

Service Management Strategy for CDN

Xinhua E¹⁺, Binjie Zhu², Hui Zhang² and Yanjun Shi²

¹Faculty of Information Technology, Beijing University of Technology, Beijing, 100124, China

²China Mobile Group Beijing Company Limited, Beijing, 100007, China

Abstract. A content distribution network is a kind of overlay network to improving Internet access speed for clients. The management of cache server of content distribution network is an important issue. This paper studies the edge server resource management strategy of streaming media service in the multi-service CDN. A management strategy was proposed in this paper. It can manage the edge server resources in the effective under the premise of guaranteeing QOS, and improve resource utilization in the CDN.

Keywords: CDN, Overlay network, service management

1. Introduction

Content distribution network is a kind of overlay network, which was used to improve the user access latency [1]. The traditional Internet access model is the C / S access patterns, in where the user directly access to a server directly from the servers to obtain the requested contents. C / S model has many shortcomings. When the load of the server was high, the server will be congestion. The user was far from the server, resulting in greater user response delay.

A layer of edge servers were added in content distribution network between the users and the servers [2]. Contents were distributed through a certain algorithm to be sent to the appropriate edge server. When client requesting a content, the request were redirected to the nearest edge server. If the edge server has the content requesting, then the edge servers service the client. If it has not the content, the request was redirected to other servers until it finds the content of user requests.

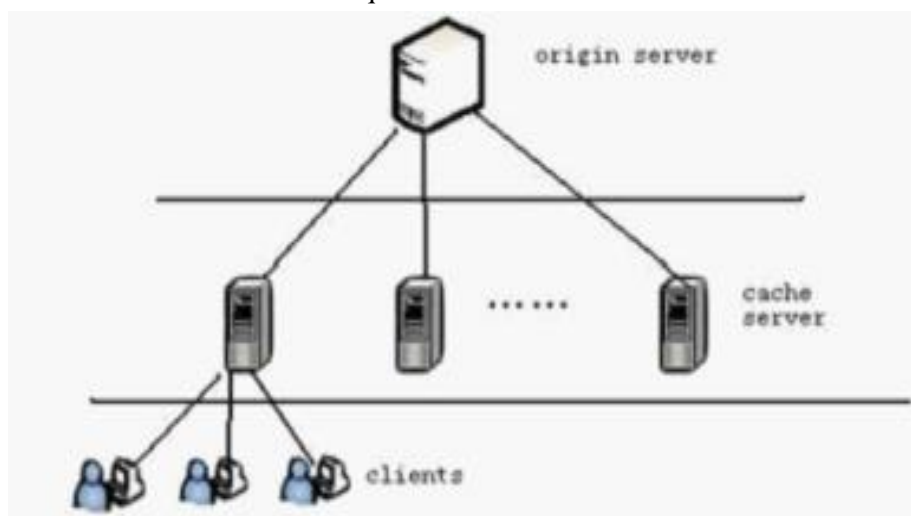


Fig. 1: CDN

⁺ Corresponding author. Tel.: + 86 13811324402
E-mail address: Elson1900@163.com

2. Background and Motivation

The CDN architecture is one of the key issues for content delivery network. Different network architecture needs to be designed based on different delivery demands. The network architecture can directly affect content delivery, request routing and other strategies.

P2P-based content delivery network can use three fusion ways, namely, direct overlay fusion architecture, core architecture and complete peer-to-peer architecture.

The design concept of overlay fusion architecture is to overlay the loose coupling of the CDN and P2P system. Generally, the overlay fusion architecture is divided into two layers, namely, CDN and P2P, of which, the CDN system architecture is divided into two layers. Therefore, this is a three-layer architecture including the CDN management layer, the CDN edge server layer and the user P2P layer. The CDN management layer manages the entire system. The CDN edge server layer is at the edge of the network, and the CDN edge server pushes contents to the network edge near the users by caching content objects. Several user nodes close to an edge server form a peer-to-peer network, generally an unstructured network. The user nodes not only can obtain content objects from peer nodes, but also can access contents objects from CDN edge server nodes.

The nodes in a CDN in a complete peer-to-peer architecture can be turned into peer nodes by two ways. CDN is generally divided into management layer and core functional layer. The management layer is responsible for optimizing and managing the function of the entire CDN system. The core functional layer is responsible for realizing specific content storage, service, search and other core functions. The first method is to turn all nodes in the management layer into peer nodes in the P2P network, that is, both of nodes in the management layer and nodes in the functional layer are peer nodes in the P2P network. The second method is to distribute the functions of the CDN management layer into nodes at each core layer. In this way, each node has a part of management functions and the management function can be realized through the cooperation of the management node modules in each of the peer nodes. The first method is complete peer-to-peer at the network level and the second one is complete peer-to-peer at the function level.

