

iSCSI를 사용한 계층적 웹 캐싱 스키의 설계

동서대학교 컴퓨터공학전공
htlim@dongseo.ac.kr

Design Considerations for Hierarchical Web Caching Scheme Using iSCSI

Lim, Hyotaek
Dept. of Computer Engineering, Dongseo University
htlim@dongseo.ac.kr

Abstract

The sharing of caches among Web proxies is an important technique to reduce Web Traffic and alleviate network bottlenecks. Additionally, due to emerging network technologies cooperative Web caching among proxies shows great promise to become an effective approach for reducing Web document access latencies. Nevertheless it is not widely deployed due to the overhead of existing protocols such as ICP. We propose iSCSI-based hierarchical web caching scheme which provides more improved performance than existing web caching scheme.

I. 서 론

웹 서비스의 이용이 폭발적으로 증가함에 따라 캐싱은 대역폭의 손실을 줄이기 위한 중요한 기법의 하나로 간주되어 왔다. 특별히 웹 프록시내의 캐싱은 매우 효과적인 것으로 나타났다. 캐싱의 이점을 충분히 얻기 위해 프록시 캐쉬들은 서로 협력할 필요가 생기게 되었는데 이를 위해 협력(Cooperative) 웹 캐싱 기법이 등장하게 되었다. 협력 웹 캐싱은 세가지 부류로 구분될 수 있는데 ICP(Internet Caching Protocol)와 같은 메시지 기반, CARP(Cache Array Routing Protocol)와 같은 해쉬 기반 그리고 Summary 캐쉬와 같은 디렉토리 기반이다. ICP는 간단한 질의 메시지를 사용하여 프록시간에 정보를 주고받기 위한 프로토콜로서 UDP 상에서 동작한다 [7, 10, 11].

협력 웹 캐싱 스키의 대안으로서 본 논문은 상·하위 레벨의 프록시간에 iSCSI(Internet SCSI)를 사용하는 새로운 계층적 웹 캐싱 스키를 제안한다. iSCSI는 IP 네트워크 상에서 블록 데이터 전송을 위한 새로운 기술로서 현재 IETF에 의해 국제표준으로 제정되었다. iSCSI는 하위 계층으로서 TCP를 사용하기 때문에 IP 네트워크 상에서 신뢰성 있는 스토리 트래픽의 전송을 제공한다. 또한 iSCSI는 원격에서도 일관성 있는 스토리지의 접근이 가능하며 네트워크 관리기능과 같은 기존 IP 분야에서의

연구결과를 그대로 이용할 수 있다[6].

본 논문은 다음과 같이 구성되어 있다. 2장은 관련 연구 및 계층적 웹 캐싱 스키의 개요에 관하여 설명한다. 3장은 iSCSI 기반의 계층적 웹 캐싱 스키를 제안한다. 이장에서 새로운 캐싱 절차와 프록시 서버 구조를 설명한다. 마지막으로 4장은 본 논문을 요약하며 결론을 맺는다.

II. 관련 연구 및 계층적 웹 캐싱

2.1 관련 연구

iSCSI는 최근에 만들어진 스토리지 프로토콜이므로 iSCSI 관련 연구가 많이 이루어지지 않았으나 대표적인 연구를 선별하면 다음과 같다. He 등은 인터넷 기반에서 스토리지의 성능을 개선하기 위해 SCSI-To-IP 캐쉬 스토리지(STICS)를 개발하였다[6]. 성능 측정 결과에 의하면 STICS는 평균 시스템 처리량(throughput)의 53~78%까지 그리고 응답시간의 90%까지 성능이 향상되었음을 보였다. Voruganti 등은 세 가지 SAN 종류인 iSCSI, Infiniband 그리고 파이버 채널을 분석하였다[5]. 이 논문은 SAN이 WAN 상에서 이용될 때 iSCSI가 파이버 채널과 Infiniband보다 우수함을 지적하였다. Nishan도 최대 iSCSI 처리량이 Nishan 제품으로 어떻게 도달될 수 있는지에 관한 iSCSI 성능을 분석하였다[3]. 이 성능시험은 iSCSI 제품이 단일 iSCSI 포트상에서 최

대의 회선속도(219.64Mbps)에 도달 될 수 있음을 보였다.

