

계층적 압축방식 지원을 위한 멀티미디어 파일시스템

양승헌 박진연 원유집
한양대학교 전자전기컴퓨터 공학부
{shyang|jypark|yjwon}@ece.hanyang.ac.kr

Harmonic Data Placement for Layered Encoded Multimedia File

Seungheon Yang, Jinyeon Park, Youjip Won
Division of Electrical and Computer Engineering, Hanyang University

요 약

본 논문에서는, 적응적인 스트리밍 시스템 환경에서 계층적으로 압축된 멀티미디어 데이터의 저장구조와 자료인출 원리를 분석한다. 또한 네트워크 시뮬레이션을 이용, 네트워크 상에서 다양한 멀티미디어 스트리밍 부하의 특성을 분석한다. 두 가지 분석을 통해 기존의 멀티미디어 저장방식이 다양한 네트워크 환경과 시간에 따라 전송율이 변화하는 적응적 스트리밍 시스템의 요구를 충족시키지 못한다는 것을 확인한다. 이를 개선하기 위해 제안된 하모닉 인터리빙은 하모닉 인터리빙 방식은 전송비율이 높은 멀티미디어 데이터 블록의 위치를 연속적으로 배치하기 위해서 전송비율이 높은 계층과 그렇지 않은 계층을 별도의 그룹으로 나누어서 저장하는 방식이다. 실험을 통해서 네트워크 상황이 전송하고자 하는 영화의 대역폭을 보장하지 못하는 환경에서는 기존의 저장방식보다 새로 제안된 저장방식이 디스크 사용률에 있어 우수한 성능을 나타냄을 보인다.

1. 개요

1.1 연구동기

본 논문에서는 계층적 인코딩 방식으로 압축된 멀티미디어 파일의 적응적 스트리밍 환경에서 데이터의 연속성을 보장하기 위한 파일내 데이터의 새로운 배치 전략을 제안하고, 실험을 통해 이 성능을 검증하고자 한다.

1.2 관련연구

다양한 미디어들을 스트리밍 서비스에서 서로 구매 받지 않고 전송해 주는 문제에 관한 연구가 이루어졌다[1]. Shenoy외[2]는 미디어 타입에 맞게 최적화되어 분리된 파일 시스템 파티션보다 1개의 파일 시스템 프레임 워크 안에서 이기종간의 workload를 효과적으로 다룰 수 있는 것이 가능하다고 결론지었다. Chen외[3]은 2배속 재생, 역2배속 재생을 위한 화면 생략 기법을 제안하였다. Chen외[4]는 MPEG로 압축된 데이터의 우선순위를 고려한 데이터 저장방법을 제안하였다. 이것은 GOP내에 우선순위가 높은 프레임들을 앞쪽에 배치함으로써 적응적인 스트리밍 환경에서 디스크 효율을 높였다. Rejaie외[5]는 유니캐스트 전송에서 현재 네트워크의 상태에 따라 전송 스트리밍의 품질을 조절하는 알고리즘을 제안하였다.

디스크가 한바퀴 회전할 때 걸리는 시간인 T_R , 실린더간의 이동시간인 T_{l-to-l} 의 조합으로 구해질 수 있다. 1개의 실린더를 읽는데 걸리는 시간인 T_{cyl} 은 수식1 과 같이 수식화 될 수 있다.

$$T_{cyl} = ((\text{실린더당 디스크 표면의 수}-1) \times T_{\text{head_switch}} + T_{\text{cyl_switch}} + (\text{디스크 표면의 수}) \times T_R) \quad (1)$$

3. 멀티미디어 자료 저장 기법

이 장에서는 계층적 인코딩된 데이터의 파일 내 배치 전략을 설명하고 있다. n 은 멀티미디어 데이터의 총 계층을 나타낸다. T 는 전체 영화의 상영 시간을 나타내며, T_{seg} 는 논리적 저장 단위의 재생 시간(playback time)을 나타낸다. 각 계층의 총 세그먼트의 개수는 $m = T/T_{seg}$ 로 나타낼 수 있다. L_j 는 i 번째 계층의 j 번째 논리적 데이터 단위를 나타낸다($1 \leq i \leq n$, $1 \leq j \leq m$).