The core transformation architecture constitutes the P2P layer of the core functional layer through the P2P transformation of the core functional layer of CDN. In this architecture, the management mode of the management layer is not changed. Instead, only peer-to-peer of the core functional layer is implemented. The goal of peer-to-peer is to improve the overall performance by enhancing collaboration at the core functional level. This approach does not require the participation of the clients. The clients only need to make a content request and access content services, so that the overall network resources can be centrally managed and billed.

P2P-CDN is a special CDN, introduced the idea of P2P into CDN. CDN's caches were organized by P2P. It can share resources between caches. Figure 2 is a CDN, the principle: the user redirects the user's request through the system-oriented nearest cache server. As shown in figure 2. The user's request is redirected to the nearest edge server redirected system. If the edge server has the request contents, it will response the request. If the edge server has not the request content, the edge server finds the request resources and pulls to the local server. The edge servers form a resource sharing network [3].

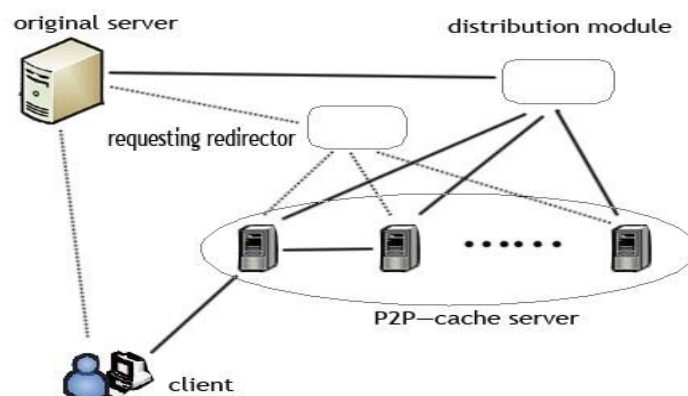


Fig. 2: Multi-service P2P-CDN

In order to ensure the QOS of cache server, most of streaming media CDN will limit the maximum number, and then request as a queue according to the rules FCFS (First Come First Servers). When a user requests occupy server resources, the server will always be for this user until users quit the requests. Management strategy for the request take a video file as a full-service, there are two problems. First, the resources were limited, while the maximum number of users in services is limited .Second, the speed that the content download from the server faster than the speed of the user view, so when the user has the exit behavior, the download the file at this time far beyond the length of time users need to watch. So it is waste of resources of the server.

This paper designs a server resource management strategies, the performance be enhanced in two ways: First, in the case of the same server resources, enhance the number of simultaneous users of the service. Second, the services were effective management avoiding the waste of server resources.

3. CDN Resource Management Strategy

The access frequency between each section within a program is a difference. Client typically demand weather continue a movie after looking at the beginning of time, and according to the contents of it form an overall evaluation of the film, and then decide whether to continue . The client’s behavior characteristics analysis in paper [4], described by the probability of a user exit during the viewing behavior.

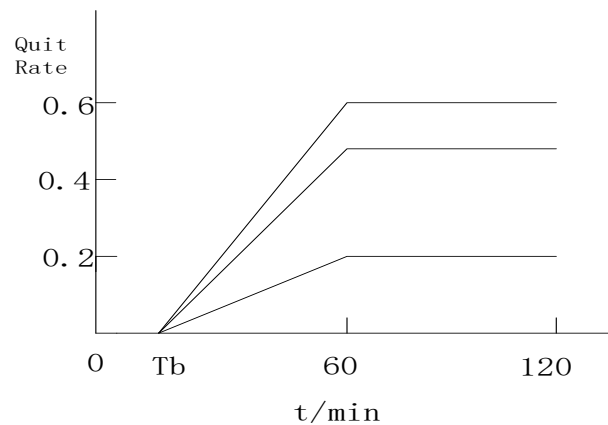


Fig. 3: Quit rate of a content

The download speed was taken as the service speed, the watching speed as the speed of service consumption. Service speed is great than service consumption rate. If take a video file download as a service, we can design a strategy for the server to efficiently manage bandwidth resources.

In order to ensure the QOS of cache server, most of streaming media CDN will limit the maximum number, and then request as a queue according to the rules FCFS (First Come First Server). When a user requests occupy server resources, the server will always be for this user until users quit the requests.

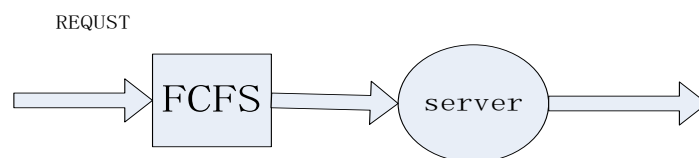


Fig. 4: First come first service

The basic idea of this patent is to create a new rule to manage the request queue on the server resources. Suppose there are R requests coming, the rule is that each request queue was divided into N sub-requests. The first M sub-requests were selected from the R queuing requests as a group. Take the first sub-requests from the first request in the queue group, and then the first sub-requests from the second request in the queue group. And then traverse from 1 to M, taking the second ub-request. Until the last traverse 1 to M, remove the last N sub-requests. Handle the request in accordance with the queuing algorithm to form a new queue, and the server's queue according to the new service.

