adjusting the pheromone level and reducing the probability that a path is chosen more than once - i.e., to avoid falling into the locally optimal solution. If the $n_{th}$ ant chooses the path between nodes $i$ and $j$, pheromones can be adjusted using Formula (5); otherwise, they should be adjusted using Formula (6), which eliminates any stagnation caused by the gap between the optimal path and the other paths.

$$\tau_{i,j}(n,s,d) \leftarrow (1 - \alpha) \tau_{i,j}(n,s,d) + \alpha \cdot F(i,j)$$  \hspace{1cm} (5)
$$\tau_{i,j}(n,s,d) \leftarrow (1 - \alpha) \tau_{i,j}(n,s,d) + \beta \cdot \frac{F(i,j)}{F_{\text{chosen}}}$$  \hspace{1cm} (6)

where $0 < \alpha < 1$ is the local pheromone evaporation coefficient and $F$ is the heuristic function of the ant releasing the pheromone.

### 3.2. QoS Routing Policy

Best effort service uses the shortest path to connect sources and destinations, but has two notable disadvantages: First, significant congestion may occur due to limited resources of only one path; second, resources are wasted if many links are idle.

In a complex network, the costs associated with bandwidth, delay, and communication for each path may contain slight differences or even sizable disparities. In the method proposed here, a flow segmentation mechanism allows the better link to encompass a larger proportion of the flow while less favorable links take smaller proportions. If there is a plethora of optional links, the proportional coefficient of the link flow can be determined by each path’s resource priority.

Source nodes $s_1$ and $s_2$ are the content server or cache server in the network, and destination node $d$ requests the content. One of the multiple paths $P_{s,d}^1$ from $s_1$ to $d$ by the way of $(s_1, i_1, \ldots, i_n, d)$ can be described as follows:

$$P_{s,d}^1 \leftarrow p_{s,d}(s_1, i_1) \cdot p_{s,d}(i_1, i_2) \cdot \ldots \cdot p_{s,d}(i_n, d)$$  \hspace{1cm} (7)
Another way of \( P_{sd}(s_2, j_2, \ldots, j_m, d) \) can be expressed as:

\[
P^d_{sd} \leftarrow P_{sd}(s_2, j_2) \times P_{sd}(j_1, j_2) \times \ldots \times P_{sd}(j_m, d)
\]

(8)

In AMPR, optimization is achieved, naturally, by multiple routes. Having more than five paths only slightly improves transmission effectiveness but increases the size of the routing table significantly [15]. Because the AMPR policy finds \( k \) available paths, the top \( K \) paths with the larger \( P_{sd} \) are used here to split the flow into \( K \) proportions:

\[
\text{proportion(i)} = \frac{P^i_{sd}}{\sum_{i=1}^{K} P^i_{sd}} \times 100\%
\]

(9)

\[
K = \begin{cases} 
  k & k \leq 5 \\
  5 & k > 5 
\end{cases}
\]

(10)

After the proportional coefficient is determined, the traffic distributed on path \( i \) is defined as \( T \):

\[
T(i) = \text{proportion(i)} \times \text{content}
\]

(11)

4. EVALUATION

As described above, content distribution efficiency is maximized via multi-path routing. The following sections discuss simulation scenarios, the setting being used, and our conclusive findings based on the experiments [16 - 18].

4.1. Simulation Setting

To test the efficiency and effectiveness of the proposed multi-path routing cache distribution policy, we built a simulation environment using OPNET in Windows i386. OPNET is a network simulation software which can be used to accurately analyze the performance and behavior of complex networks [19]. We then processed the content data through MATLAB R2010a. MATLAB can be used for algorithm development, data visualization, data analysis, and numerical calculation of advanced technology computing language and interactive environments [20, 21]. We utilized three simulation scenarios for the experiment.

Scenario A corresponds to a setting where a request for content is missed in the cache server, so the content server opens a session to reply. The content is distributed by open shortest path first (OSPF) [16, 17]. In Scenario A, when the user requests content, the data is transmitted across the path with the least steps.

Compared to Scenario A, the distribution routing algorithm of Scenario B uses the multi-path routing algorithm equal-cost multi-path routing (ECMP) [18]. In Scenario B, when the user requests content, the data is distributed on average across each path.

The method we propose is Scenario C, in which the ant multi-path routing algorithm delivers the content on each path in different proportions. When the user requests content, multiple paths are selected in optimized proportions to improve transmission effectiveness.

Eleven nodes (N1-N11) with respective storage capacity were deployed on the network topology during the experiment to analyze the content routing capacity and attribute link values, as shown as Fig. (2). To best represent a typical user experience, we deployed a subnet with a user client at the network edge and one server node on the network topology which supplied the original content. The server node and subnet were connected by N1 and N11, respectively.

Simulation parameters are summarized in Table 1.

| Parameter name                  | Parameter value |
|---------------------------------|-----------------|
| Number of Content servers       | 1               |
| Number of users in Subnet       | 100             |
| Cache hitrate                   | 50%             |
| Number of Routers               | 17              |
Bandwidth and delay time were measured in ms and Mbps. In Scenario C, the content was distributed with multiple route methods (ECMP or AMPR) which selects $K$ paths to transfer the requested content. In this section, the path selections between $N1$ and $N11$ of different algorithms are compared.

