

대화형 서비스를 고려한 디스크 스케줄링의 디스크 이용률 향상 방안

이용수* 이승원 김인환 정기동
부산대학교 전자계산학과
{peace123, bluecity, ihkim}@melon.cs.pusan.ac.kr, kdchung@hyowon.cc.pusan.ac.kr

Scheme for Improving Disk Utilization of Disk Scheduling Considering Interactive Service

Young-Soo Lee Won-Sung Lee In-Hwan Kim Ki-Dong Chung
Dept. of Computer Science, Pusan National University

요 약

본 논문에서는 디스크의 이용률을 향상시키기 위하여 효율적인 디스크 스케줄링 알고리즘을 제안한다. 디스크 요구는 제공되는 서비스의 레벨에 따라 주기적인 요구, 비 주기적인 요구, 대화형 요구로 나눈다. 기존의 멀티미디어 서비스를 위한 디스크 스케줄링은 공정성 및 디스크 대역폭 보장만을 고려하였는데 본 논문에서는 대화형 요구를 고려하여 디스크 스케줄링의 공정성 및 주기적인 요구의 대역폭을 보장하고 디스크 이용률을 향상시켜 처리 가능한 요구 수를 증가시킬 수 있는 디스크 스케줄링을 제안한다.

1. 서론

기존의 디스크 스케줄링은 크게 3가지 범주로 분류할 수 있다. 첫째, best-effort 응용을 위한 알고리즘으로서 디스크 헤드의 탐색의 최적화를 위한 알고리즘이다. 둘째, 실시간 요구의 마감시간을 보장하는 알고리즘이다[1][2][3][4][5]. 셋째, 위 두 가지 목적을 동시에 만족하는 알고리즘들이다 [6][7][8].

VOD 서버에서 멀티미디어 객체는 주기적인 연성 실시간 특징을 가지는 객체 외에도 빠른 응답 시간을 요구하고 불규칙적으로 발생하는 객체가 있다. 이런 객체들의 특성을 고려하여 디스크의 이용률을 향상시키고 처리 가능한 요구 수를 증가시킬 수 있도록 한다. 그러나 위에서 언급한 세가지 디스크 스케줄링 알고리즘들은 이러한 조건들을 만족시키지 못한다. 즉 best-effort 응용을 위한 디스크 스케줄링의 목적은 평균응답시간을 최소화 시키기 위한 것이고, 실시간 응용을 위한 디스크 스케줄링의 경우 실시간 시스템을 목적으로 설계 되었기 때문에 범용 운영체제에서는 적합하지 못한 디스크 스케줄링 기법이다. 마지막으로 혼합형 또한 기존 방법의 변형에 그치고 있다.

본 논문에서는 대화형 요구를 고려하여 VOD 서버의 디스크 이용률을 향상시켜 처리 가능한 요구 수를 증가시키면서 실시간 태스크의 디스크 대역폭을 보장할 수 있도록 한다.

실시간 태스크는 시간의 제약성에 따라서 마감시간 상실이 전체 시스템 성능에 치명적인 영향을 미치는 경성 실시간 태스크와 시스템 성능 저하가 완만한 연성 실시간 태스크로 나눌 수 있다. 본 논문에서는 범용운영 체제에서 멀티미디어 서비스를 위한 연성 실시간 태스크 스케줄링을 다룬다.

다음 장에서는 본 연구와 관련된 관련 연구를 살

펴보고, 3장에서는 디스크 배열상에서 액세스 시간을 설명한다. 4장에서는 디스크 스케줄링 모델과 디스크의 이용률을 향상시킬 수 있는 디스크 스케줄링 기법을 기술한다. 5장에서는 제시된 기법을 시뮬레이션 하여 측정결과를 제시하고 6장에서는 결론 및 향후 연구과제에 대해서 언급한다.

2. 관련연구

탐색 시간의 최적화를 위한 알고리즘으로 FCFS(First Come First Served), SSTF(Shortest Seek Time First), SCAN, LOOK, STR(Shortest Total/Access Time First), ASAF(Aged Shortest Access Time First) 등이 있는데 이들은 대화형 best-effort 응용을 위하여 평균 응답 시간을 최소화하는데 목적이 있다. 그러나 이러한 알고리즘은 VOD 응용의 실시간 요구를 보장할 수 없다.

