

2. 운영체제(I)

계층적 자료저장을 위한 광주크박스 관리 시스템

이 종민[○] 차 호정
광운대학교 컴퓨터과학과

A Magneto-Optical Jukebox Management System for Hierarchical Data Storage

Jongmin Lee and Hojung Cha
Department of Computer Science, Kwangwoon University

요 약

본 논문에서는 대용량 데이터의 계층적 자료저장을 위한 광주크박스 관리 시스템을 설계 하고 구현한다. 광주크박스를 이용한 계층적 자료저장을 함으로써 적은 비용으로 대용량의 저장공간을 확보하며, 적절한 스케줄링 정책에 의해 효과적인 자료 관리를 할 수 있다. 관리 시스템은 사용자와의 인터페이스를 위한 부분, 미디어의 관리를 위한 부분, 주크박스를 제어 하기 위한 부분으로 구성된다. 각 구현모듈들은 운영체제 독립적으로 설계되어 재사용이 가능하며 다양한 운영체제에 적용할 수 있다.

1 서론

컴퓨터 하드웨어 기술의 발달로 인하여 고속 계산과 네트워크 전송이 가능하게 되었다. 이러한 컴퓨팅 환경을 통하여 많은 양의 자료를 처리하는 다양한 응용 프로그램들도 나타나고 있다. 최근 방대한 양의 자료를 처리하고, 저장해야 할 필요성이 증가함에 따라 백업용으로만 인식되었던 제 3 저장장치(Tertiary Storage)에 대한 인식이 바뀌고 있다. 마그네틱 디스크는 자료의 임의 검색이 가능하고 속도가 빠른 반면 고가이다. 따라서 대용량의 데이터를 저장하기 위하여 가격대 성능비가 좋은 마그네틱 테이프같은 저장장치를 제 3 저장장치로 쓰게 되었다. 이는 자료 저장 능력이 좋으나 자료 검색, 전송 속도등이 상대적으로 낮다. 이러한 점들을 개선하기 위하여 광디스크 등을 이용하는 저장장치가 개발되고 있다. 주크박스(Jukebox)는 많은 수의 저장 미디어를 보관 할 수 있는 슬롯과 미디어를 읽고 쓰기 위한 다수개의 드라이브, 그리고 슬롯과 드라이브간에 미디어를 옮기기 위한 로봇암(Robot-arm)으로 구성된다. 일반적으로 주크박스에 쓰이는 미디어는 광디스크, 마그네틱 광디스크 등이 있다. 이러한 저장장치들이 출현 함에 따라 저장장치를 효율적으로 관리하기 위한 계층적 자료저장 관리(HSM: Hierarchical Storage Management)^[1]와 같은 다양한 연구들이 진행되고 있다. 계층적 자료저장 관리는 저장장치를 각각의 특성에 맞게 계층화하여 관리 하는 방법이다. 계층적 자료저장을 효율적으로 하기 위해서는 관리 시스템의 개발이 필수적이다. 본 논문에서는 범용 운영체제에서 광주크박스를 사용하여 계층적 자료저장을 위한 관리 시스템의 구현을 기술한다.

논문의 구성은 다음과 같다. 2절에서는 계층적 자료저장 관리와 광주크박스에 대하여 기술한다. 3절에서는 자료 관리 시스템의 구성, 주요 기능과 설계, 그리고 범용 운영체제에서 구현한 모습을 기술한다. 4절에서는 결론 및 향후 연구과제를 논한다.

2 계층적 자료저장 관리와 광주크박스

다음에는 계층적 자료저장 관리와 이를 구현하기 위해 사용한 광주크박스에 대하여 기술한다.