반면에 웹 캐싱에 관한 연구는 상당히 오랜 기간동안에 많이 이루어졌다. 무엇보다도 Danzig 등은 인터넷상에 웹 객체를 위한 계층적 객체 캐시의 필요성을 언급하였다. 이들은 외부로의 FTP 트래픽이 FTP 캐시를 통해서 42% 줄일 수 있음을 보였다[1]. 이러한 결과는 계층적 웹 캐시, Harvest 객체 캐시를 설계하는 동기가 되었다. 본 논문에서 제안된 스킴은 계층적 웹 캐싱에서 iSCSI 프로토콜을 이용한다.

2.2. 협력 웹 캐싱 개요

협력 웹 캐싱은 인터넷에서 많은 통신을 하는 프록시 캐시들 사이의 캐시 되어진 웹 문서의 공유 및 협조를 의미한다[4, 7]. 데이터의 협력 캐싱은 높은 전송 속도의 랜 환경에서 분산되어진 파일과 가상 메모리 시스템에 그 근원을 두고 있다. 협력 캐싱은 대체로 분산되어진 컴퓨팅 시스템에서 지연시간을 줄여준다. 그 이유는 네트워크 전송 시간은 오류가 발생하는 디스크 접근 시간보다 더 적기 때문이다. ISP(Internet Service Provider)에 의해 개발되어진 이전의 IP 네트워크 백본망의 대역폭이 일반적으로 34Mbps로 제한되어져 있는 반면, 현재 그리고 미래의 IP 네트워크는 155Mbps에서 2.4Gbps 범위의 대역폭을 제공한다. 따라서 프록시 캐시 사이의 협력 웹 캐싱 네트워크 기술의 등장은 웹 문서의 접근 지연시간 감소에 효율적으로 이용될 수 있다.

캐시들을 협력하기 위한 한 방법은 캐싱 계층구조(hierarchical structure)를 설정하는 것이다. 계층적인 캐싱은 다음과 같이 동작한다. 계층구조의 하위 레벨에 클라이언트 캐시가 있다. 요구되어진 문서가 클라이언트 캐시에 존재하지 않으면, 그 요구는 말단기관(institutional)의 캐시에게로 향한다. 만일 문서가 말단기관 캐시에서 발견되지 않으면 요구는 다시 지역적인(regional) 캐시에게로 향한다. 그 다음은 국가적인(national) 캐시에게로 향한다. 만일 그 문서가 어느 캐시 레벨에서도 발견되지 않으면 국가적인 캐시는 웹 서버에게 직접적으로 요구한다. 그 문서가 캐시나 웹서버에서 발견되어지면 그 문서는 계층적으로 내려오면서 캐시 각각에 즉시 그 문서를 복사한다. 따라서 그 문서가 어떤 캐시 레벨에서 발견되어지기 전까지 같은 문서에 대한 요구는 캐싱 계층구조의 상위로 올라가면서 검색되어 진다.

협력 캐싱은 이미 인터넷의 많은 곳에서 사용되어 지고

있다. 인터넷에 연결되어진 많은 ISP들이나 기관들은 대역폭 감소와 클라이언트의 지연시간을 감소하기 위해 캐시들을 설치하였다.[2, 8] 그러나, 캐싱 계층구조에는 몇몇의 문제점이 있다. (i) 모든 계층구조의 레벨은 추가적인 지연이 발생할 수 있고 (ii) 더 높은 레벨의 캐시는 병목현상과 대기시간의 오랜 지연도 발생할 수 있으며 (iii) 많은 문서의 복사본이 각각의 캐시 레벨에 저장되어 진다. 다음 섹션에서 이 문제들을 해결하기 위해 iSCSI 기반의 계층적인 웹 캐싱 스킴을 제안한다.

III. iSCSI 기반의 계층적 웹 캐싱 스킴

3.1 개요

iSCSI 기반의 계층적 웹 캐싱 스킴은 상·하위 레벨 프록시 서버간에 iSCSI 프로토콜을 사용한다는 점에서 종전의 캐싱 스킴과 다르다. 제안된 캐싱 스킴에서 iSCSI는 원격의 상위 레벨 스토리지가 하위 레벨의 프록시에 가상으로 마운트 되는 것을 가능하게 한다. 그림 1은 제안된 스킴의 전체적인 개요를 보여준다.