마감시간 보장을 위한 알고리즘으로 EDF[1], PSCAN (Priority SCAN)[2], ED -SCAN(Earliest Deadline SCAN) [3], FD-SCAN(Feasible Deadline SCAN)[4], SCAN-EDF[5] 등이 있는데 이들은 마감시간 이전에 디스크 블록을 서비스 하는 것이 목적이다. 또한 이 알고리즘들은 주로 경성 실시간 응용으로 개발되어서 긴박한 요구는 만족할 수 있으나, 시스템 전반적인 처리율은 저하되고, 특히 최대 사용자를 서비스를 목적으로 하는 VOD에는 적절하진 않다.

탐색시간의 최적화와 마감시간 보장을 동시에 만족하기 위한 방법으로 비실시간 요구를 마감시간을 가지는 실시간 요구로 전환하여 모든 요구를 실시간으로 처리하거나[6], 각 요구들을 우선순위로 분류하는 방식[7], 그리고 각 요구들을 서비스 클래스로 분류하여 서비스 시간을 다양하게 할당하는 방식[8] 등이 있다

3. 디스크 배열상에서 디스크 액세스 시간

디스크 액세스 시간은 식 (1)로 나타낼 수 있다.

$$T_{access} = T_{seek} + T_{rot} + T_{tr_zone(i)} \quad (1)$$

T_{seek} 은 탐색 시간, T_{rot} 은 회전 지연 시간, $T_{tr_zone(i)}$ 은 MZR 디스크의 데이터 전송시간이다. 위와 같은 식을 기준으로 하여 기존에 연구에서 주기적 요구와 비주기적 요구를 처리하는데 걸리는 디스크 액세스 시간은 식(2)와 같다[9]. 그리고 하나의 디스크 상에서 정밀한 회전 지연 시간 모델을 이용하여 디스크 액세스 시간을 예측하는 방법이 있다[10]. 또한 하나의 MZR 디스크 상에서 배치기법에 따른 디스크 액세스 시간을 계산하는 방법 등이 있다[11].

$$T_{i,j} = 2 \times t_{seek} + \sum_{i=1}^n \left\lfloor \frac{t_{i,j}}{B} \right\rfloor \times t_{rot} + \sum_{i=1}^n \left\lfloor \frac{t_{i,j}}{R} \right\rfloor \times B \quad (2)$$

본 논문에서는 디스크 배열상에서 디스크 개수, 스트라이핑 유닛의 크기 등 여러 가지 요소들을 고려하여 정확한 디스크 액세스 시간을 구할 수 있는 공식을 구한다. 이 액세스 시간을 사용하여 현재 서비스 받을 수 있는 주기적인 요구에 대한 스트림 수를 결정하고 주기마다 서비스 될 수 있는 각각의 요구들의 수를 결정한다. 본 논문에서는 액세스 시간을 정확히 구한다는 가정 하에서 디스크 스케줄링을 제안한다. 그리고 디스크 액세스 시간에 대해서는 향후 연구과제로 남긴다.

4. 디스크 스케줄링

본 논문에서는 VOD 서버의 실시간, 비실시간 그리고 대화형 요구를 동시에 만족하고, 실시간 요구의 독립성과 공정성도 동시에 만족하면서 또한 디스크 이용률을 향상시킬 수 있는 2단계 스케줄링을 제안한다.

4.1 시스템 모델

본 논문에서는 리눅스 커널과 같은 운영 체제에서 VOD 서버를 위한 디스크 스케줄링을 위해 다음의 세 가지를 고려한다.