계층적 자료저장 관리 계층적 자료저장 관리란 저장장치를 각각의 특성에 맞게 계층화 하여 이를 관리하는 방법을 말한다. 그림1은 계층적 자료저장 관리의 개념을 보여준다. 계층적 자료저장 관리를 구성하는 기본적인 방법은 높은 자료 전송 능력을 가지며 상대적으로 고가인 저장장치를 계층구조의 상위에 놓고, 자료의 전송 능력은 떨어지나 가격이 저렴한 저장장치를 계층구조의 하단에 위치시킨다. 계층구조의 상단에 위치하는 저장장치의 예로는 마그네틱 디스크 드라이브가 있고, 계층구조의 하단에 위치하는 저장장치의 예로는 테이프 저장장치를 들 수 있다. 계층적 자료저장 관리의 동작과정은 계속적으로 자주 요구되는 자료는 계층구조의 상단에 저장하고 상대적으로 요구가 적은 자료는 계층구조의 하단에 저장하는 것이다. 이러한 계층적 자료저장은 전체적인 물리적 저장장치의 비용을 줄일 수 있는 장점이 있다. 같은 용량의 저장용량을 확보할 때 고가인 마그네틱 디스크로만 구성하는 것보다 상대적으로 가격이 저렴하다. 또한 자주 요구되는 자료는 계층구조의 상단에 위치하므로 빠른 제어가 가능하다는 장점도 있다.

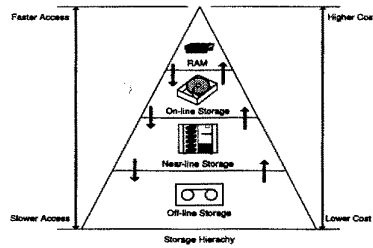


그림 1: 계층적 자료저장 구조

계층적 자료저장 관리에서 계층간의 자료 이동은 관리 시스템에 의하여 자동으로 이루어진다. 특히 언제 어떤 자료를 계층들 사이로 이동할 것인가에 대한 연구가 많이 되고 있다. Legato사의 Networker에서는 자료의 생성, 갱신 시간 및 자료의 종류, 크기의 조건이 만족 되었을 때 자료의 이동이 이루어지게 한다^[2]. 자료의 이동은 시스템이 유휴상태에 있을 때 하게 되어 시스템의 성능을 높인다. UniTree사의 UniTree는 상위 계층의 저장장치에 일정량 이상의 자료가 채워지게 되면 하위 계층의 저장장치로 이동하는 방법을 사용한다^[3]. 이러한 방법들이 자동으로 이루어져 사용자에게 저장장치의 투명성을 제공하며, 효과적인 자료 관리를 가능하게 한다.

광주크박스 광주크박스는 많은 수의 저장 미디어를 보관할 수 있는 슬롯과 미디어를 읽고 쓰기 위한 다수개의 드라이브, 그리고 슬롯과 드라이브간에 미디어를 옮기기 위한 로봇암으로 구성된다. 그림 2는 주크박스의 기본구조를 보여준다^[4]. 주크박스는 방대한 양의 자료를 저장할 수 있다. 주크박스에 쓰이는 미디어는 Tape, CD-ROM, MO, DVD 등이 있으며, 사용되는 미디어와 슬롯의 갯수, 로봇암의 갯수, 드라이브의 갯수에 따라 종류도 다양하다. 주크박스에서 쓰이는 미디어는 저장할 수 있는 자료의 양이 마그네틱 디스크보다 많으나 읽고 쓰는 작업은 상대적으로 느리다. 미디어에 있는 자료에 접근하기 위해서는 로봇암이 슬롯에 있는 미디어를 꺼내어 드라이브에 옮겨놓는 지연시간이 있게 된다. 이러한 지연시간은 시스템 성능의 저하를 가져오며, 지연시간을 줄이기 위한 여러가지 연구가 진행되었다. 입출력 요구를 재배열 하는 입출력 스케줄링 방법이 대표적인 예이다. 입출력 스케줄링은 다수의 자료요청 요구가 있을 때 미디어의 교환 시간과 찾는 시간을 최소한으로 하면서 모든 요구들을 만족시킬 수 있도록 입출력 요구를 재배열하는 방법이다^[5]. 또 다른 방법으로는 자료 선반입, 캐싱 등이 있다^[6].