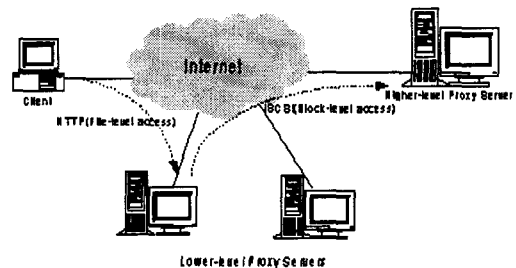


그림 1 iSCSI 기반 캐싱 스킴의 개요

클라이언트는 파일-레벨의 HTTP 요구를 로컬 프록시 서버에게 보낸다. 로컬 프록시 서버가 클라이언트로부터 요구를 받을 때 먼저 자신의 캐시를 조회한다. 만일 프록시가 요구받은 데이터를 자신의 캐시에 저장하고 있으면 직접 클라이언트로 그 데이터를 보낸다. 만일 요구받은 데이터가 로컬 프록시와 이웃(neighbors)에 존재하지 않으면 로컬 프록시는 상위 레벨의 프록시에게 데이터 요청을 하는데 이때 기존의 HTTP 대신에 블록-레벨의 iSCSI 프로토콜을 사용한다. 상세한 프로시듀어는 다음절에서 설명된다.

3.2 프로시듀어

상세 프로시듀어는 그림 2와 그림 3에서 보여주고 있다. 임의의 객체가 발견되는 위치에 따라 두

가지 경우가 존재하는데 객체가 로컬 프록시와 같은 레

별에 발견되는 케이스(케이스 1)와 상위레벨에서 발견되는 케이스(케이스 2)이다.

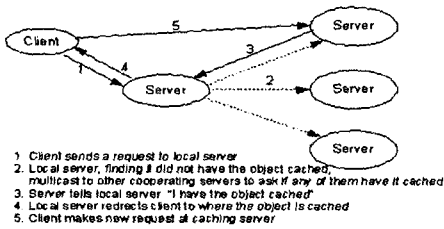


그림 2 iSCSI 기반의 프로토콜 절차(케이스 1)

그림 2는 객체가 로컬 프록시의 이웃중 하나에서 발견되는 경우의 프로시더를 보여주고 있다(케이스 1). ICP는 로컬 프록시와 그 이웃 간에 사용되는 프로토콜로서 사용될 수 있다. ICP는 UDP/IP 상에 동작이 되므로 HTTP 보다 상대적으로 간단하고 빠른 응답을 제공한다.

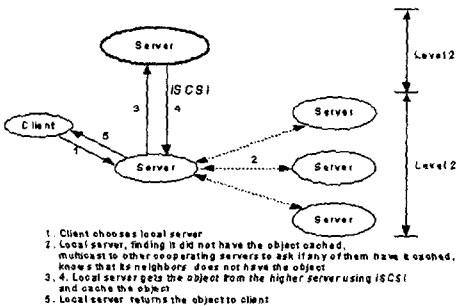


그림 3 iSCSI 기반의 프로토콜 절차(케이스 2)

그림 3은 객체가 로컬 프록시의 이웃에서 발견되지 않는 경우로 원하는 객체는 iSCSI를 통해 상위 레벨 프록시로부터 가져오게 된다. 위의 그림에서 보여진 바와 같이 iSCSI 기반의 캐싱 절차는 다음과 같다.

1. Client randomly picks a local proxy, say A, from a group of proxies
2. It sends the request to the proxy A
3. If proxy A has the object cached then
 - It replies back to the client with the cached copy
- Else
 - It multicasts to other cooperating servers using ICP about the client's request
 - If a cooperating server, say B, has the object then
 - It sends an ACK to proxy A
 - Proxy A sends a response to client to contact proxy B
 - Client contacts proxy B
 - Proxy B sends cached object to client
 - Else if none of the cooperating servers also have the object
 - It fetches the object from the higher-level proxy using iSCSI

- Caches a copy and replies to client with the newly retrieved object

HTTP는 persistent 연결을 지원한다고 할지라도 iSCSI는 이미 확립된 TCP 연결을 사용한다는 점에서 HTTP보다 개선된 성능을 제공한다. 즉, 하위 레벨의 프록시는 TCP 연결 확립시간 없이 iSCSI를 통해 상위레벨로 요구를 보낼 수 있다. 실제의 환경에서 상·하위 레벨의 프록시시간에 TCP 연결을 위한 왕복 지연시간(round trip time)은 무시할 수 없을 정도로 큰데 그 이유는 이들 프록시시간 WAN상에 멀리 떨어져 있을 수 있기 때문이다.

3.3 프록시 서버 구조

일단 프록시 서버가 iSCSI를 사용하여 원격의 상위 스토리지를 마운트하게 되면 프록시는 HTTP와 iSCSI간을 연결하기 위한 더 이상의 모듈도 필요하지 않다.