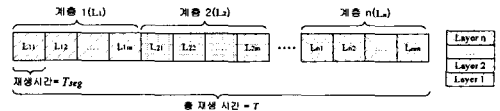


그림 1. 계층 별로 분리 저장 방식

계층 별로 분리 저장하는 방식은 그림 1과 같이 논리적 데이터 단위를 디스크에 배치 할 때, 각 계층 별로 나누어서 저장하는 방식이다. 이와 같은 저장 방식은 네트워크의 대역폭이 낮아서 하위 계층만을 전송할 때 유리하다.

인터리빙 저장방식은 그림 2와 같이, 각 계층을 세분하고 인터리빙(interleaving)하여 저장한다.

2. 디스크의 작동원리와 멀티미디어 자료인출의 효율성

디스크 입출력 대기시간을 유발하는 중요한 요소가 검색 오버헤드(seek overhead)이다. 특히 스트리밍 부하와 같이 파일을 순차적으로 접근하는 부하의 경우에는 검색에 따른 시간을 극소화 시키는 것이 절대 중요하다. 디스크 읽기 시간은 플래터간의 헤드가 이동하는 $T_{\text{head_switch}}$, 실린더의 마지막 데이터 블록에서 다음 데이터 블록으로 이동하는 시간인 $T_{\text{cyl_switch}}$

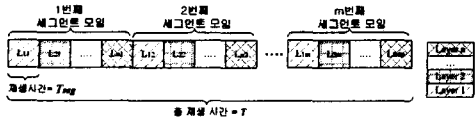


그림 2. 인터리빙 저장 방식

이와 같은 저장 방식은 상위 계층을 많이 전송할수록 디스크 회전 지연 시간이 감소한다. 하지만, 하위 계층 전송 비율이 높을수록 데이터만을 읽는데 걸리는 시간은 작아지는 대신에, 디스크 회전 지연 시간이 증가하는 단점이 있다.

하모닉 인터리빙 저장 방식은 그림 3과 같이 멀티미디어 데이터를 재생 시간 Δt 로 나는 후에, Δt 시간안에서 하위 계층 그룹과 상위 계층 그룹으로 나누어서 저장을 한다.

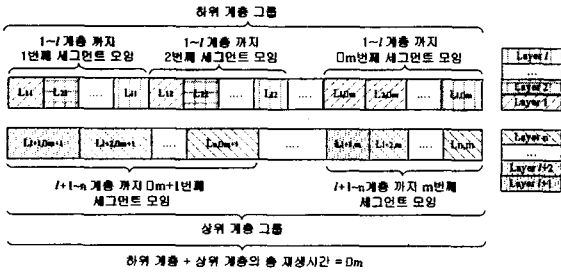


그림 3. 하모닉 인터리빙 저장 방식

즉 Δt 시간 동안의 논리적 데이터 유닛 ($Om = \Delta t / T_{seg}$)의 하위 계층을 앞 부분에 저장하고, 상위 계층을 뒷 부분에 저장하는 방식이다. 이 방식은 상대적으로 요청 빈도수가 낮은 상위 계층을 분리함으로써, 디스크 탐색 시간을 줄일려고 한다.

4. 모델링

각 저장 방식의 효율성을 나타내는 기준으로 디스크 사용률(utilization)을 사용하였다. 데이터를 읽는데 걸리는 시간을 T_{data} , 디스크를 읽는데 걸리는 총시간을 T_{total} , 회전 지연 시간을 T_{rot} , 검색시간을 T_{seek} 이라 하면 디스크 사용률(utilization)은 다음과 같이 수식화 될 수 있다.

$$utilization = \frac{T_{data}}{T_{total}} = \frac{T_{data}}{T_{data} + T_{overhead}} \quad (2)$$

$$T_{overhead} = T_{rot} + T_{seek} \quad (3)$$

T_{seg} 시간 동안의 계층 i 의 논리적 데이터 유닛을 읽는데 걸리는 시간을 t_i , 계층 i 의 접근 비율을 α_i 라 하면, 데이터를 읽는데 걸리는 총 시간은 수식 4과 같이 나타난다.