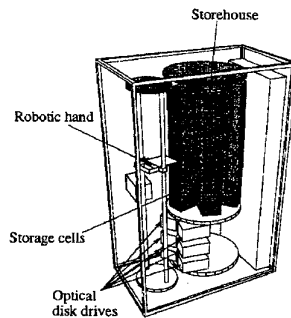


그림 2: 광주크박스 구조

3 광주크박스 관리 시스템

본 논문에서 구현한 광주크박스 관리 시스템은 마그네틱 디스크를 상위 계층의 저장장치로 사용하고 마그네틱 광디스크 주크박스(Magneto-Optical Jukebox)를 하위 계층의 저장장치로 사용한다. 마그네틱 광디스크는 마그네틱 디스크보다 읽고 쓰는 속도가 느린 반면, 자료의 안전성이 높고 일반 광디스크와는 달리 반복적으로 읽고 쓸 수 있다. 관리 시스템이 제공하는 기능은 주크박스를 사용하기 위한 기능과 계층적 자료저장을 위한 기능으로 구분된다. 주크박스를 사용하기 위한 기능은 사용자가 주크박스의 미디어를 자유롭게 사용할 수 있도록 한다. 이를 위해 가상 드라이브를 사용하여 주크박스내의 미디어를 통합된 인터페이스로 제공한다. 또한 주크박스내 작업의 투명성을 제공하기 위해 미디어 검색, 미디어 교환 등의 작업을 자동으로 한다. 한편, 계층적 자료저장을 위한 기능은 저장장치의 투명성을 제공한다. 즉, 사용자에게 요청된 자료가 어느 저장장치에 있는지 상관없이 동일한 서비스를 받을 수 있도록 한다.

3.1 관리 시스템 설계

관리 시스템의 각 모듈들은 운영체제 독립적으로 설계되어 대부분의 모듈들은 재사용이 가능하다. 운영체제 의존적인 부분은 각 운영체제별로 수정, 삽입, 교체가 가능하도록 구성하였다. 그림3은 마이크로소프트사의 Windows NT를 기반으로 설계한 관리 시스템의 구성을 보여준다. 관리 시스템은 사용자와의 인터페이스를 담당하는 부분 (가상 디스크 관리자), 미디어를 관리 하는 부분 (스케줄러, 볼륨 관리자), 주크박스를 제어하는 부분 (로봇암 관리자, 로봇암 디바이스 드라이버, 파일 입출력 관리자)의 세부분으로 구분한다. 각 모듈의 역할은 다음과 같다.

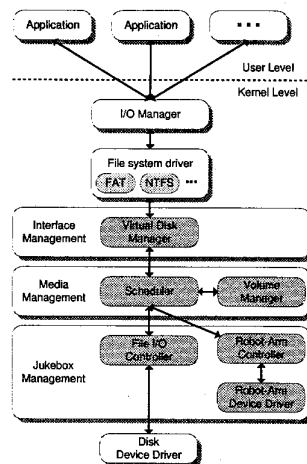


그림 3: 관리 시스템 구성

가상 디스크 관리자 사용자와의 인터페이스를 담당하는 부분으로 여러개의 물리적인 미디어들을 통합된 방법으로 접근 할 수 있는 방법을 제공한다. 주크박스를 나타내는 하나의 가상 드라이브를 제공하여 사용자가 이 가상 드라이브를 통해 주크박스를 제어할 수 있도록 한다. 즉, 사용자에게 주크박스를 하나의 디스크드라이브처럼 보이게 하여 내부적으로 일어나는 작동에 대하여 투명성을 제공한다. 사용자의 자료에 대한 제어 요구를 허가할 것인지 거절할 것인지를 결정도 이 모듈에서 수행된다. 허가된 사용자의 요구는 스케줄러에게 전달된다.

스케줄러 관리 시스템의 가장 핵심적인 부분으로 자료관리를 위한 여러가지 정책을 적용하는 모듈이다. 계층적 자료저장을 위해 상위 계층의 저장장치의 자료 저장량을 검사하는 방법과 시스템의 유휴시간을 이용하여 자료의 생성 및 갱신시간 또는 종류등의 선택을 통해 자료를 이동시키는 방법을 적용할 수 있다. 사용자의 자료 요구에 대해 효율적인 응답을 하기 위한 관리정책으로는 입출력 스케줄링 및 자료 선반입, 캐싱 등을 사용할 수 있다. 스케줄러는 가상 디스크 관리자로 부터 순차적으로 받은 요구를 큐를 이용하여 관리한다. 주기적으로 현재 큐에 들어있는 작업의 내용을 확인하고 각 작업이 요구하는 자료가 저장되어 있는 미디어와, 그 미디어의 현재 위치, 상태등의 정보를 볼륨 관리자로부터 얻는다. 이러한 정보들을 근거로 미리 정해진 자료관리 정책에 따라 큐의 내용을 재배열한다. 실행되기 위해 큐로부터 꺼내어진 자료 요청 요구는 미디어에 관계된 명령과, 자료의 입출력에 관계된 명령으로 나누어져 각각 로봇암 관리자와 화일 입출력 관리자에게 전달된다.

