System Area Network 에 기반한 Scalable VOD Server 의 설계 및 구현

0
*조재호, *심재홍, *강봉직, *최경희 ,*김재훈,**정기현
*아주대학교 정보 및 컴퓨터 공학부 . **아주대학교 전자 전기 공학부

Design and Implementation of Scalable VOD Server based on System Area Network

0
*Jae-Ho Jo, *Jae-Hong Shim, *Bong-Jik Kang*Kyung-Hee Choi, **Gi-Hyun Jung
* Information & Computer Division, Ajou University.
**Division of Electronical & Electronics Engineering, Ajou University.

요 약

컴퓨터 하드웨어 기술의 발달은 물론 고속 통신 기술의 발달로 컴퓨터의 사용분야가 단순 계산에서 멀티미디어를 사용하는 VOD 서버 기술을 필요로 하게 되고 경제적이며 효율적으로 멀티미디어 데이터를 다루는 기술이 요구 되면서 서버의 확장성과 원활한 데이터 서비스가 필요하게 되었다. 이에 본 논문에서는 초고속 통신망에 기반한 SAN(System Area Network)에 기반한 클러스터링 시스템을 제안하여 최대한의 대역폭을 이용한 병렬 I/O 를 통하여 SAN 상에서 입출력 성능의 개선할 수 있는 Scalable VOD 서버를 제안하여 구현 성능 평가 하려 한다.

1 서 론

컴퓨터 하드웨어 기술의 발들은 물론 고속 통신 기술의 발달로 컴퓨터 사용 분야 및 사용환경이 단순 계산 분야로부터 점차 멀티미디어를 사용하는 환경으로 변하고 있다. 특히 최근 5 년 사이에 발달한 VOD 서버 처리 기술은 컴퓨터 기술 분야에 많은 영향을 주었을 뿐만 아니라 사회 여러 곳에 많은 영향을 주고 있다. VOD 서버의 가장 핵심이 되는 기술은 무엇보다도 온-디멘드(on-demand)로 디지털 멀티미디어 데이터를 클라이언트에게 보내도록 서버와 통신 망을 구성하는가 이며, 서버의 구현은 VOD 서버를 효율적으로 저장하고 검색하는 기술을 필요로 하고 있으며 이러한 관점에서 MOD(Multimedia On-Demand) 서버를 경제적으로 구현하려는 노력이 있으며 고속 통신망으로 서버를 연결하여 병렬 VOD 서버를 구현 하려 하고 있다. 이에 같은 가격으로 사용량 증가에 따른 시스템 증설에서 용이하게 하는 확장성(scalable)있고 많은 사용자들을 서비스 할 수 있는 Scalable VOD 서버가 요구 되게 되었다. 이러한 요구에 의해 많은 사용자 서비스를 위한 목적으로 개발된 N-VoD[1] 시스템이 있으나 서비스를 위하여 특정한 사용자는 특정한 상태까지 대기 해야 하는등 제약 사항이 있었다 이에 원할한 서비스를 위하여 AT&T Bell LAB 에서 버퍼 구조에 중심을 둔 *Fellini* 멀티미디어 저장 장치와[2],확장성에 중심을 둔 Berkeley 에서의 VOD 시스템을 제시하였다[3] 그러나 이 시스템은 Archive Server 의 개념을 인터넷 상에 두고 서버에서 연속된 데이터를 전송 해주기 때문에 네트웍 상에서 병렬 I/O 가 제공되지 않고 단순한 데이터 서버로의 분산된 환경을 제시 하였다. 이에 본 논문에서는 대용량의 데이터를 다루는 VOD 서버에서 네트웍을 단순한 접근 서버로 이용하는 것이 아니라 고성능 입출력을 위한 구조로 SAN 상에서 클러스터된 병렬 검색을 위한 서버를 그림 1 과 같이 제시하였다. 본 구조는 디스크 저장 매체를 SAN(System Area Network)상에 분리하여 디스크 저장 매체의 증설과 서버의 증설을 용이하게 하여 확장성(scalable)이 있으며, 각 펌프에 병렬 I/O 를 통하여 시스템의 성능을 극대화 시키며, 시스템의 부하에 따라 영화를 서비스 하는 비디오 서버를 자유롭게 선택하게 하는 Scalable VOD 서버의 구조를 제시 하여 구현하여 SAN 상에서의 접근 패턴에 따른 서버의 성능테스트를 하려 한다.

2 장에서 시스템의 모델을 제시하고 그에 따른 자세한 구현 방안에 대하여 논하며 3 장에서는 성능측정 방법을 4 장에서는 향후 계획에 대하여 살펴 보기로 한다.

