NOD 시스템에서 오디오와 비디오 데이터 Push에 따른 Precaching 연구

[0]박성호, 김광문, 송기욱, 정기동
부산대학교 전자계산학과 병렬멀티미디어연구실

A Study on Precaching according to Push Technology for Audio and Video Data in NOD(News On Demand) System

Sung-Ho Park, Gwang-Moon Kim, Ki-Uk Song, Ki-Dong Chung
Parallel Multimedia Lab., Dept of computer science, Pusan National University

요 약

컴퓨터와 통신망 기술의 발전으로 많은 언론 기관에서 전자신문 서비스를 제공하고 있다. 그러나 현재 서비스되는 전자신문은 텍스트 위주의 정적인 정보를 주로 서비스 하며, 사용자가 필요한 정보를 찾아 다니는 pull 기술을 기반으로 서비스한다. 그리고 사용자 맞춤 기능을 제공하지 못하므로, 불필요한 정보를 사용자에게 전송함으로써 시스템 자원을 낭비하는 단점을 가지고 있다.

본 논문에서는 멀티미디어 데이터를 지원하는 NOD(News On Demand)시스템에서 텍스트 데이터 뿐아니라 오디오/비디오 데이터를 push하므로써 서버의 load를 분배시키고 사용자에게 실시간성을 제공하는 NOD분배 서버를 설계하고 프로토타입을 구현하였다. 특히 본 논문에서는 대용량 데이터인 오디오/비디오 데이터를 사용자 시스템으로 Push할 때, 사용자 시스템의 디스크 공간상태 등을 고려하여 실시간성을 유지할 수 있는 적정 Push량을 시뮬레이션을 통해 측정하였다.

1.서론

인터넷과 초고속정보통신망 그리고 디지털 기술의 급격한 발전으로 기존의 종이와 잉크로 표현되던 신문의 정보 표현 방식이 일대 전환점을 맞이하게 되었다. 기존의 신문은 2차원적 평면상에 텍스트 및 이미지등 정적인 정보를 표현하므로 동영상이나 오디오와 같은 멀티미디어 정보를 수용하는데에는 한계가 있다. 이러한 한계를 극복하고 독자들이 원하는 정보를 데이터통신망을 통하여 interactive하게 실시간으로 제공하기 위한 전자신문이 출현하게 되었다. 전자신문은 인터넷과 WWW(World Wide Web)의 등장, 고성능 PC의 출현에 기반하여 국내외적으로 서비스를 제공하고, 많은 독자를 확보하고 있다. 그러나 기존의 전자신문은 텍스트위주의 정적인 정보 제공, Client Pull기술의 채택, 주문(맞춤)형 정보 제공 기능의 결여 등의 단점을 지니고 있다. 이로 인하여 독자들은 필요한 정보를 보기 위하여 서버에서 제공하는 불필요한 headline 정보를 모두 본 후에 필요한 정보를 선택하여야 하며 또한 그 기사가 다운로드(download)될때까지 기다려야하는 등 독자들에게 많은 작업을 요구된다.

이러한 단점을 극복하기 위하여 현재 몇몇의 전자신문 사이트에서는 사용자 설정에 따른 맞춤 정보를 사용자의 로컬 시스템으로 Push하는 기법을 이용하여 전자신문 서버 시스템의 로드 분배하고 맞춤 정보에 대해서는 사용자의 정보 요구시 로컬

시스템에서 데이터를 로드함으로써 기다리는 시간을 최소화한다. 그러나 현재 제공되고 있는 데이터는 텍스트 위주의 소량 데이터이므로 사용자 로컬 시스템의 디스크 공간, 네트워크 대역폭 등 로컬 시스템 자원의 Push기법을 지원하기위한 할당이 로컬 시스템의 성능에 영향을 많이 미치지 않으나, 대용량 데이터인 비디오나 오디오 데이터는 로컬 시스템의 자원을 많이 소모하므로, 로컬 시스템의 성능을 저하시킬 수 있다.

본 논문에서는 현재 전자신문이 제공하는 텍스트 및 이미지 위주의 서비스뿐만 아니라 비디오, 오디오 등 멀티미디어 데이터를 서비스하는 NOD 시스템 구축을 위한 기반 기술 연구로써, 사용자 맞춤 기능과 Push기능을 제공하는 NOD 분배서버 구조를 설계하고 프로토타입을 구현하였다. 특히, 대용량 데이터인 비디오와 오디오 데이터를 사용자 시스템으로 Push할 때, 데이터의 실시간성을 유지함과 동시에 NOD 서버 시스템의 로드 밸런스를 유지하고 사용자 시스템의 디스크공간 등을 고려한 대용량 데이터의 적정 Push 량을 시뮬레이션을 통해 측정하였다.