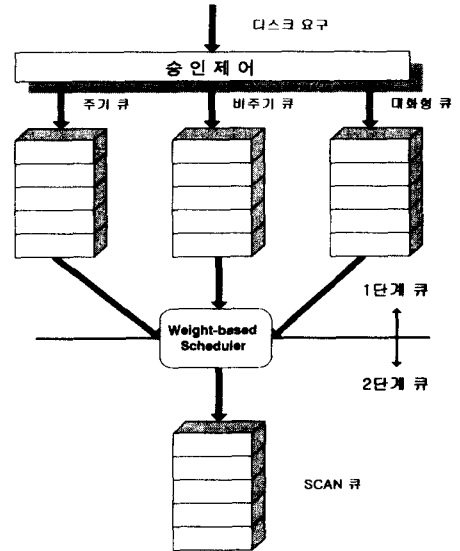
- 첫째, 주기적, 비주기적, 대화형 요구를 동시에 만족
- 둘째, 주기적 요구는 VOD를 대상으로 하므로 제일 높은 우선순위를 부여하고 다른 요구에 의해 희생되지 않는다. 대화형 요구는 VCR 연산이기 때문에 주기적 요구 다음의 우선순위를 부여한다.
- 셋째, 대화형 요구는 최소 지연시간 내에서 처리하고, 비주기적 요구는 기아현상이 발생하지 않는 범위 내에서 처리

본 논문에서 제시하는 디스크 스케줄링 기법은 2단계 큐를 사용하여 위에서 제시한 고려사항을 만족하고자 한다. 2단계 스케줄링은 <그림2>에서와 같이 보듯이 1단계 큐와 2단계 큐로 구성된다.

<그림2>에서 1단계 큐(주기적, 비주기적, 대화형)에서는 각각의 큐별로 할당된 가중치(디스크 대역폭)만큼 각각의 요구들이 SCAN 순서로 정렬된다. 2단계 큐에서는 태스크의 종류에 관계없이 여러 요구들의 탐색시간을 최적화하기 SCAN 디스크 스케줄링을 사용하였다. 가중치 기반의 스케줄러(Weight-based Scheduler)는 1단계 큐에서 각각의 큐에 가중치를 할당하고 한 주기내에 처리될 수 있는 요구들을 2단계 큐로 보낸다

4.2 가중치 기반의 스케줄러

가중치 기반의 스케줄러(Weight-based Scheduler)는 1단계 큐에서 각각의 큐에 가중치를 할당하고 한 주기내에 처리될 수 있는 요구들을 2단계 큐로 보낸다. VOD 서버에서는 실시간 클래스의 요구가 많기 때문에 각각의

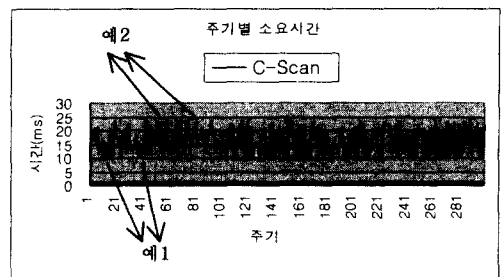


<그림1> 2단계 디스크 스케줄링

큐에 할당되는 가중치는 주기적인 큐에 많은 가중치를 부여한다. 대화형 요구는 VCR 연산이 많으므로 그 다음으로 많은 가중치를 부여한다. 서버 관리자는 초기에 각각의 큐에 가중치를 지정하여(예: 주기적 : 0.7, 대화형 : 0.2, 비주기적 : 0.1) 서버를 동작시킨다.

<그림2>와 같이 주기적인 요구들은 주기마다 요구하는 블록들의 액세스 시간의 편차가 생긴다. 이 액세스 시간의 편차를 이용하여 디스크 이용률을 향상시킨다. 해당 주기내에 서비스 중에 주기적 요구가 희생 시키지 않는 범위 내에서 남은 디스크 대역폭을 필요로 하는 큐에 전달하고 반대로 주기적 요구가 더 많은 요구를 할 때는 다른 큐의 디스크 대역폭을 희생시킨다. <그림2>에서 예1, 예2를 통하여 디스크 이용률 향상 방법 및 대화형 요구의 최소 지연시간 보장과 비주기적 요구의 기아현상 방지 방법을 설명한다. 또한 예2에서 발생할 수 있는 디스크 액세스 시간의 차이에 의하여 처리되는 요구들을 설명한다.

1단계 큐에서 각각의 큐에 할당된 초기 가중치는 아래 <표 1>와 같다.

