2006년도 대한전자공학회 하계종합학술대회 제29권 제1호

클라이언트 기반 분산 웹캐싱 시스템

박종호*, 유성구**, 정길도***
전북대학교 공과대학 전자정보공학부

A Client-based distributed web caching system

Jong Ho park*, Sung Goo Yoo**, Kil To Chong
Department of Electronics and Information Engineering
Chunbuk National University
E-mail : *jhpark@chonbuk.ac.kr, **ding5@chonbuk.ac.kr, ***kitchong@chonbuk.ac.kr

Abstract

A distributed web caching system can transmit information to a user quickly and stably, avoiding a congested internet network by storing and later supplying requested content to a cache that is distributed and shared like a proxy server. This paper proposes a client-based distributed web caching system that assigns an object and controls the load using a user's direct connection to shared caches, without the aid of additional domain name system (DNS) requests. The proposed system simplifies information transmission by reducing both DNS queries and delay time.

I. Introduction

Network bottlenecks and too many requests for a server may increase delays in downloading information and possibly result in failure of the system. To address these issues, many studies have focused on reducing total delay time by avoiding network bottlenecks, distributing the requests for a specific web server among caches near users' clients, and then supplying this cached information to clients. These studies can be classified into two categories: those that use a communication method according to object assignment and a searching method between shared caches [4-7], and others that apply a hash function [1, 3, 8-10], called a hash routing method. Communication methods assign a single configured cache based on the characteristics of the client in the system. Thus, it is not able to use the storing space effectively because requested objects may exist among shared caches. This may increase inquiry delay time, lower system performance, and waste network resources.

CARP controls loads in shared caches by manipulating the time to live (TTL) value of the domain name system (DNS), and optimizing the assignment of objects to caches. Similarly, consistent hashing requires the DNS to assign objects and control loads. Therefore, these methods require multiple DNS queries (hereafter, "additional DNS") in a distributed web caching system. Such a system, operated in an average web environment, using DNS to find an object in DNS, results in a delay time between inquiry and response, while the DNSs communicate. In other words, CARP [1] and consistent hashing [3] have a complex domain delegation for DNS that grows, and the system becomes increasingly complex, further increasing DNS delay time.

If a client can be changed to find a cache directly, DNS use will be reduced, thereby decreasing delays. Moreover, it is

.

possible to configure a simple system that does not require additional DNS in the CARP or GHS distributed web caching systems. This would also reduce DNS delay time and thus improve the performance of distributed web caching systems. This paper develops an alternative client-based distributed web caching system that does not require additional DNS for CARP [1] and consistent hashing [3].

II. Implementation Results

The experimental system was configured to compare consistent hashing to 2HRCS on the private network. Figure 1 shows the average delay time of the hash routing scheme and proposed method with respect to object size. Although similar, the proposed system performed slightly better. Figure 2 shows the hit ratio of each system. The proposed scheme completed the service with a lower delay time and a better hit ratio. By analyzing the DNS log files of each system; we verified a more than 90% reduction in DNS queries using 2HRCS, as compared to the other systems.

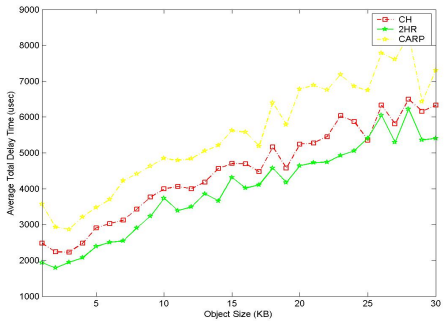

Figure 1. Average delay time in each system.

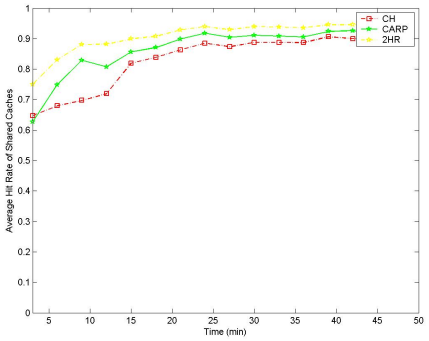

Figure 2. Hit ratio of each system.

Ⅲ. Conclusions

This study has evaluated the distributed web caching systems CARP and consistent hashing, and has proposed a simplified client-based system that does not require additional DNS queries. We verified the proper operation of the system by implementing it in a real local network. It outperformed both the consistent hashing scheme and CARP with regard to delay time and hit ratio. In conclusion, the advantage of the proposed scheme lies in the simplicity of the system and the shorter delay time of service.

References

[1] Valloppillil, V. et al., Cache array routing protocol v1.1. Internet Draft, http://www.globecom.net/ietf/draft/draft-vinod-carp-v1-3.html (1998).

[2] Leighton, et al., Global hosting system. US Patent 6108703,http://www.delphion.com/cgibin/viewpat.cmd/US06108703, (2000).

[3] Karger, D. et al., "Web caching with consistent hashing," In Proceedings of the 8th International World Wide Web Conference (1995).

[4] Chankhunthod, A. et al., "Hierarchical internet object cache," In USENIX (1996).

[5] Malpani, R. et al., "Making world wide web caching servers cooperate," In Fourth International World Wide Web Conference, (1995), 107–110.

[6] Gadde, S. A. et al., "A taste of crispy squid," In Workshop on Internet Server Performance, http://www.cs.duke.edu/ari/cisi/crisp (1998).

[7] Fan, L. et al., Summary cache: A scalable wide-area web-cache sharing protocol, Technical Report 1361, Computer Science Dept., University of Wisconsin (1998).

[8] Wessels D. et al., Internet cache protocol version 2, Internet Draft, http://icp.ircache.net/rfc2187.txt (1997).

[9] Valloppillil V. et al., Hierarchical HTTP routing protocol. Internet Draft, http://icp.ircache.net/draft-vinod-icp-traffic-dist-00.txt (1997).

[10] Thaler D. G. et al., Using named-based mappings to increase hit rates. To appear in IEEE/ACM Transactions on Networking (1997).