볼륨 관리자 미디어 관리를 위한 데이터베이스를 관리한다. 현재 미디어의 위치 정보를 위한 데이터베이스와 각 미디어가 가지고 있는 내용 정보에 관한 데이터 베이스를 관리한다. 각각의 정보들은 효과적인 자료의 관리를 위한 다양한 자료구조로 구성될 수 있으며, 자료의 검색을 위한 여러가지 검색 방법등이 적용될 수 있다. 대부분의 작업은 시스템의 메모리에서 이루어지고 시스템의 비정상 종료등에 대처하기 위해 주기적으로 메모리내의 자료를 저장장치에 백업한다.

로봇암 관리자와 로봇암 디바이스 드라이버 주크박스내에 있는 로봇암을 제어하여 슬롯에 있는 미디어를 드라이브에 로드시키거나, 언로드 시키는 작업등을 수행한다. 디바이스 드라이버는 사용되는 운영체제에 의존적인 부분이다. 이 부분은 운영체제에 따라 각각 다르게 작성 해야 한다. 이 위에 디바이스 드라이버를 제어하는 모듈을 추가하여 다른 모듈들에게 공통된 명령어 집합을 제공한다. 이 방법을 통해 다른 모듈들에게 디바이스드라이버의 투명성을 제공한다. 주크박스는 SCSI를 이용하여 시스템과 연결된다. 로봇암 관리자가 로봇암 디바이스 드라이버에 명령을 주면 로봇암 디바이스 드라이버는 받은 명령을 알맞은 SCSI명령어로 만들어 주크박스에 전달하게 된다.

화일 입출력 관리자 화일의 읽기, 쓰기등의 작업을 수행한다. 기존 운영체제에서 제공되는 디바이스 드라이버를 이용하며 자료의 읽기, 쓰기등의 작업을 쉽게 할 수 있는 명령어 집합을 제공한다. 로봇암 관리자와 마찬가지로 운영체제 의존적인 하위 디바이스 드라이버를 추상화 시킴으로써 다른 모듈들에게 디바이스 드라이버의 투명성을 제공한다.

관리 시스템이 사용자의 자료 접근 요구를 수용하는 과정은 다음과 같다. 먼저 사용자 요구가 가상 디스크 관리자를 통해서 들어오게 된다. 가상 디스크 관리자는 사용자 요구의 승인, 거절을 결정하며 승인된 요구는 스케줄러에게 순차적으로 보내진다. 스케줄러는 가상 디스크 관리자로부터 받은 자료를 내부의 큐를 이용하여 관리한다. 큐를 이용하여 자료 관리 정책에 따라 수행할 작업의 우선순위를 결정하고 큐의 내용을 재배치한다. 실행될 요구는 로봇암 관리자에 대한 명령과 화일 입출력 관리자에 대한 명령으로 나누어 각각의 관리자에게 적절히 분배된다. 이러한 과정에서 발생하는 미디어나, 미디어 내용의 변화는 바로 볼륨 관리자에게 보내져 내부의 데이터베이스를 갱신하도록 하며 볼륨 관리자는 시스템의 비정상 종료등에 대비하기 위해 데이터베이스를 주기적으로 저장장치에 백업한다.

계층적 자료를 위한 방법으로 관리 시스템에 특정한 조건과 수행 작업을 할당할 수 있다. 관리 시스템은 디몬(daemon) 형태로 시스템에서 실행된다. 시스템 관리자가 설정한 특정 시간대나 특정 조건이 맞았을 때 스케줄러가 실행되어 미리 설정된 작업이나 자료 계층구조의 상위 계층과 하위 계층의 자료 이동을 자동으로 수행한다.

3.2 관리 시스템 구현

구현한 관리 시스템은 마이크로소프트 Windows 2000 운영체제에서 동작한다. 사용된 주크박스는 HP SureStore 80ex 이며, 사용되는 미디어는 5.2GB 용량의 마그네틱 광디스크이다. 주크박스는 펜티엄-II 시스템과 SCSI-2로 연결되어 있다.

