

Tertiary 저장장치를 사용한 비디오 서버

이종민^U 오재학 차호정
광운대학교 컴퓨터학과
{jmlee,ojh,hojungc}@cs.kwangwoon.ac.kr

Video Server with Tertiary Storage

Jongmin Lee^U Jaehak Oh Hojung Cha
Dept. of Computer Science, Kwangwoon University

요약

본 논문은 Tertiary 저장장치를 이용한 대용량 비디오 서버의 설계와 구현에 대해서 기술한다. 서버는 서비스와 프로세스 관리 서브시스템, 자원 관리 서브시스템, 입출력 관리 서브시스템의 세 부분으로 구성된다. 입출력 대역폭이 낮은 Tertiary 저장장치의 효율적인 사용을 위해 디스크 캐쉬 풀 모델을 사용하고, 적절한 스케줄링 정책과 캐싱 알고리즘을 이용하여 가능한 많은 사용자를 수용한다. 서버는 Windows 2000에서 시스템 서비스로 구현되었으며 디스크 캐쉬의 효과적인 사용으로 Tertiary 저장장치에 낮은 대역폭을 최대한 활용하여 접속된 사용자에게 안정된 비디오 스트리밍 서비스를 제공한다.

1. 서론

최근 컴퓨터 하드웨어와 네트워크 기술의 발전으로 인터넷을 이용한 방송 및 주문형 비디오(VOD) 서비스가 활성화되고 있다. 주문형 비디오 서비스는 안정된 서비스를 위한 연속된 스트리밍 보장과 다양한 사용자의 욕구를 만족시키기 위한 많은 콘텐츠의 확보가 필요하다. 기존의 디스크 기반의 비디오 서버는 고품질의 서비스와 안전성을 보장하는 반면에 저장장치 구성에 많은 비용이 든다[1].

광 주크박스나 테잎 드라이브 같은 Tertiary 저장장치는 미디어들을 보관하는 슬롯과 미디어를 읽기 위한 다수개의 드라이브 그리고 슬롯과 드라이브 간의 미디어 교환을 위한 로봇암으로 구성된다. Tertiary 저장장치는 비교적 적은 비용으로 대용량의 저장공간을 확보할 수 있는 반면에 미디어 교체 시간 등의 지연으로 대역폭이 낮은 단점이 있다. 이러한 Tertiary 저장장치를 멀티미디어 서버로 활용하기 위해 디스크 버퍼를 두는 방법이 있다[2].

본 논문에서는 광주크박스 Tertiary 저장장치를 비디오 저장장치로 사용하고 디스크를 캐싱 장치로 이용하는 비디오 스트리밍 서버를 설계하고 구현한다. 서버 설계시 Tertiary 저장장치를 효율적으로 활용하기 위한 Tertiary 저장장치, 디스크, 메모리의 계층적 자료 관리와 Tertiary 저장장치의 낮은 대역폭을 보완하기 위한 적절한 알고리즘 및 스케줄링 정책이 필요하다. 논문의 구성은 다음과 같다. 2장에서 서버의 구조와 각 구성 요소들의 기능들에 대하여 설명하고 3장에서는 결론 및 향후 과제를 기술한다.

2. 서버 구현

구현한 VOD 서버는 비디오 저장용도로 광 주크박스 저장장치를 사용하고 중간에 캐싱을 위한 디스크 장치를 갖는다. 서버는 디스크 캐쉬를 두어 Tertiary 저장장치의 낮은 대역폭을 보완하고 EDF에 기반한 스케줄링 정책으로 사용자에게 안정된 연속적인 비디오 스트림을 보장한다[3]. 이 때 서버는 수용제어 과정을 통하여 안정된 비디오 스트리밍 서비스를 클라이언트에게 제공한다.