본 논문은 2장에서 관련 연구, 3장에서 시스템 개요에 대해 설명하였다. 그리고 4장에서는 시뮬레이션 및 결과를 설명하고 5장에서 결론을 맺는다.

2. 관련 연구

국내 각 신문사에서 현재 WWW(World Wide Web)을 통해 서비스중인 전자신문은 위에서 제시한 맞춤형 정보제공 기능과 멀티미디어 데이터 서버스 기능, 데이터 Push기능을 제공하지 못하고 단지 사용자가 interactive하게 선택한 정적인 정보만을 제공한다[1]. 그리고 맞춤형 정보제공에 대한 개발이 동아일보의 IIC (Internet Information Center)등에서 이루어지고 있으나, 주로 문자 정보 위주로 서비스가 제공되고 있으며, 동영상이나 음성등 멀티미디어 서비스가 제공되고 있지 못하여 사용자들의 요구를 충분히 반영하지 못하고 있다.

KBS NOD와 MBC의 뉴스데스크 등에서는 Streamwork나 Real Video/ Audio를 이용한 서비스가 이루어지고 있으나, 영상의 질이 현저히 떨어지며, 네트워크 트래픽이나 서버의 overload시 유동적으로 대응하지 못하고 맞춤정보 서비스 및 데이터 Push 기능을 제공하지 못하고 있다.

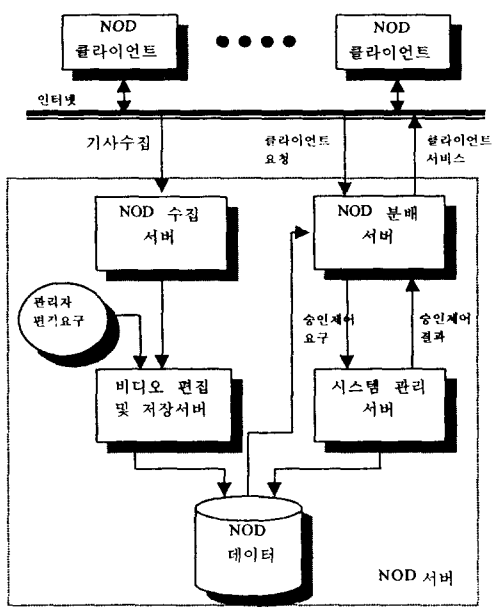
미국의 카네기멜런대학과 캐나다의 앨버트대학등에서는 VOD연구의 일환으로 NOD를 연구가 이루어지고 있다. 미국 카네기멜런대학에서는 Informedia라는 NOD에 대한 연구를 하고 있으나 전자신문이라기보다는 주로 디지털라이브러리에 대한 연구이다[2]. 그리고 캐나다의 앨버트대학에서는 객체지향데이터베이스 개발의 일환으로 연구가 이루어지고 있다.

포인트 캐스트사의 Pointcast는 사용자의 요구에 따라 데이터를 Push해주는 기술를 적용하고 있다. 그러나 현재, 텍스트 데이터 위주의 서비스를 지원하고 있으며, 대용량 멀티미디어 데이터의 Push서비스에 대한 연구는 미비한 상태이다.

3. 시스템 개요

3.1 NOD 시스템 구조

NOD 시스템은 사용자가 멀티미디어 뉴스 데이터를 원하는 시간에 원하는 기사를 볼 수는 기능을 제공한다. NOD 시스템은 Internet을 통하여 다양한 데이터를 수집, 분류, 저장한 다음 사용자가 원하는 뉴스 기사를 검색, 시청 또는 청취할 수 있는 서비스를 제공한다. 이런 NOD 서비스가 원활하게 제공되기 위해서 NOD 시스템은 사용자 정보 관리 및 뉴스 데이터 수집, 편집, 저장, 분배 등과 같은 기능들을 수행해야 한다. NOD 시스템은 [그림 1]과 같이 NOD 클라이언트와 NOD 수집서버, NOD 분배서버 , 비디오/오디오 편집 및 저장 서버, 시스템 관리 서버로 구성된다.