2 시스템 모델 및 구현

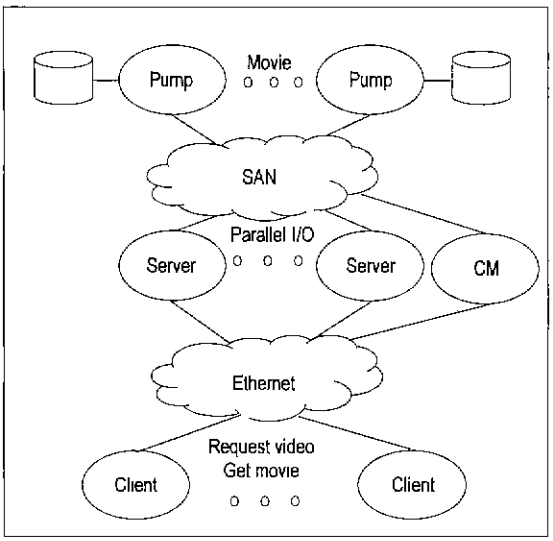

그림 1. 시스템 구조

그림 1 과 같이 본 VOD 시스템은 사용자에게 영화를 보여 주기 위한 서비스 서버(Service Server)와 초고속 망위에서 분산된 네트웍 환경인 SAN(System Area Network)를 통하여 서비스 서버에 영화를 공급하는 펌프(Pump)와 시스템의 성능을 위하여 제공되는 버퍼링이 기본적인 구성 요소이다. 이러한 서비스는 CM(Central Manager)이라는 사용자 접속 매니저를 통해서 이용 가능 하다. 다음에 기술할 내용은 이러한 시스템의 상세 구조다.

2.1 CM (Central manager)

CM 은 서버내의 부하요소, 사용자의 수 등에 의하여 사용자에 대한 서비스 가능 여부를 결정 하며, 사용자에게 서비스 가능한 영화 목록을 보여주는 디렉토리 서비스를 수행한다. CM 의 가장 중요한 기능으로 선택된 영화의 서비스를 위하여 서버내의 부하요소에 기반하여 원할 하게 서비스 할 수 있는 최선의 서버를 선택하여 실제 서비스 받는 사용자와의 재 연결을 확립한다. 사용자와의 재확립시에 사용자가 요청한 파일의 구성 정보를 해당 서버에게 전송해 주어 서비스 서버가 서비스 요구를 할 수 있게 한다

2.2 서비스 서버(Service Server)

서비스 서버(Service Server)는 CM 에 의해 선택되며, 사용자에게 영화 스트림(stream)을 공급하는 일을 수행한다. 서버는 영화 서비스를 위하여 선택된 영화, QoS(Quality of Service) 파라메터 등을 CM 으로부터 전송 받아 서비스를 위한 버퍼를 구축한다. 서버는 영화 서비스를 위하여 데이터를 요청하기 위하여 CM 으로 부터 전송되온 영화에 대한 정보를 보고 펌프에게 SAN 을 통하여 각 펌프에 스트라이핑(striping) 된 필요로 하고 있는 데이터를 요청한다 영화에 대한 요청에는 스트라이핑되어 있는 영화에 대한 인덱스를 CM 으로 부터 전송 받게 되며 실제 데이터를 소유하고 있는 펌프의 아이디를 가지고 있어 메시지를 펌프에 자유롭게 보낼 수 있다. 서비스 서버는 그림 2 와 같이 이루어 진다.

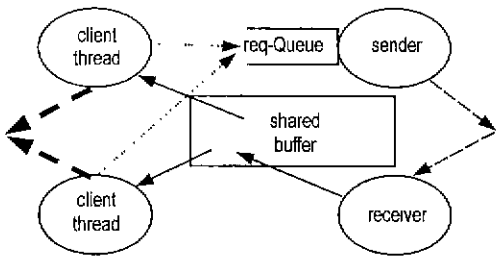

그림 2. 서버 구조

서비스 서버는 사용자의 요청이 있을 때 각각의 사용자 마다 **client thread** 를 만들며 이들은 각각 현재 주기동안에 요청해야 할 페이지 수, 접근하는 페이지 등의 요소를 가지고 **sender thread** 의 **req-Queue** 에 현재 접근 페이지의 서비스를 요청한다. **sender** 는 이 **req-Queue** 의 정보에 의하여 **shared buffer** 상에 버퍼에 접근하여 필요한 버퍼를 사전에 할당하며, 버퍼 안에 서비스 해야 할 페이지가 들어 와 있는지 결정을 내리게 되고 버퍼에 존재 하지 않는다면 페이지를 디스크 구조에 의하여 요청한다. 이렇게 요청된 정보는 각

펌프들에 의해서 읽혀 져 **receiver**로 병렬 I/O로 전송 되 요청한 데이터를 버퍼 알고리즘을 통하여 **client thread**가 사용 할 수 있게 배치 하여 준다. 이러한 서비스 서버의 버퍼는 캐쉬의 역할을 위하여 제공되어 SAN상의 시스템 부하를 줄여 시스템 전반적으로 부하를 감소 시켜 서비스를 원할 하게 해줄 수 있으며, 병렬 I/O 때문에 요구 될 수 있는 데이터의 순서 제어 역할을 한다.