그림 1은 서버의 구조를 보여준다. 서버는 크게 서비스와 프로세스 관리 서브시스템, 자원 관리 서브시스템, 입출력 관리 서브시스템의 3가지 관리 서브시스템으로 구성된다. 서비스와 프로세스 관리 서브시스템은 사용자 수용제어와 사용자에게 연속된 비디오 서비스를 위한 스케줄링 정책을 결정한다. 자원 관리 서브시스템은 서버의 자원 사용량을 동적으로 관리한다. 입출력 관리 서브시스템은 Tertiary 저장장치에서 디스크 캐쉬로의 자료 입출력을 담당하며, 효과적인 자료의 입출력을 위한 스테이징 스케줄링, 캐쉬 관리 정책 등을 수행한다[4]. 다음은 서버의 관리 시스템과 스테이징 방법에 대하여 설명한다.

서비스와 프로세스 관리 서브시스템 클라이언트의 서비스 요구를 처리하며 클라이언트 수용제어 및 안정된 스트리밍 서비스를 위한 쓰레드 스케줄링을 수행한다. 클라이언트의 요구를 받아들이는 Listener, 수용제어기, 시스템 스케줄러가 여기에 속한다. 수용제어기는 자원 관리 서브시스템을 통해 현재 시스템의 자원 사용량을 검사하여 사용자의 수용 여부를 결정한다. 수용제어기에 의해 수용 허가된 클라이언트의 정보는 자원 관리 서브시스템에 등록되어 자원 관리 대상에 포함된다.

• 본 연구는 한국과학재단의 특정기초연구사업(97-01-00-12-01-5)에 의해 지원받았음.

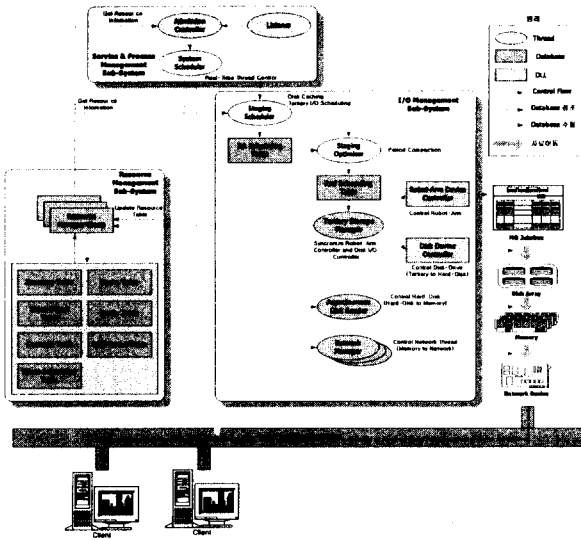


그림 1. 서버 구조

시스템 스케줄러는 실시간 스케줄러로 주기적으로 실행되는 주기모델이며 클라이언트의 요구율에 따른 위급도를 정하여 EDF 방법에 의해 서비스 순서를 결정한다. 결정된 순서에 의해 입출력 관리 시스템의 네트워크 매니저를 제어하여 비동기 디스크 읽기 모듈에 의해 메모리에 옮겨진 자료를 네트워크를 통해 클라이언트에게 전달될 수 있도록 한다. 시스템 스케줄러는 입출력 관리 서브시스템의 스테이징 스케줄러를 활성화하여 광 주크박스와 디스크 캐쉬간의 자료이동에 관한 여러 정책을 수행할 수 있도록 한다.

자원 관리 서브시스템 서버의 시스템 자원의 사용량을 동적으로 기록, 관리한다. 관리되는 시스템 자원으로는 네트워크, 디스크 대역폭 사용량이 있다. 또한 서비스를 요청한 클라이언트에 대한 정보를 가지고 있는 클라이언트 정보 테이블, 서비스 가능한 영화의 이름, 대역폭 등을 저장한 비디오 객체 테이블, 비디오의 디스크 캐쉬 정보를 가지고 있는 캐쉬 테이블, 광 주크박스의 미디어 정보를 가지고 있는 미디어 정보 테이블, 광 주크박스의 상태를 기록하는 Tertiary 저장장치 상태 테이블 등, 시스템 내에 필요한 여러 가지 정보 등을 관리한다. 관리되는 각각의 다른 테이블들을 공통적인 함수의 집합으로 제어할 수 있도록 하여 내부 자료구조나 관리 방법에 대해 다른 모듈들에 대해 투명성을 제공한다. 내부 정보 테이블은 용도에 따라 이진 트리나 해쉬 테이블 구조로 관리되며 시스템의 비정상 종료들에 대비해 주기적으로 저장장치에 백업된다.