OSPF computes the shortest path for each route using a method based on Dijkstra's algorithm. ECMP forwards content to a single destination with multiple paths with equal scale. Ant multi-path routing uses the bandwidth and delay time as an expected heuristic function defined by Formula (4). The $P_{i,j}$ can be calculated by the Formula (7) or (8) with the policy proposed above.

The content distributed at each path can be calculated by Formulas (9) and (11). The path selection results of the above three routing methods are listed in Table 2.

| Algorithm/policy | Path Selection | Hops | Transmission ratio |
|------------------|----------------|------|--------------------|
| OSPF             | Path1:N1,N15,N11 | 2    | 100%               |
|                  | Path1:N1,N2,N3,N4,N11 | 4    | 16.7%              |
|                  | Path2:N1,N5,N6,N11   | 3    | 16.7%              |
|                  | Path3:N1,N7,N8,N9,N10,N11 | 5    | 16.7%              |
|                  | Path4:N1,N12,N13,N14,N11 | 4    | 16.7%              |
|                  | Path5:N1,N15,N11     | 2    | 16.7%              |
|                  | Path6:N1,N16,N17,N11 | 3    | 16.7%              |
| ECMP             | Path1:N1,N2,N3,N4,N11 | 4    | 18.9%              |
|                  | Path2:N1,N5,N6,N11   | 3    | 44.1%              |
|                  | Path3:N1,N7,N8,N9,N10,N11 | 5    | 13%                |
|                  | Path4:N1,N12,N13,N14,N11 | 4    | 10.6%              |
|                  | Path5:N1,N16,N17,N11 | 3    | 13.4%              |
| AMPR             | Path1:N1,N2,N3,N4,N11 | 4    | 18.9%              |
|                  | Path2:N1,N5,N6,N11   | 3    | 44.1%              |
|                  | Path3:N1,N7,N8,N9,N10,N11 | 5    | 13%                |
|                  | Path4:N1,N12,N13,N14,N11 | 4    | 10.6%              |
|                  | Path5:N1,N16,N17,N11 | 3    | 13.4%              |

4.2. Simulation Results

We next analyzed the load of the original server to determine whether pressure on the content server was eased under the proposed method. This measured value reflects the optimal routing system for the content provider. User response time and the specified $k$ required to retrieve instant content with all objects contained online were also measured. The smaller these values, the better the system performance.
In this experiment, users requested content halfway through the simulation. Fig. (3) shows user response time for each scenario. When content was delivered in the cache server, users received the majority of the requested content from the local edge content server. In most cases, however, the capacity of the cache server was lower than that of the original server and the bandwidth of the backbone network was larger than those of other networks. Acquiring content from distant cache servers for an extended period was rather inefficient accordingly.

Fig. (3). User response time.

Fig. (4) shows the content send rate of the simulated systems. Content send rate (y axis) within the experiment smoothed with increased simulation time (x axis). All the flows were sent by OSPF in Scenario A and by ECMP in Scenario B; Scenario B exhibited the more efficient content send rate of the two. As shown in Fig. (4), Scenarios A and B showed lower average content send rate than Scenario C.

Fig. (4). Content send rate.

Fig. (5) is a transmission capacity overlay chart for OSPF, ECMP, and AMPR. As discussed above, each of the three represents a different approach to distributing content. OSPF chooses the shortest path, while ECMP equalizes the distribution of content on multiple paths. The path bandwidth, delay, and reliability of links all differ. If all links have the same weight, not only is bandwidth useless, but other disadvantageous effects result - differences in path status, in particular. AMPR maintains real-time adjustment for resource allocation according to path states.
Based on the simulation results, we were able to conclude that the proposed content routing policy produces ideal outcomes. It certainly showed better performance in a wide range of traffic features and network scenarios compared to the other content routing and distribution policies examined here. The AMPR policy can be utilized to disseminate content efficiently between the original server and cache servers to accommodate public demand by routing the content proportionally on different links according to path resources.

As far as noteworthy benefits for network operators and ICPs, the proposed method distributes content preferentially to edge cache servers thus minimizing burst flow in the backbone network and cutting down the server load. Routing content through more than one path increases the efficiency of the network transmission while distributing content for possible future access improves user QoS simultaneously.

CONCLUSION AND FUTURE RESEARCH DIRECTIONS

Our primary goal in this study was to address the challenges of routing efficiency in distributed caching systems. We developed a network state information-based routing policy which uses an ant multi-path routing policy to deliver data.

The primary contributions of this paper can be summarized as follows:

- Our routing optimization policy uses architecture that separates the policy management plane and data plane.
- The AMPR policy also effectively divides content between multiple paths to facilitate efficient transmission.

In the future, we plan to focus on network user behavior in effort to manage content generated in complex networks, and will also investigate routed content size regularity. We intend to deploy the proposed method in a real network as well to investigate its implementation at length.

CONFLICT OF INTEREST

The author confirms that this article content has no conflict of interest.

ACKNOWLEDGEMENTS

The authors would like to thank all project team members for their many insightful contributions to this topic. We would also thank the anonymous reviewers for their thoughtful comments.

REFERENCES

[1] Cisco White Paper, Cisco Visual Networking Index: Forecast and Methodology 2012-2017, Available at: www.itworldcanada.com/assets/cisco-visual-networking-index-forecast-and-methodology-2012-2017

[2] Cisco White Paper, Cisco Visual Networking Index: Forecast and Methodology, 2014-2019, Available at: http://www.cisco.com/