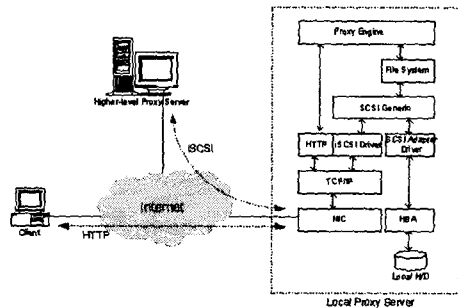


그림 4 프록시 서버 구성요소

그림 4는 프록시 서버의 전체적인 구조를 보여주고 있다. 프록시 서버가 클라이언트로부터 요구를 받을 때 프록시 엔진은 파일 시스템을 통하여 로컬 디스크에 데이터가 있는지 체크한다. 만일 요구받은 데이터가 로컬 프록시와 그 이웃에 없으면 iSCSI 프로토콜을 통하여 상위 레벨 프록시와 통신한다. 프록시 엔진이 상위 레벨의 프록시 서버에게 요구를 보내고자 할 때 파일 시스템을 통하여 iSCSI 드라이버를 활성화한다. 이때 iSCSI는 인터넷상에서 블록-레벨 스토리지 데이터의 전송이 가능하게 하며 로컬 프록시는 가상으로 로컬 프록시에 마운트되어 있는 상위 레벨 스토리지로부터 데이터를 얻게 된다.

IV. 결론

본 논문은 상·하위 레벨의 프록시 서버간에 iSCSI 프로토콜을 사용하는 iSCSI 기반의 계층적 웹 캐싱 스킴을 제안하였다. 제안된 스킴은 iSCSI 프로토콜의 사용으로 상·하위 레벨의 프록시시간에 이미 확립이 된 TCP

연결을 사용한다는 점에서 개선된 성능을 제공한다. 실제의 인터넷 환경에서 임의의 프록시와 상위 레벨의 프록시간의 TCP 연결확립을 위한 왕복지연 시간이 무시할 수 없을 정도로 큰데 이는 대부분의 프록시 서버들이 인터넷과 같은 WAN상에 존재하기 때문이다.

본 논문의 저자는 제안된 스킴에서 사용될 수 있는 네가지 메타-데이터 갱신 기법을 시뮬레이션에 의하여 분석하였으나 제한된 논문 페이지로 인해 성능분석의 결과를 포함시킬 수 없었다. 현재 실제의 테스트베드를 구성하여 iSCSI 기반의 웹 서버를 기존의 웹서버와의 성능을 비교하는 작업을 수행하고 있다.

참 고 문 헌

- Networking, Vol. 8, No. 3, June 2000
- [10] Keith W. Ross, Hash Routing for Collections of Shared Web Caches, IEEE Network, Nov. 1997
- [11] Microsoft, Microsoft Proxy Server: Cache Array Routing Protocol and Microsoft Proxy Server 2.0, White Paper, 1997
- [1] Anawat Chankhunthod , Peter Danzig, A Hierarchical Internet Object Cache, USENIX Annual Technical Conference, 1995
- [2] P. Rodriguez, C. Spanner, E.W. Biersack, Web Caching Architecture: Hierarchical and Distributed Caching, Proc. of the 4th International Caching Workshop, San Diego, Apr. 1999
- [3] Nishan Systems, Achieving Wire-Speed iSCSI Performance, Technical Note, 2002
- [4] Alec Wolman, Geoffrey M. Voelker, et al, "On the Scale and Performance of Cooperative Web Proxy Caching" , ACM Symposium on Operating Systems Principles, Dec. 1999
- [5] Kaladhar Voruganti, Prasenjit Sarkar, "An Analysis of Three Gigabit Networking Protocols for Storage Area Networks" , 20th IEEE Intl. Performance, Computing, and Communications Conference (IPCCC), April 2001
- [6] Xubin He, Qing Yang, "Architecture and Performance Potential of STICS-SCSI-To-IP Cache Storage" , 2001
- [7] Christoph Lindemann, Oliver P. Waldhorst, "Evaluating Cooperative Web Caching for Emerging Network Technologies" , Workshop on Caching, Coherence and Consistency, Italy, June 17, 2001
- [8] Jia Wang, "A Survey of Web Caching Schemes for the Internet" , ACM CCR Vol. 29, Nov. 1999
- [9] Li Fan, Pei Cao, et al, Summary Cache: A Scalable Wide-Area Web Cache Sharing Protocol, IEEE/ACM Transactions on