

멀티미디어 콘텐츠 전달 및 분산(CDD) 시스템의 성능 평가

이영석, 윤창일, 이용우
서울시립대학교 전자전기컴퓨터공학부

e-mail:sok2000@netian.com, cruise45@hanmail.net, ywlee@uoscc.uos.ac.kr

Performance Evaluation of a Contents Delivery and Distribution(CDD) System in Intranet

Young-Sok Lee Chang-Il Yoon Yong-Wo Lee
Department of Electrical & Computer Engineering, The University of Seoul
90 Cheonng-Dong, Dongdaemun-gu Seoul S.Korea
Phone: (02) 2210-2827

요 약

최근에 인터넷을 통한 WWW(World Wide Web)이 전 세계적으로 널리 사용되고 있다. 따라서, 매년 인터넷을 사용하는 사람의 수도 매우 빠른 속도로 증가하고 있다. 그러나, 인터넷이나 인트라넷 모두, 현재의 네트워크 용량으로 모든 사용자들의 요구를 적절하게 수용할 수 없다. 따라서, 학계와 산업계의 전문가들이 WWW 개성에 많은 노력을 하고 있다. 적절히만 구성된다면, 웹 캐싱 시스템은 현저한 네트워크 대역폭의 사용량 감소를 가져올 수 있고, 네트워크 레이턴시를 줄일 수 있다. 이것은 VOD(Video On Demand)의 경우와 같은 멀티미디어 스트리밍 데이터 전송 환경에서도 적용된다. 본 연구에서, 인트라넷 환경에서 멀티미디어 데이터를 전송하는데 가장 비용이 적게드는 최적의 솔루션을 찾으려한다. 이를 위해서, 테스트환경을 구축하고, 실험적 측정을 수행하였다. 그리고 그 결과를 토대로 여러 가지 현상에 대해서 분석하였다.

Keywords: CDD, Caching, Intranet

1. 서 론

최근 웹(World Wide Web)은 학교, 대학, 공공기관, 기업, 일반인 등에게 경쟁력을 제공하는 매우 중요한 도구가 되었다. 모든 데이터의 공유, 통신, 소식을 알리는 일, 물건을 사고 파는 일, 심지어 은행의 금융서비스까지 다양하게 이용되고 있는 실정이다.

그런데, 이렇게 좋은 환경을 제공하는 웹도 여러 가지 문제가 있다. 그 중에서도 웹의 가장 큰 문제점은 너무 느리다는 것이다. 단일 웹 페이지를 다운받는데 몇 초에서 심지어는 몇 분까지 걸리는 경우도 있다. 이는 웹을 사용하는 사용자의 증가와 더불어 웹에서 오고가는 데이터 타입의 변화에 기인한다. 한국 전산원[1]에 따르면, 2000년 6월에 집계된 인터넷 이용자 수(이는 최근 3개월 이전에 인터넷을 사용한 경험이 있는 자)는 1400만 명에 이르는 것으로 나타났다. 이는 전체 가구원 3900만 명중에서 36%의 비중을 차지한다. 또한, 인터넷 데이터 타입이 기존의 문서를 기반으로 하던 매우 작은 량의 데이터에서 크기가 매우 커진 멀티미디어 데이터(Multi Media Data)로 전환되고 있는 실정이다. 이처럼 사용자가 늘어나고, 데이터 타입이 대용량의 멀티미디어 데이터로 변화하면 네트워크를 사용하는 대역폭자체의 사용량이 늘어나게 되었다. 이렇게 네트워크의 사용량이 늘어나면 네트워크의 전송속도는 자연히 느려질 수밖에 없다.

많은 사람들이 이처럼 느려진 네트워크를 좀더 빠

르게 하기 위하여 많은 연구를 거듭하고, 여러 가지 방법을 제시하고 있다. 가장 간단하게 생각할 수 있는 방법은 성능이 좋은 컴퓨터를 사용하면, 네트워크 대역폭을 증설하는 공사를 단행하는 일이다. 하지만, 네트워크 대역폭의 증설은 비용이 많이 들뿐만 아니라 이것을 증설하는 데는 한계가 있기 때문에 이 문제를 근본적으로 해결하는 대책은 되지 못한다. 또 다른 방법으로 제시된 것이 웹 데이터를 사용자 가까이 두게 함으로써 웹 사용자에게 빠른 응답시간을 제공하는 방법이다. 이것은 멀리 있는 데이터를 자신의 근처에 가져다가 서비스를 제공하기 때문에 속도나, 네트워크 교통량(Network Traffic)의 양에 있어서 좋은 효과를 내게 하고 있다. 이것이 웹 캐싱 기술[2]이다.