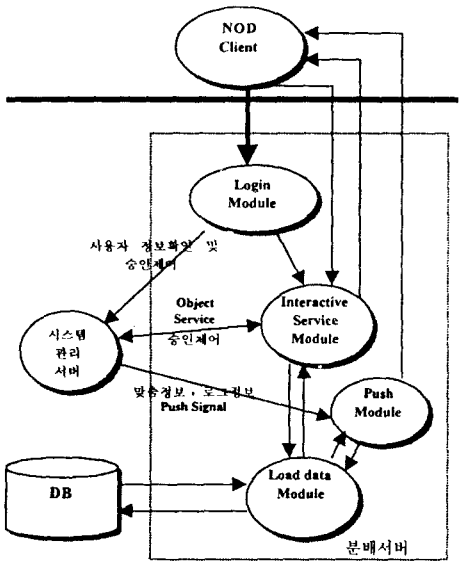


[그림 1] NOD 시스템의 구조

NOD 클라이언트는 사용자 로컬 시스템에서 구동되는 응용프로그램으로 NOD 서버로부터 전송된 멀미티미디어 데이터를 디스플레이하는 기능을 제공하며, 서버로부터 Push된 데이터를 관리하는 역할을 한다. NOD서버의 수집 서버는 Internet상을 자동을 돌아다니면서 매일 새롭게 발생하는 뉴스 기사를 수집하는 모듈이다[3]. 비디오/오디오 편집 및 저장 서버는 수집 서버가 수집한 데이터를 분류하고 저장하며, 저장되어 있는 비디오/오디오 데이터의 편집 기능을 제공하는 모듈이다. 시스템 관리 서버는 사용자의 QoS(Quality of Service)를 고려한 승인제어 및 사용자 정보관리 등의 기능을 수행한다. 마지막으로 본 논문에서 중점적으로 논의할 분배서버는 사용자에게 interactive한 서비스 제공뿐아니라 비디오 및 오디오 데이터 위한 실시간 서비스 제공, 맞춤정보 서비스를 위한 push기능을 수행한다.

3.2 분배 서버 구조

분배서버는 데이터베이스로부터 사용자가 요구한 기사를 검색하여 사용자에게 전송하는 통합 모듈로써 [그림 2]와 같은 구조를 가진다.



[그림 2] 분배 서버 구조

Login모듈은 사용자를 확인하고 사용자 단위의 승인제어를 수행한다. 사용자 단위의 승인제어는 시스템 자원에 대한 확률적 승인제어를 지원함으로써 보다 많은 사용자를 수용할 수 있다[4]. Interactive 서비스 모듈은 각 NOD 클라이언트당 하나의 쓰레드가 생성되며, 클라이언트의 interactive한 요구에 따라 텍스트와 이미지 데이터뿐만아니라 오디오와 비디오 데이터를 전송한다. 여기서 분배서버는 오브젝트단위의 승인제어기능을 통해 현재 서비스되고 있는 사용자의 QoS를 보장함과 동시에 새로운 요구를 서비스할 수 있는 정책을 고려한다. 그리고 사용자의 맞춤정보 및 사용자 로그정보를 시스템 관리 서버에게 제공한다. 데이터 로드 모듈은 사용자가 요구한 데이터를 데이터베이스로부터 검색, 캐쉬하는 기능을 수행한다. 마지막으로 본 논문에서 중점적으로 논의할 Push모듈은 맞춤정보와 로그정보를 이용하여 사용자의 선택이 없이도 사용자가 원하는 기사를 사용자 로컬 시스템에 전송하는 기능을 제공한다.

3.3 Push 모듈

Push모듈은 시스템관리 서버의 맞춤정보 및 로그정보를 이용하여 NOD서버의 로드를 고려하여 사용자 로컬 시스템으로 예상 데이터를 미리 전송하는 모듈이다.

▶ 장점

- 사용자의 선택이 없이 데이터가 미리 전송되므로 사용자 대기시간을 줄인다.
- 비디오나 오디오데이타는 실시간 서비스가 필요하므로 대용량 대역폭이 필요하나 Push 기술은 사용자가 동시에 폭주하는 시간의 시스템 사용 대역폭을 분산시키므로써 NOD서버 시스템의 성능을 향상 시킬 수 있다.
- Multicast 방식의 네트워크 서비스가 가능하므로 NOD서버의 디스크 대역폭이나 캐쉬 관리에 있어 효율성을 제공한다.