$$T_{data} = (T / T_{seg}) \times \sum_{i=1}^n \left(t_i \times \sum_{j=i}^n \alpha_j \right) \quad (4)$$

각 계층 별로 분리하여 저장 할 경우에는 디스크 회전은 계층 2 이상을 읽을 때는 항상 나타난다. 또한 계층 2 이상을 읽은 후에는 다시 계층 1로 돌아 와야 된다. i 번째 계층의 대역폭을 B_i , 영화의 대역폭을 B_{movie} , 계층 i 의 전송 비율, 디스크의 최대 전송율을 B_{max} , 영화가 차지하는 디스크 실린더의 수를 k 라고 하면, 각 계층이 차지하는 실린더의 수는 아래와 같이 정의된다.

$$S_i = \frac{B_i}{B_{movie}} \times k \quad (5)$$

T_{i_seek} 는 S_i 디스크 실린더를 이동 할 때의 탐색 시간이라 정의한다. 총 디스크 탐색 지연 시간은 수식 6와 같이 나타내어진다.

$$T_{overhead} = (T / T_{seg}) \left(\sum_{i=1}^n \left(T_{i_seek} \sum_{j=i+1}^n \alpha_j \right) + \sum_{i=2}^n \alpha_i \left(\sum_{j=1}^{i-1} T_{j_seek} \right) \right) \quad (6)$$

인터리빙 저장 방식은 멀티미디어 데이터가 연속적으로 배치되어 있다고 가정하면, 1개의 디스크 실린더를 읽은 후에 다음 디스크 실린더를 순차적으로 읽게 된다. 데이터를 읽는데 걸리는 총 시간은 아래와 같이 수식화 된다.

$$T_{total} = S_{total} \times T_{cyl} + (S_{total} - 1) \times T_{1-to-1} \quad (7)$$

하위 계층의 요청이 많을수록 디스크 읽기 시간 증가하고(T_{total}), 데이터를 읽는데 걸리는 시간(T_{data})은 감소하므로 디스크 사용률은 떨어지게 된다.

하모닉 인터리빙 저장 방식은 하위 계층 전송시는 인터리빙 저장 방식과 같이 수식화 되고, 상위 계층까지 전송시는 분리 저장 방식과 유사하게 수식화 될 수 있다.

하위 계층이 차지하는 실린더의 개수는 Δt 시간 단위로 저장을 하기 때문에 아래와 같이 수식화 된다.

$$S_{low} = \left(\frac{k}{T / \Delta t} \right) \times \frac{\sum_{i=1}^l B_i}{B_{movie}} \quad (8)$$

하위 계층에서 일어나는 디스크 읽기 시간은 아래와 같이 수식화 된다(수식 9).

$$T_{low_read} = S_{low} \times T_{cyl} \quad (9)$$

Δt 시간동안 상위 계층이 차지하는 실린더의 개수는 Δt 시간 단위로 저장을 하기 때문에 아래와 같이 수식화 된다.

$$S_{high} = \left(\frac{k}{T / \Delta t} \right) \times \frac{\sum_{i=l+1}^n B_i}{B_{movie}} \quad (10)$$

상위 계층까지 요청이 일어났을때의 디스크 읽기 시간은 해당 데이터를 읽고 다시 돌아오는 하위 계층으로 되돌아오는 시간의 합으로 수식화 된다.(수식 11).

$$T_{high_read} = T_{cyl} \times S_{1-to-l} + \Delta T / T_{seg} \times \left(\sum_{j=1}^{l-1} T_{j_seek} \right) \quad (11)$$

하위 계층과 그룹과 상위 계층 그룹을 모두 고려한 디스크 탐색 시간은 아래와 같이 수식화 된다.

$$T_{total} = T / \Delta t \left[T_{low_seek} \sum_{i=1}^l \alpha_i + \sum_{i=l+1}^n \alpha_i \times T_{high_read} \right] + (S_{total} - 1) \times T_{1-to-1} \quad (12)$$

5. 실험

모델검증을 위해 모델링한 수식과 디스크 인자를 사용해서 디스크 이용율을 수학적으로 계산하였고, 이를 대용량의 파일을 저장방식 따라 순차적으로 읽는데 걸리는 시간과 비교하였다.