2.3 펌프 (Pump)

펌프는 서버에서 영화를 요청했을 때 그림 3과 같이 영화를 SAN을 통하여 서비스 서버에게 전달해 준다. 펌프 역시 데이터 요구가 들어 왔을 때 마다 디스크에서부터 데이터를 읽어서 서버에게 전달 한다. 이때, 디스크의 상대적으로 느린 접근 시간(access time)때문에 서버에게 영화를 정해진 시간 안에 전달 하지 못 할 수도 있다. 따라서 디스크에서 읽어 들이는 속도와 네트웍으로 보내주는 속도차를 줄이기 위하여 펌프에서도 디스크와 프로세스간의 버퍼링이 필요하며, 이는 프리-페치(pre-fetch)방식을 취하여 서비스 한다

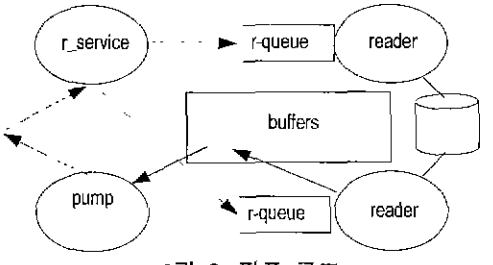

그림 3. 펌프 구조

r_service는 서버의 **sender**로 부터 videoid, pagcoffset, requestsize를 데이터를 읽는 **reader**로 전송받은 패킷을 큐잉 시킨다. 큐잉된 정보로부터 **reader**는 데이터를 읽어 **buffers**안에 데이터를 넣게 되고 읽혀진 데이터는 **pump**가 자신의 주기에 전송을 요청한 서버에게 전달 된다.

3 실험 방법

본 시스템은 SAN을 통하여 펌프와 서버를 PVM[4]을 통하여 연결하여 사용자의 접속에 대하여 확장성(scalablity)이 있음을 실험을 통하여 입증하려 한다. 본 VOD 시스템의 실험 환경을 다음과 같이 구성한다. Pentium Pro를 사용하는 서비스 서버 2대와 Pentium II를 사용하는 펌프 2대, SAN(System Area Network)으로 Myrinet[5]을 이용하여 VOD 시스템을 구성 하여 시스템의 성능 분석은 사용자의 접근 패턴에 따라서 펌프와 서버의 수를 점차적으로 늘려 감에 따라 인증할수 있는 노드의 수가 증가 함을 보이는 것을 통하여 본 시스템의 확장성(scalable)을 입증하며, 또한 CM의 서버 선택 알고리즘에 따라 시스템의 성능을 비교 하여 최선의 스케쥴링 전략을 찾고자 한다 비교 할 전략은 임의의 선택, 라운드-로빈(round-robin), 시스템 로드와 버퍼캐쉬의 정보에 따라 수용할 수 있는 최대 사용자 수를 실험 하려 한다.

4 향후 계획

본 구조에 입각하여 시스템을 구성 한 후, Scalable VOD 서버 시스템에 필요한 디스크 스트라이핑, 서비스 서버에서의 요청 스케쥴링 전략, 펌프와 서버에서의 최적의 버퍼링 알고리즘을 적용 개발 하려 한다.

5 Reference

[1] A.dan. D.M.Dias. R. Mukherjee, D Sitaram, and R.Tewari, " Buffering and caching in large-scale video servers," *in Proc. Compcon*, pp.217-224, March 1995

[2] Ozden. B., R. Rastogi, and A. Silberschatz, "Buffer Replacement Algorithms for Multimedia Databases," *IEEE International Conference on Multimedia Computing and Systems*, June 1996

[3] L.A. Rowe, D.A.Berger. and J.E Baldeschwieler,"The Berkeley Distributed Video-on-Demand System". *Multimedia Computing – Proceedings of the Sixth NEC Research Symposium*, 1996

[4] A. Geist, A Beguelin. J.Dongarra, W.Jiang,R Manchek. and V. Sunderamm, "PVM 3 User's Guide and Reference manual." *Oak Ridge National Laboratory*, Oak Ridge, Tennesee 37831, Sept. 1994

[5] Myricom , "MYRINET USER'S GUIDE" . *http //www.myri com.80/scs/documentation/mug*