본 연구는 웹 캐싱 기술을 기반으로 기존에 제작되어 있는 스퀴드(Squid)[3]라는 소프트웨어를 이용한다. 인트라넷(Intranet) 상에서 학내방송, 회의장면 등등의 멀티미디어 스트리밍 콘텐츠를 끊임 없이 볼 수 있는 최적의 환경 구축을 위한 솔루션(Solution) 탐색을 위하여 제반 관련 요소들의 성능에 대한 영향(Effect)을 분석, 평가한다.

2. 본 론

2. 1. 실험 환경

실측에 사용되는 시스템은 [그림 1]에서 보듯이 하

나의 리얼 웹 서버가 존재하고, 그 밑에 하나의 캐시 서버가 있어서 서비스를 제공하는 구조이다. 이런 배치 방식을 전방배치캐싱(Forward Proxy Caching)이라고 한다. 클라이언트는 하나의 컴퓨터에 하나의 요청을 가지는 것을 기본으로 생각하며, 동시에 여러 개의 요청이 요구되는 경우에는 여러 대의 컴퓨터를 사용하여 하나의 클라이언트 당 하나의 요청이 요구되는 방향으로 진행한다. 웹 서버에 멀티미디어 스트림 데이터를 올려놓고, 이것을 클라이언트에서 요청하는 형식을 취했다. 클라이언트가 웹을 통해서 멀티미디어 스트림 데이터를 요청하게 되면, 이 요청을 캐시서버가 훔쳐서 자신의 내부 디스크에 해당하는 데이터가 존재하는지를 판단하여, 이것을 서비스한다. 이러한 구조를 가지고 실측에 임하게 된다.

실험에 사용된 Client computer는 38대이며, Cache server와 Original server는 각각 1대씩이다. 실험에 사용된 컴퓨터 사양들은 [표 1]과 같다.

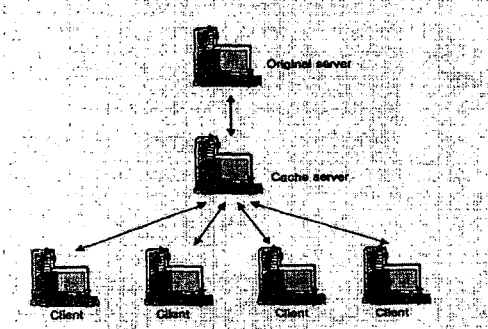


그림 1 실측 시스템

2. 2. 테스트 방식

실측의 정확도를 높이기 위해서 우선 캐시서버 내에 캐싱되어 있는 모든 데이터를 제거하고, 실측에 사용될 실제 데이터를 가지고 있는 웹서버에 클라이언트가 멀티미디어 스트림 데이터를 요청하게 한다. 이때 클라이언트는 컴퓨터 하나 당 하나의 요청이 요구되게 한다.

실측에 사용되는 멀티미디어 스트리밍 데이터는 다음과 같다.

용도	CPU	RAM	Network
Original server	400Mhz	196M	100M
Cache server	866Mhz	128M	100M
Client	1.7Ghz	512M	1G/100M

표 1 실측에 사용된 시스템의 사양

파일이름	파일 크기	대역폭	플레이 시간
test3.asf	4.543 KB	103 KBPS	6분 1초
test703k.asf	30 MB	703 KBPS	6분

표 2 측정에 사용된 동영상 정보

2. 3. 실측 결과와 분석

2. 3. 1. test3.asf 파일을 이용한 실측과 결과

아래는 동시에 38대의 클라이언트가 web server에 test3.asf(멀티미디어 스트리밍 데이터)를 요구하게 하여, 측정된 데이터 값들이다.