관리 시스템은 디몬으로 시스템의 메모리에 상주한다. 스케줄러를 제외한 그 외의 모듈들은 DLL의 형태로 작성되었다. 가상드라이브는 윈도우 셸을 이용하여 탐색기에 포함시킴으로써 사용자가 사용하기 쉽도록 하였다. 또한 미디어의 한면당 하나의 폴더를 할당하여 미디어의 내용이 그 하부 폴더로 나타나도록 하는 미디어 기반의 인터페이스를 제공한다. 볼륨 관리자는 미디어를 관리하기 위한 데이터베이스와 미디어의 내용을 관리하기 위한 데이터베이스를 관리한다. 내부적으로 이진트리를 이용하여 자료관리를 하며 자료의 삽입, 삭제, 검색등의 기본적인 명령어 집합을 제공한다. 화일 입출력 관리자는 미디어에 관련된 기본적인 화일 제어에 관한 명령어 집합을 제공한다. 미디어에 쓰이는 화일 시스템은 윈도우에서 제공하는 FAT, NTFS등의 화일 시스템을 쓴다. 이렇게 함으로써 관리 시스템이 아닌 다른 응용 프로그램들도 광주크박스를 사용하기 쉽도록 하였다. 그림 4는 윈도우 탐색기의 실행 화면이다. 사용자가 탐색기의 자료를 클릭하면 자료 요청이 디몬으로 떠있는 스케줄러에게 전달된다.

스케줄러는 전달된 요청을 자료 관리 정책에 따라 처리하여 사용자에게 서비스 한다.

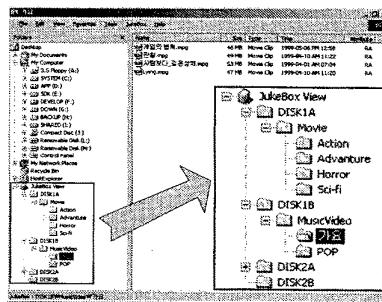


그림 4: 탐색기 실행화면 예

4 결론 및 향후 과제

본 논문에서는 계층적 자료저장 관리를 위한 광주크박스 관리시스템을 설계하고 구현하였다. 관리 시스템은 여러개의 모듈로 구성되어 있으며, 운영체제 의존적인 부분은 차후에 추가 및 교체가 용이하도록 설계되었다. 관리 시스템의 적절한 자료 관리 정책에 따라 광주크박스를 이용한 계층적 자료저장 관리를 함으로써 적은 비용으로 대용량의 저장공간을 확보하며, 적절한 스케줄링 정책에 의해 효과적인 자료 관리를 할 수 있다.

향후 연구 과제로는 주크박스의 미디어 교환시 발생하는 지연시간등을 최소화 하기 위한 입출력 스케줄링의 연구와 계층적 자료저장 관리를 효과적으로 하기위한 정책의 연구등이 있다. 또한 Linux와 같은 다른 운영체제에서의 구현과 일반적인 용도가 아닌 특수 목적의 자료 관리 정책 연구가 있다. 현재 본 연구실에서 개발한 VOD 시스템^[7]을 위해 광주크박스를 스트리밍 저장장치로 사용하기 위한 연구가 진행중이다.

참고 문헌

- [1] IBM , <http://www.storage.ibm.com/storage/software/adsm/adwhhsm.htm>.
- [2] Legato Systems Ltd., http://www.legato.com/Products/html/product_literature_library.html.
- [3] UniTree Software Ltd., <http://www.unitree.com/overview/overview.html>.
- [4] Makoto Mizukami, Shigetaro Iwatsu, Nobuyoshi Izawa, 'Real-Time Staging in Optical Disk Library', *Japanese Journal of Applied Physics*, January 1997, pp.568-571
- [5] Bruce K. Hillyer, Avi Siberschatz, 'Random I/O Scheduling in Online Tertiary Storage Systems', *ACM SIGMOD*, Vol.25, No.2, June 1996, pp.195-204.
- [6] S. Prabhaker, D.Agrawal, A.El Abbadi, A.Singh, 'Tertiary Storage: Current Status and Future Trends', *Technical Reports*, Computer Science Department, University of Columbia, TRCS96-21 , August 1995
- [7] 차호정, 박한규, 'QoS를 제공하는 비디오 서버 자원관리 시스템', 정보과학회 논문지(C), Vol.4, No.4, 1998년 8월, pp .449-459.