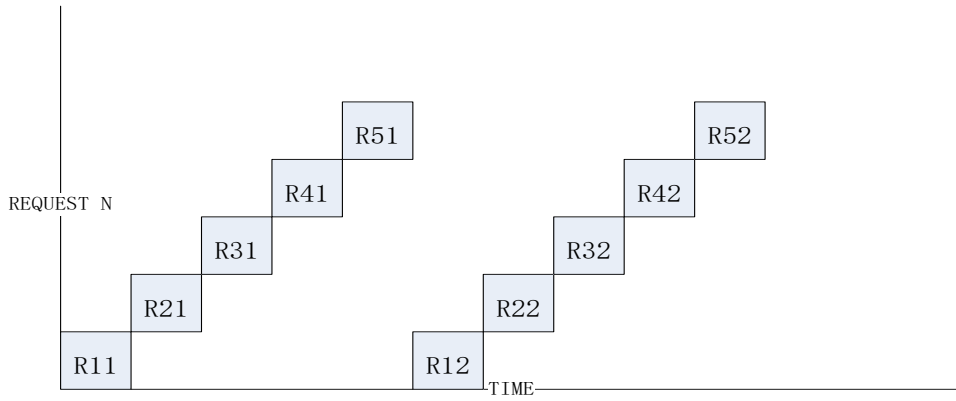


Fig. 5: First come portion first service

4. Analysis

For contrast analysis, a scenario was supposed. The large of all the video files were 120 Minutes. The storage spaces for every minute file were 1MB. The download speed is 0.1MB/S. The download time for one video is 12 Minutes.

The service time of FCFS strategy is 12 Minutes. The average service time of FCPFS strategy is 6.6 Minutes.

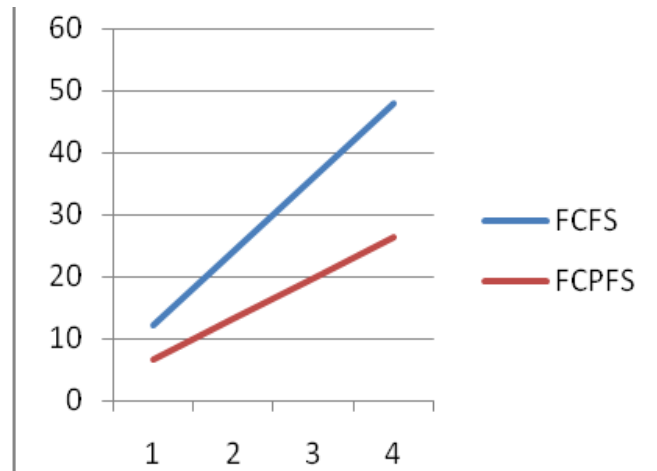


Fig. 6: Service time

Each edge server was taken as M|M|n|0 queue system with arrival rate 0.2. The number of the server in queue system is 2. FCFS strategy as situation A. FCPFS as situation. In the situation A,

$$\rho_0 = \frac{1}{1 + \rho + \frac{\rho^2}{2!}} = 0.158 \quad (1)$$

$$\rho_1 = \rho \rho_0 = 2.410 * 0.158 = 0.381 \quad (2)$$

$$\rho_2 = \frac{\rho^2}{2!} \rho_0 = 0.459 \quad (3)$$

The contrast of Pr as figure 7. The contrast of Q as figure 8.

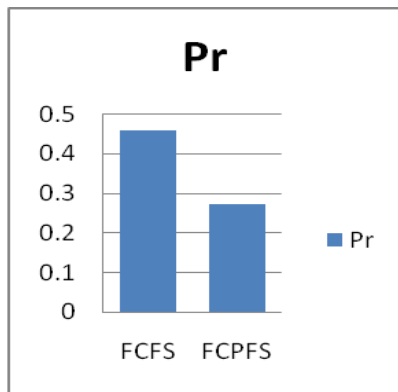


Fig. 7: Pr of strategy

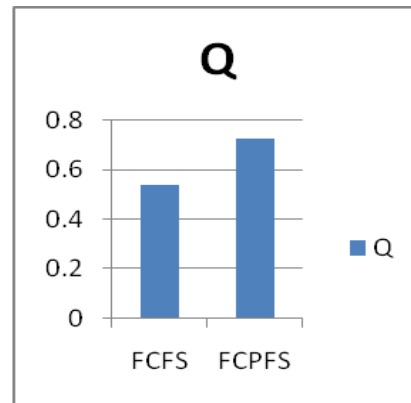


Fig. 8: Q of strategy

5. Conclusion

The management of cache server of content distribution network is an important issue. This paper studies the edge server resource management strategy of streaming media service in the multi-service CDN. A management strategy was proposed in this paper. It can manage the ES resources in the effective under the premise of guaranteeing QOS, and improve resource utilization in the CDN. Form the contract, we can conclusion that the proposed strategy was effective.

6. 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.