CPU utilization

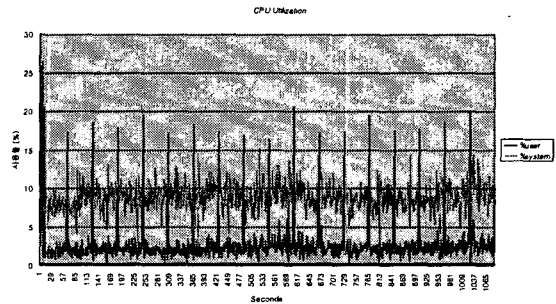


그림 2 test3.asf파일을 이용했을 때의 CPU utilization

Memory utilization

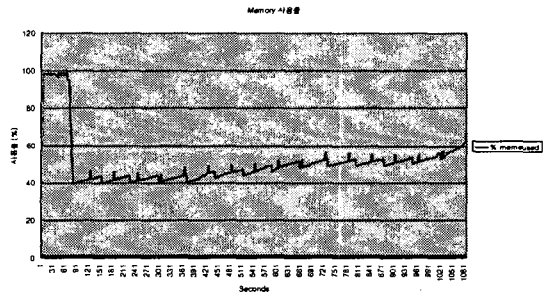


그림 3 test3.asf를 이용했을 때의 memory utilization

Network Utilization

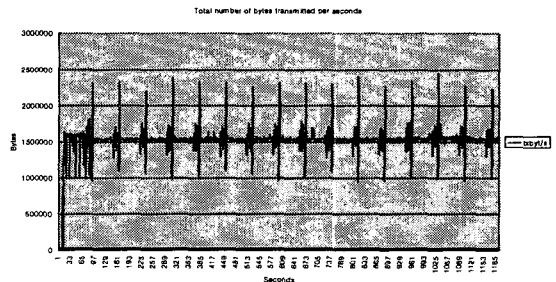


그림 4 test3.asf파일을 이용했을 때의 network utilization

[그림 2]을 보면 실제 캐시서버를 동작시키고 있는 상황에서 CPU의 이용률이 잘 나타나있다. 먼저 그림에서 확연히 나타나는 특징을 살펴보면 CPU의 이용률이 일정하다가 아주 높은 값으로 피크가 일어나는

부분이 있다. 이 부분이 동영상이 끝나고 또다시 멀티미디어 스트림 데이터를 요구하는 부분이다. 실측에 사용된 멀티미디어 스트림 데이터의 재생 시간이 6분이기 때문에 아래의 피크는 6분에 한번 꼴로 발생하고 있다. 이점을 미투어 요청을 요구하는 동안에는 CPU를 많이 사용하게 됨을 알 수 있다. 사용자 모드에서 사용하는 CPU의 이용률은 6분마다 피크가 나타나는 주기성이 분명하게 나타나는 반면에 시스템 모드에서 사용하는 CPU 이용률은 10%대에서 상하 5%를 반복하면서 나타나고 있다. 멀티미디어 데이터를 전송하는데 시스템 모드에서 CPU작업을 많이 요구하는 것을 알 수 있다.

메모리 이용률은 전체 리눅스 시스템에서 캐시서버가 동작하고 있을 때, 시스템에서 사용하는 메모리의 양이 얼마인지를 의미한다. [그림 3]은 캐시서버가 동작하는 동안 시스템에서 사용한 메모리 양의 변화를 그래프로 나타낸 것이다. 처음에 메모리의 사용률이 100%에 육박하는 것은 멀티미디어 데이터를 web server에서 가져오기 때문에 메모리를 많이 쓰게 된다.

[그림 3]을 보면, 일정한 주기마다 피크가 발생한 것을 알 수 있다. 이는 실측에 사용된 멀티미디어 스트림 데이터의 재생 길이가 6분임을 감안한다면, 반복 재생으로 반복적으로 데이터를 요청하고 있는 상황에서 한번의 재생이 끝나고 다시 데이터를 요청할 때, 피크가 발생하고 있다. 이 같은 현상이 발생하는 이유는 한번의 재생이 끝나고 새롭게 클라이언트에게서 요청을 요구받기 때문에 캐시서버 자신의 내부 디스크에 내용이 있는지를 검색해야 한다. 이때, 시스템에서 사용하는 메모리의 양이 늘어나는 것이다.