5.1 실험환경

본 실험은 펜티엄 III 750Mhz 시스템에서 리눅스 커널 버전 2.4.18-3 상에서 실시하였다. 네트워크 트래픽 특성을 얻기 위해 사용된 시뮬레이터는 네트워크 시뮬레이터 2[6]로 버전은 2.1 beta 8a 를 사용하였다. 하드디스크는 SCSI 인터페이스 방식의 IBM Ultra 36LP DPSS-309170를 사용하였다.

네트워크 가정은 그 단순함을 위해 두개의 RAPP기반 전송과 TCP기반 전송만을 가정한다(그림 4). 각 전송은 유니캐스트 전송으로 가정하고 각 세션의 양단 노드는 일대일로 대응한다.

본 실험에서는 병목 연결구간의 대역폭을 조절함으로써 얻어낸 트래픽의 변화를 계층별 전송율로 사용한다.

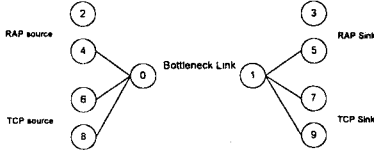


그림 4 Simulation Topology

5.2 모델 검증

본 논문에서 사용된 샘플 파일은 5개의 계층으로 구성된 40분(2400초) 분량의 영화로써 각 계층의 구분은 표 1과 같다. 멀티미디어 트래픽인 RAP 프로토콜과 병목 연결구간에서 경쟁하게 되는 TCP는 Sack1을 사용하며 병목큐(Bottleneck Queue)는 Drop tail 기법을 사용한다.

표 1 사용자 네트워크 환경을 고려한 계층의 분할[7]

계층	대역폭	누적대역폭	대용 네트워크 환경
1	34 kbps	34 kbps	56k Dial-up
2	46 kbps	80 kbps	128k Dual ISDN
3	270 kbps	350 kbps	384k DSL or Cable Modem
4	350 kbps	700 kbps	768k DSL or Cable Modem
5	800 kbps	1500 kbps	DVD Quality

모델을 검증하기 위해서 병목 연결구간의 대역폭을 변화시켜 가며 구한 레이어별 전송율을 이용, 계층별로 저장방식, 인터리빙 저장방식, 그리고 하모닉 인터리빙 방식에서 상위 계층과 하위 계층의 구분인 l 을 2~4까지 변화시켜가면서 실험을 하였다. 병목구간의 대역폭이 2400kbps인 환경에서 각 계층의 전송비율은 1계층부터 0.1%,0.3%,5.3%,94.3%,0%로 나타난다. 이 경우 모델링과 디스크 읽기 실험의 결과는 그림 5와 같다.

Disk Utilization(Bottleneck Link BW:2400Kbps)

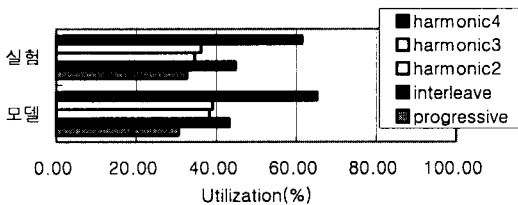


그림 5 디스크 사용율(Bottleneck bandwidth : 2400Kbps)

그림 5에서 가로축은 디스크 효율을 나타내고 세로축은 모델링한 결과와 디스크 읽기 실험에 의한 결과의 각 저장방법 디스크 사용율을 나타낸다. Harmonic2, Harmonic3, Harmonic4는 각각 l 값이 2,3,4이고 $\Delta t = 20$ 인 하모닉 인터리빙 방식을 나타낸다. 병목 연결구간의 대역폭이 2400Kbps인 경우 4개의 각 세션이 약 800kbps의 평균 전송율을 기록한다 이 경우 4 계층까지의 데이터 연속성에 최적화된 l 이 4인 하모닉 인터리빙 저장 방식이 가장 좋은 효율을 나타낸다. 1.5Mbps의 대역폭을 갖는 영화를 서비스 할 때 1Mbps급의 서비스 수준(Service Level)을 보장하는 네트워크 환경을 가진 ADSL이나 케이블모뎀으로 접속하는 단말이 주 서비스 대상이라면 l 이 4인 하모닉 저장방식으로 저장하는 것이 디스크 효율을 높일수 있다

병목 연결구간의 대역폭이 1440kbps인 환경에서 각 계층의 전송비율은 1계층부터 0%,14%,85.3%,0.6%,0%로 나타난다.