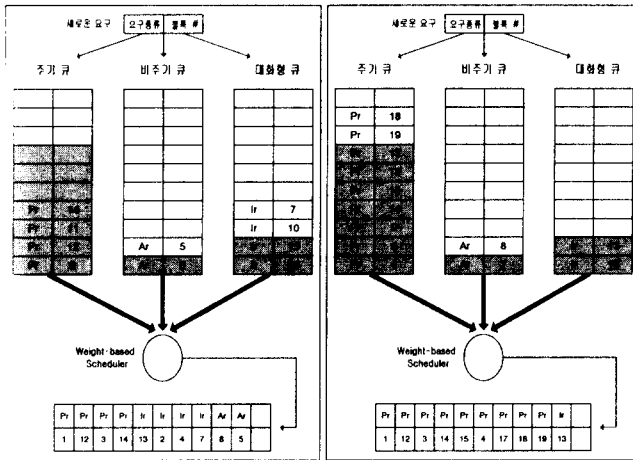


<그림2> 주기적인 요구의 주기별 액세스 시간의 편차

<표1> 초기 가중치 할당

Periodic Request(P _i)	Interactive request(I _i)	Aperiodic request(A _i)
70%	20%	10%

예1과 예2를 <그림3>의 (a)와 (b)를 통하여 설명한다



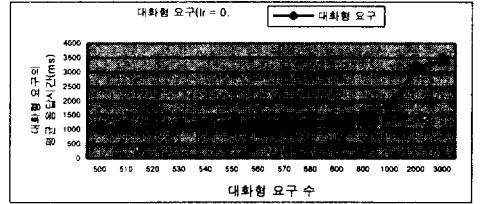
<그림3> (a) 예1, (b) 예2 설명

만약 디스크 대역폭이 주기 당 10블록을 처리 가능하다면 주기적인 요구는 7 블록, 대화형 요구는 2 블록, 비주기적인 요구는 1블록을 처리 가능하다. <그림3>에서 (a)는 주기적인 요구 4블록, 대화형 요구 4블록, 비주기적인 요구 2블록을 요구하고 있다. 주기적인 요구는 4블록을 처리하고 남은 3블록의 디스크 대역폭을 각각 대화형 요구와 비주기적인 요구에게 2블록, 1블록씩 할당하여 최대한 디스크 이용률을 향상시킨다. <그림3>에서 (b)는 주기적인 요구 9블록, 대화형 요구 2블록, 비주기적인 요구 2블록을 요구하고 있다. 주기적인 요구는 7블록을 처리할 수 있는 데 9블록을 요구하였으므로 추가적인 2블록은 대화형 요구와 비주기적인 요구를 희생하여 우선 주기적인 요구를 다 처리한다. 그 다음으로 우선 순위가 높은 대화형 요구를 처리한다. 그리고 나면 대화형 요구 1블록, 비주기적 요구 2블록은 다음 주기로 지연된다. 여기서 대화형 요구는 최소 지연시간을 지킬 수 있을 때 까지 다음 주기로 넘기고, 비주기적 요구는 기아 현상이 발생하지 않을 때 까지 다음 주기로 넘긴다. 만약 예2에서 주기적 요구의 마지막 블록이 18번이 아니고 5번이라고 가정하고 대화형 요구의 처음 블록 번호는 25번, 비주기적 요구의 처음 블록은 7번이라고 가정한다. 이때 주기적 요구의 마지막 요구 블록에서 떨어졌던 대화형 요구 25번 블록 하나를 처리하는 것보다 주기적 요구의 마지막 요구 블록 5에 인접한 비주기적 요구 블록 7번과 8번을 처리하는 액세스 시간의 거의 비슷하다면 디스크 대역폭의 이용률 향상을 위해 대화형 요구 25번 블록 보다는 비주기적인 요구 7번, 8번 블록을 먼저 처리한다.

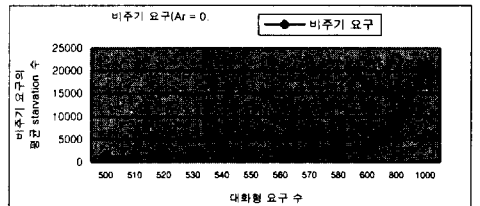
5. 실험 및 성능 분석

2단계 큐를 이용한 디스크 스케줄링의 성능을 평가하기 위하여 본 논문에서는 두 가지 실험을 하였다. 먼저 주기적 요구의 디스크 대역폭을 보장하면서 대화형 요구가 최대 응답시간 내에 처리 될 수 있는 방법을 제시하였고, 이들 실험한 결과를 <그림4>에 나타내었다. 그리고 대화형 요구 수가 증가 함에 따른 비주기 요구의 기아 수를 측정하였다. 이를 실험한 결과를 <그림5>에 나타내었다. 이들 실험에서 주기적인 요구는 MPEG1 비디오 20개를 포아송 분포를 따르는 50명의 사용자가 Zipf분포에 따라 비디오를 선택한다. 실험에서 멀티미디어 저장 서버의 라운드 길이는 1초이고, 대화형 요구의 최대 응답시간을 4초로하고, 비주기 요구는 1분으로 하였다. <그림4>의 결과는 대화형 요구가 증가 함에도 불