네트워크 대역폭 이용률은 캐시서버를 통해서 이루어지는 네트워크의 전송 양에 관계한다. [그림 4]은 캐시서버를 구동하고 있을 때, 캐시서버로부터 클라이언트에게로 전송한 초당 Byte수를 그래프로 나타낸 것이다.

[그림 4]을 보면, 일반적으로 1.5Mbps의 전송속도로 데이터를 전송하고, 일정한 주기를 가지고 약 2.3Mbps의 피크치를 볼수 있다. 이것은 실측에 사용된 대상 파일의 재생 길이가 6분이기 때문에 한번의 재생이 끝나고 또 다시 캐시 서버에 요청을 요구하는 부분이다. 또, 6분전에 바닥으로 떨어지는 이유는 클라이언트로부터 요청을 받고 난 다음, 이것을 처리하는 시간이 필요하다. 그렇기 때문에 실제로 네트워크를 통해서 전송되는 데이터는 거의 없다. 이때, 이용률이 바닥으로 떨어지는 것이다.

Cache Server의 네트워크는 100Mbps이므로 네트워크에 부하를 더 많이 주려면 더 큰 대역폭을 요구하는 멀티미디어 스트리밍 데이터가 필요하다는 것을 알 수 있다. 이와 같은 인식을 가지고 실험을 한 것이 다음의 실험이다.

2. 3. 2. test703k.asf 파일을 이용한 실측, 결과

Cache server의 네트워크에 더 많은 부하를 줄 목적으로 대역폭이 test3.asf보다 7배 정도 더 큰 멀티미디어 스트리밍 데이터를 가지고 실험을 하였다. 이번

의 실험은 위의 실험과는 약간 다른 방식으로 테스트를 하였다. 초기에는 8대의 클라이언트가 web server에 멀티미디어 스트리밍 데이터를 요청한다. 그리고 10분마다 추가로 6대의 클라이언트가 web server에 멀티미디어 스트리밍 데이터를 요청하게 된다. 이렇게 되면 50분 뒤에는 38대의 클라이언트가 web server에 데이터를 요청하게 된다. 캐시서버에 대한 실측은 2시간을 하였다. 결과는 아래와 같다.

CPU utilization

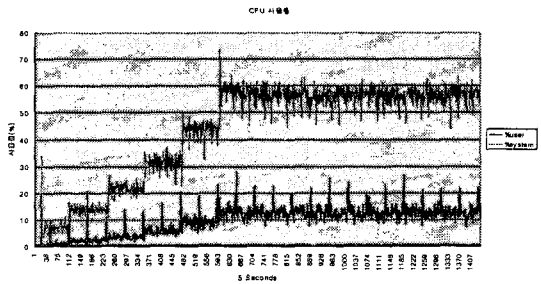


그림 5 test703k를 이용했을 때의 CPU Utilization

Memory utilization

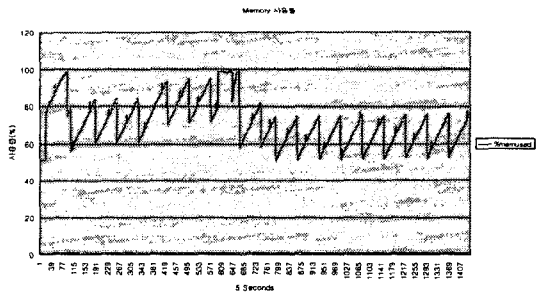


그림 6 test703k를 이용한 memory utilization

Network Utilization

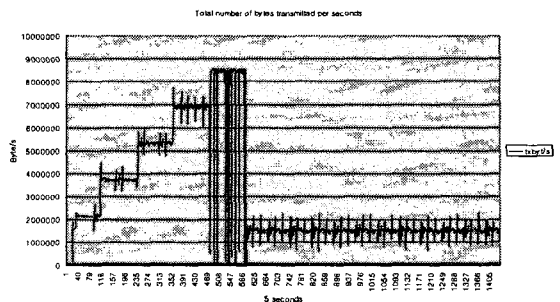


그림 7 test703k를 이용한 Network Utilization

[그림 5]는 6분 주기로 시스템 모드에서의 CPU 이용률이 10%씩 증가하는 것을 알 수 있다. 또한 사용자 모드에서 CPU 이용률도 증가하는 것을 알 수 있

으나 시스템 모드에서의 증가보다는 매우 적은 2%정도 증가하는 것을 볼 수 있다. 이를 보면 멀티미디어 스트리밍 데이터를 보내는데 시스템 모드의 작업이 영향이 크게 미치는 것을 알 수 있다.