▶ 단점

- 사용자 시스템에서 Push되는 데이터를 관리할 특별한 응용프로그램이 필요하다.
- Push가 발생할 경우 로컬 시스템의 CPU time, 네트워크 대역폭, 디스크 공간 등의 자원이 소요된다. 특히 비디오나 오디오와 같은 데이터를 Push 받아 관리하기 위해서는 많은 디스크 공간이 필요하다.

위의 Push기술의 장단점을 고려하여 사용자 로컬 시스템의 자원을 적정 수준으로 요구하는 Push 데이터량을 측정할 필요가 있다. 예를 들면, 뉴스 데이터 비디오 클립 한편은 대개 2 - 3분 정도로 MPEG-I으로 압축하였을 경우, 24 - 36Mbyte의 저장 공간이 필

요하다. 그러므로 하루 10편 이상의 뉴스 비디오 클립 데이터를 Push받기 위해서는 수백 Mbyte의 사용자 로컬 시스템의 디스크 공간이 요구된다. 이는 사용자 로컬 시스템 자원의 낭비를 초래하여 로컬 시스템의 성능을 저하시킨다. 그러므로 적정수준의 데이터 Push량을 선택하여 NOD서버의 성능을 향상시키는 반면, 사용자 로컬 시스템의 로드를 최소화하여야 한다. 본 논문의 4장에서는 시뮬레이션을 통해 적정한 push데이터량을 측정하였다.

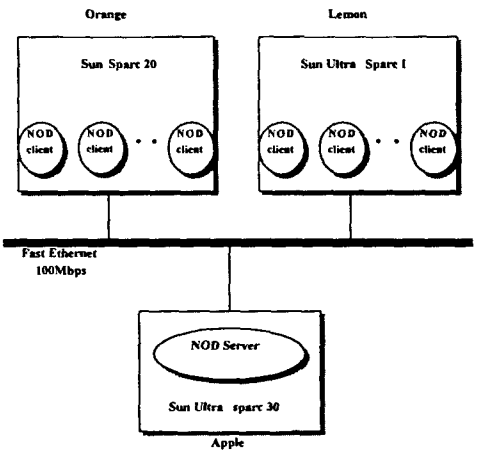
4. 시뮬레이션 및 결과

4.1 시뮬레이션 환경

NOD 클라이언트와 NOD 서버를 아래와 같은 환경에서 구축하였다.

- 시뮬레이션 시스템 :
  - 서버 : Sun ultra-30
  - 클라이언트 : Sun Sparc-20, Sun ultra-I
- 사용 언어 : Java 1.1.5
- 사용 O.S : Solaris 2.5.1
- 데이터 : MPEG-I, 14Mbyte(70초)

시뮬레이션은 [그림 3]과 같이 Apple에 NOD 서버 시뮬레이션 프로그램을 구동하고 Orange와 Lemon에 NOD 클라이언트 시뮬레이션 프로그램들을 구동하여 수십명의 사용자가 동시에 사용하는 환경과 같은 시뮬레이션 환경을 구축하였다.



[그림 3] 시뮬레이션 구조도

시뮬레이션을 위한 뉴스 비디오 클립의 수는 20개로 가정하였다. 그리고 사용자의 기사 접근 빈도는 특정 전자신문의 하루 접근 빈도[5]를 기준으로 작성한 [표 1]의 값을 사용하였다.

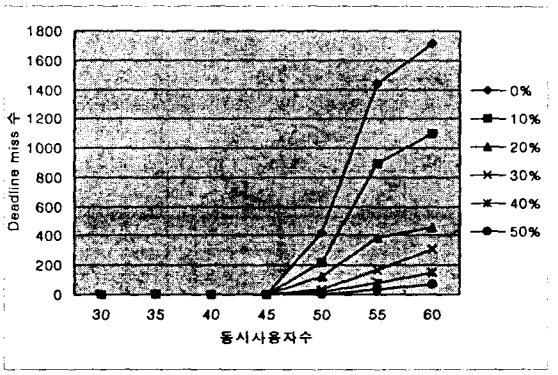
[표 1] 20개의 뉴스 비디오 클립의 접근 확률

스트림 No.	1	2	3	4	5	6	7	8	9	10
접근 확률	.110	.103	.096	.094	.091	.086	.083	.072	.063	.047
스트림 No.	11	12	13	14	15	16	17	18	19	20
접근 확률	.031	.024	.019	.015	.012	.012	.012	.011	.010	.009

4.2 시뮬레이션 분석

가. 데이터 Push량에 따른 분배서버의 Deadline miss 측정

Push된 데이터량이 NOD 시스템에서 어떠한 영향을 미치는가를 측정하기 위하여 Push된 데이터량이 0%, 10%, 20%, 30%, 40%, 50%에 대하여 동시 사용자수를 증가시키면서 분배서버의 deadline miss를 측정하였다[그림 4].