구하고 최대 응답 시간을 초과하지 않는다. <그림5>의 결과는 대화형 요구의 응답 시간을 만족하기 위하여 어느 정도 기아현상이 발생한다.



<그림4> 대화형 요구의 평균 응답 시간



<그림5> 대화형 요구수의 증가에 따른 비주기요구의 기아의 수

6. 결론 및 향후 연구과제

본 논문에서는 디스크 요청을 주기적, 비주기적, 그리고 대화형의 세가지 클래스로 분류하고, 각 클래스가 사용할 수 있는 디스크 대역폭인 클래스 가중을 할당하였다. 요청들은 먼저 1단계 큐 내에서 SCAN 순서로 정렬되고 가중치 기반 스케줄러에 의해 해당 클래스 가중의 한도 내에서 2단계 큐로 삽입되어 서비스된다. 동적 가중을 이용한 디스크 스케줄링과 가중치 기반 스케줄러에 의해서 디스크 이용률을 향상시켜 처리 가능한 요구 수를 증가 시켰고 주기적인 요구의 디스크 대역폭을 보장하면서 대화형 요구의 최대 지연 시간을 보장하였다. 향후 계획으로는 디스크 배열상에서 디스크 액세스 시간을 구하고 1단계 큐의 각각의 큐에서 2단계 큐로 요구들의 삽입 시 데이터 배치를 고려하여 삽입 위치를 결정하기 위하여 각각의 큐에 스케줄러를 삽입 할 계획이다.

7. 참고 문헌

- [1] C.L.Liu and J.W.Layland, "Scheduling algorithms for multiprogramming in a hard read-time environment," Journal of ACM, 20(1):46 -61, 1973
- [2] M. J. Carey, Jauhari, and M. Linvy. "Priority in DEMS Resource Scheduling", In Proceedings of the 15th VLDB Conference, 1989.
- [3] S.Chen, J.A. Stankovic, J.F. Kurose, and D. Towsley. "Performance Evaluation of Two New Disk Scheduling Algorithms for Real-Time Systems," Journal of Real-Time Systems, pp 307-336, 1991.
- [4] R. K. Abbott and H. Garcia -Molina, "Scheduling I/O Requests with Deadlines: A Performance Evaluation," Proceedings of RTSS, pp.113 - 124, 1990.
- [5] A.L. Narasimha Reddy and J. Wyllie. "Disk Scheduling in Multimedia I/O System," ACM Multimedia pp 225 -234, 1993.
- [6] J.P. Lehoczyk and S. Ramos -Thuel. "An optimal algorithm for scheduling soft-aperiodic tasks in fixed priority preemptive systems," In Proceedings of Real Time Systems Symposium, pp. 110/123, 1992
- [7] D.P.Anderson, Y.Osawa, and R. Govindan, "A File system for Continuous Media," ACM Trans. on Computer Systems, Vol.10 No.4 Nov, pp 313/337, 1992
- [8] J.S Prashant and M.V. Harick, "Cello: A Disk Scheduling Framework for Next Generation Operating Systems." SIGMETRICS, pp.44 -55, 1998
- [9] E. Park, N. Kim, S.Park, J.kim, and H.Shin. "Dynamic disk scheduling for multimedia storage servers". IEEE Region 10 Conference, 1999
- [10] 박상수, 박문정, 이수형, 신현식, "리눅스 멀티미디어 서버에서 정밀한 회전 지연 시간 모델에 의한 디스크 서비스 시간의 예측", 2000년 춘계 정보과학회
- [11] 양민영, "구간별 선반비에 기반한 VBR 데이터 스케줄링 및 수용제어", 학위 청구 논문, 1998