[그림 6]은 Memory 이용률을 나타낸 것이다. 이는 그림 3과 같이 멀티미디어 재생시간인 6분의 주기로 피크치가 나오는 것을 볼 수 있다. 동일한 멀티미디어 스트리밍 데이터를 요청할 때는 클라이언트 수가 많고 적음에 상관없이 동일하게 메모리를 이용한다고 할 수 있다.

[그림 7]은 네트워크 이용률을 나타낸 것이다. 50분이후로 네트워크 이용률이 매우 불안정하게 나타나다가 60분이 넘어가면 네트워크 이용률이 매우 저조하게 나타나는 것을 볼 수 있다. 이는 [그림 5]의 CPU이용률을 보고 설명할 수 있다. [그림 5]에 나타난 CPU이용률에서 50분대이후에는 시스템 모드에서의 이용률은 60%대이고 사용자 모드에서의 이용률은 15%대에 가깝다. CPU에 많은 부하가 걸리므로서 Cache서버의 작업을 다 처리하지 못하여 Cache서버가 각각의 client들에게 멀티미디어 스트리밍 데이터를 전송해 주지 못한다고 할 수 있다.

3. 결 론

현재 네트워크에는 멀티미디어 스트림 데이터를 인터넷상에서 실시간으로 전송하는데 있어서 장애요소가 있다. 본 논문에서는 이러한 장애요소를 극복하기 위하여 쓰이는 웹 캐싱 솔루션에 사용되는 기술들을 찾아내는데 그 목적을 두고 있다. 이러한 목적을 달성하기 위해서 여러 가지 요소들에 대하여 실측을 통하여 성능을 평가해 보았다.

메모리나 CPU의 이용률의 추이를 살펴보면, 처음에 요청을 요구하는 부분에서 많은 자원을 사용하고 있음을 확연하게 나타나 있다. 그리고 요청의 수에 따른 변화를 보면, 요청의 증가에 따라 CPU이용률과 네트워크 이용률이 비례적으로 증가하는 반면에, 메모리의 이용률은 증가가 없었다.

네트워크에 부하를 주게 하려고 의도한 실험이 CPU에 많은 부하를 주게 되어, 제대로 된 결과를 측정할 수 없었다. 네트워크에 부하를 줄 수 있는 멀티미디어 스트리밍 데이터를 가지고 다시 실험을 하여 원하는 결과를 얻도록 할 예정이다.

참고문헌

- [1] 한국전산원 <http://www.nca.or.kr/>
- [2] Barish, G.; Obraczke, K. "World Wide Web caching: trends and techniques", IEEE Communications Magazine, Volume: 38 Issue: 5, May 2000, Page(s): 1045 -1052
- [3] D. Wessels, Squid Internet Object Cache; Homepage of the Squid project, <http://www.Squid-cache.org>, 1996



이 영 석

2000 : 서울시립대학교 전자전기
컴퓨터 공학부 학사
2002 : 서울시립대학교 전자전기
컴퓨터 공학부 석사

관심분야 : Metacomputing, System software, parallel processing



윤 창 일

2001 : 서울시립대학교 전자전기
컴퓨터 공학부 학사

관심분야 : Metacomputing, GRID, 시스템 소프트웨어, Embedded System



이 용 우

1981 : 서울대학교 공과대학 전기
공학과 학사

영국 Edinburgh 대학 Computer
Science Department 석사(월반),박사.

1981 : Schlumberger Technical Services
Inc. International Engineer.

1982-1998 : KIST/KAIST, ETRI, 선임연구원.

1998 : 교육부 KMEC, 책임연구원

1999-현재 : 서울시립대학교 전자전기컴퓨터공학부

관심분야 : Metacomputing, GRID, CSCW, 시스템 소프트웨어,

초고속통신, 가상학교, 성능평가공학 등등.