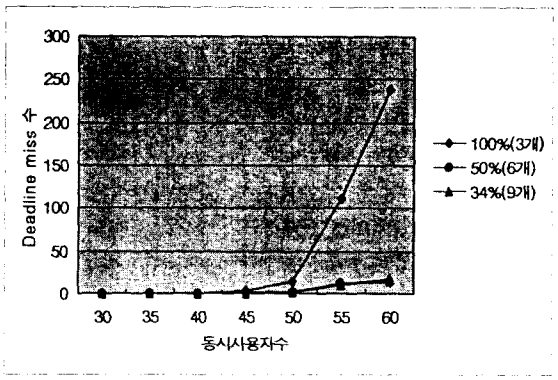
[그림 4] Push량에 따른 deadline miss

위의 [그림 4]에서 나타난 결과를 살펴보면, Push된 데이터의 량이 0%, 10%에서는 동시사용자 수가 45명을 초가하면서 급격한 deadline miss 횟수의 증가를 보이는 반면, 데이터 Push량이 20%이상인 경우 비교적 완만한 deadline miss 횟수의 증가를 보이고 있다. 그러므로 적절한 데이터 Push량을 선택하므로써 사용자 시스템의 로드를 적정수준으로 유지하고 서버 시스템의 성능을 향상시킬 수 있다.

나. Push된 데이터 수에 따른 분배서버의 Deadline miss 측정

사용자 시스템의 하드 디스크 공간이 고정되어 있을 때, Push된 데이터의 개수가 NOD 시스템에서 어

떠한 영향을 미치는가를 측정하기 위하여 접근 빈도가 높은 순서대로 3개, 6개, 9개의 데이터를 각 100%, 50%, 34% 순으로 Push 하였을 때, 사용자수를 증가시키면서 분배서버의 deadline miss를 측정하였다[그림5].



[그림 5] Push수에 따른 deadline miss

위의 [그림 5]에서 나타난 결과를 살펴보면, 사용자 시스템의 동일한 디스크 공간을 사용할 때, 데이터 push rate를 줄이더라도 많은 개수의 데이터를 push 하는 것이 더 좋은 성능을 보임을 알 수 있다.

5. 결론 및 향후 연구 과제

본 논문에서는 맞춤 정보제공 및 Push기술을 지원하는 NOD시스템의 구조 및 Push 기술의 장단점에 대하여 살펴보았다. 그리고 NOD서버의 성능을 향상시키고, 사용자 로컬 시스템의 성능 저하를 최소화하기 위한 데이터 push량을 시뮬레이션을 통해 측정하였다.

향후 연구 과제로는 Multicast 전송방식을 이용한 Push 기법 및 캐쉬 관리기법에 관한 연구 및 맞춤 정보와 로그 정보를 이용한 정확한 Push 데이터 선택, NOD 서버 시스템의 로드 균형을 고려한 Push시점 선택에 관한 연구가 필요하다.

참고 문헌

[1] 부산일보사, "CTS, 기사/화상/공공 DB에 기반한 NOD시스템 개발 진도보고서", 1997

[2] Christel, M., and Pendyala, K., Informedia Goes to School: Early Findings from the Digital Video Library Project, D-Lib Magazine, September 1996.

[3] Martijn Koster, NEXOR, "Robots in the Web: threat or treat?", ConneXions, Vol. 9, No. 4, April 1995. http://info.webcrawler.com/mak/projects/robots/threat-or-treat.html

[4] H. M. Vin, P. Goyal, "A Statistical Admission Control Algorithm for Multimedia Servers", in Proc. of ACM Multimedia 94, pp.33-40, 1994.

[5] 이주경, 박용운, 김영주, 정기동, "NOD(News-On-Demand) 데이터의 저장 기법", 한국 정보과학회, 제24권 2호, 1997