입출력 관리 서브시스템 Tertiary 저장장치와 디스크 장치로의 자료의 입출력에 관련된 부분을 담당한다. 광 주크박스를 제어하기 위한 로봇암 제어기와, 디스크 장치를 제어하기 위한 디스크 제어기 같은 Low Device를 제어하기 위한 모듈 외에

광 주크박스의 비디오 자료를 디스크 캐쉬로 옮기는 스테이징 작업을 효과적으로 하기 위한 스테이징 스케줄러, 스케줄링된 결과를 주기 압축 방법에 의해 압축하는 스테이징 최적화 모듈이 있다. 디스크 캐쉬의 내용을 메모리로 읽는 비동기 디스크 읽기 모듈, 메모리의 내용을 시스템에 접속된 사용자에게 전달하기 위한 네트워크 모듈 등도 여기에 포함된다. 스테이징 스케줄러는 시스템 스케줄러 주기의 정수 배의 주기를 가지며 광 주크박스에서 디스크로의 자료 이동 순서를 EDF 방법에 의해 결정한다. 자료이동은 캐싱 블록 단위로 이루어진다. 캐싱 블록의 크기는 광 주크박스의 미디어 교환시간과 비디오의 대역폭을 고려하여 결정되며, 네트워크 패킷 크기의 정수배의 크기를 갖는다.

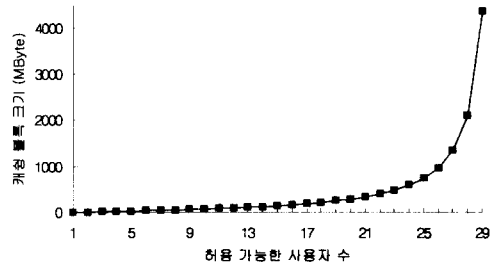


그림 2. 캐싱 블록 크기와 허용 가능한 사용자 수

그림 2는 디스크 캐쉬를 FIFO 구조로 했을 경우, 캐싱 블록의 크기와 동시 허용 가능한 사용자 수와의 관계를 보여준다. 스테이징 최적화 모듈은 스테이징 스케줄러의 마지막 단계를 독립된 모듈로 분리한 것으로 스테이징 스케줄러에 의해서 결정된 자료 이동 순서를 테드라인을 만족하면서도 광 주크박스의 미디어의 교체시간이 최소화 될 수 있도록 최적화 한다. 최적화는 주기 압축이라고도 하며 최적화의 결과로 여유 대역폭을 확보하여 더 많은 사용자를 수용할 수 있도록 한다. 로봇암 제어기와 디스크 제어기는 운영체제에서 지원하는 디바이스 드라이버를 이용한 일련의 명령어 집합들을 제공하여 디바이스 드라이버를 추상화 시킴으로써 다른 모듈들에게 디바이스 드라이버의 투명성을 보장한다. 광 주크박스 매니저는 스테이징 최적화 모듈로부터 나온 최종 결과를 바탕으로 로봇암 제어기와 디스크 제어기를 제어하며 로봇암의 움직임과 디스크의 읽기 쓰기 작업간의 동기화를 담당한다.

스테이징 요청된 비디오가 디스크 캐쉬에 없을 경우 Tertiary 저장장치에서 디스크로 자료이동을 하는 지연시간이 생긴다. 이러한 지연시간은 사용자가 직접적으로 느끼게 된다. 이를 해결하기 위하여 비디오의 일부분을 디스크로 미리 가져오는 자료 선반입을 수행한다. 비디오는 사용자 선호도에 따라 Cold, Partial, Hot 비디오로 구분된다. 자료 선반입의 양은 비디오의 선호도 구분에 따라 결정되며 캐싱 블록의 정수배의 크기를 갖는다. 선호도가 가장 낮은 Cold 비디오의 경우 비디오 요청시 그 비디오를 위한 미디어 교환 시간과 캐싱 시간 동안