이 경우 모델링과 디스크 읽기 실험의 결과는 그림 6와 같다.

Disk Utilization(Bottleneck Link BW:1440Kbps)

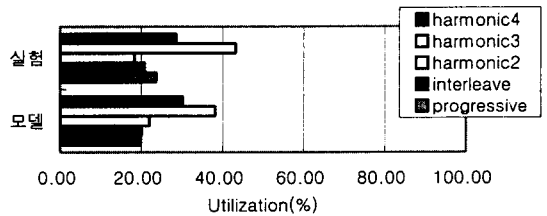


그림 6. 디스크 사용율(Bottleneck bandwidth : 1440Kbps)

병목 연결구간의 대역폭이 1440Kbps인 경우 4개의 각 세션이 약 360Kbps의 평균 전송율을 나타내며 3 계층을 요구하는 비율이 가장 많게 나타난다. 이 경우 3 계층까지의 자료 연속성에 최적화된 l 이 3인 하모닉 인터리빙 저장 방식이 가장 좋은 효율을 나타낸다. 1.5Mbps의 대역폭을 갖는 영화를 서비스 할 때 500Kbps급의 서비스 수준을 보장하는 네트워크 환경을 가진 Cable-lite급이나 홈랜으로 접속하는 단말이 주 서비스 대상이라면 l 이 3인 하모닉 저장방식으로 저장하는 것이 디스크 효율을 높일수 있다

6. 결론

본 논문에서는 계층적 인코딩 기법으로 제작된 멀티미디어 파일을 효과적으로 전송하는 목적으로 최적화된 파일 저장구조를 연구하고 이 구조의 성능 측정과 분석에 초점을 맞추었다. 이 연구의 결과로 계층적 인코딩된 파일의 전송에 최적화된 새로운 멀티미디어 데이터 배치방법인 하모닉 인터리빙 방식을 제안했다. 하모닉 인터리빙 방식은 파일을 일정한 구간으로 나누고 전송비율이 높은 계층과 낮은 계층의 데이터를 분리하여 각각 데이터를 연속적으로 배치함으로써 데이터 인출시 발생하는 디스크 오버헤드를 줄였다. 성능평가 결과 하모닉 인터리빙 방식은 고품질 대역폭이 보장되지 못하는 네트워크 상황에서 기존의 저장방식보다 디스크 회전지연시간과 디스크 검색시간을 단축시켜 디스크 효율을 향상시켰다.

7. 참고 문헌

- [1] Youjip Won and Y.S. Ryu, " Handling Sporadic Tasks in Multimedia File System", In *In Proc. of ACM Multimedia Conference '00, Los Angeles, CA, USA, Oct. 2000*
- [2] Prashant Shenoy, Pawan Goyal, Harrick M. Vin. " Architectural considerations for next generation file systems", *Proceedings of the seventh ACM international conference on Multimedia*, 1999
- [3] M. S. Chen, D. Kandlur, and P. Yu, " Support for fully interactive playout in a disk-array-based video server" *in Proc. ACM Multimedia94*, San Francisco, Oct. 1994
- [4] H. Chen, T. Little, and D. Venkatesh, " A storage and retrieval technique for scalable delivery of MPEG-encoded video," *J. Parallel and Distributed Computer.*, vol. 30, no. 2, pp. 180.189, Nov. 1995.
- [5] R.Rejaie, M.Handley, and D. Estrin, " Layer Quality Adaptation for Internet Video Streaming" , *In Proc. Of ACM SIGCOMM99*, Oct. 1999
- [6] S. McCanne and S. Floyd. Ns (network simulator). 2002. <http://www-mash.cs.berkeley.edu/ns>.
- [7] Real Networks. Inc. Helix Producer USER GUIDE, 2002