

디스크 쓰기 성능 향상을 위한 가장자리 영역 트랙의 이용

우 중 정[†] · 흥 춘 표^{††}

요 약

현재의 디스크 저장장치의 발전은 대부분 읽기 트래픽을 효율적으로 처리함에 따라 입출력 부하에 대한 쓰기 트래픽이 디스크 시스템 성능의 병목 현상이 되고 있다. 이를 큰 부담 없이 효과적으로 극복하기 위하여 본 연구에서 다영역 기록 디스크의 가장자리 영역을 2차 디스크 캐쉬로 사용하기를 제안한다. 가장자리 영역 디스크 캐쉬는 RAM 버퍼에 축적된 로그 형식의 데이터의 사용으로 인한 블록 전송과 트랙 전송에 대한 속도 차이, 가장자리 영역 트랙과 내부영역 트랙간의 전송률 차이, 디스크 캐쉬 트랙의 실질적 감소로 인한 탐색 시간 감소, 누적집단 접근 특성에 따른 휴지 기간 등을 이용함으로써 디스크 저장 장치의 성능 향상을 도모한다. 또한 실린더 단위로 캐싱 공간을 할당하므로 디스크 공간의 낭비가 거의 없다. 모의 실험 결과에 의하면 기존 디스크 시스템에 비하여 쓰기 연산에 대한 평균 응답 시간이 2.54에서 3.11배에 이르는 성능 향상을 보였다.

Using Outermost-Zone Tracks as a Cache to Boost Disk Write Performance

Jongjung Woo[†] · Chun Pyo Hong^{††}

ABSTRACT

Current disk systems are generally designed to reduce read traffic more effectively. Hence, write traffic of the I/O workload could potentially become a bottleneck of the disk system performance. In order to overcome this problem without much cost, this paper presents using outermost-zone track of multi-zoned recording disk as a secondary disk cache. The proposed disk cache improves the disk system performance by following exploitations: speed difference between block transfer and track transfer, difference in transfer rate between outermost-zone tracks and inner tracks, reduction in the seek time caused by decreasing the number of disk cache tracks, and idle period during burst accesses. In addition, it does not waste the disk space because it allocates the caching space by the cylinder unit. The simulation results show that the proposed system improves 2.54~3.11 times better in terms of average response time for write operations than existing disk systems.

1. 서 론

프로세서의 지속적인 성능 향상에 비하여 디스크 저

장 장치의 느린 발전 속도는 전체 시스템의 병목이 되고 있다. 현재 디스크 시스템은 일반적으로 RAM 캐쉬를 사용하여 성능을 향상시키고 있지만 RAM 캐쉬의 사용은 읽기 트래픽(read traffic)을 줄이는데 효과적이다. 읽기 트래픽의 축소로 시스템 성능에 대한 쓰기 트래픽(write traffic)이 상대적으로 비중이 커져 시스템 병목 현상이 되고 있다. 특히 RAID 디스크 배열

* This Paper was supported by the Grants for Professors of the Unjung Foundation in 1997.

† 종신회원 : 성신여자대학교 컴퓨터정보학부 교수

†† 정 회 원 : 대구대학교 컴퓨터정보공학부 교수

논문접수 : 1999년 8월 26일, 심사완료 : 1999년 10월 14일

구조에서의 처리율(throughput)은 소규모 쓰기에 대하여 기존 디스크 시스템에 비하여 4배정도의 영향을 받으며 이는 전형적인 입출력 부하를 지배하고 있어 시스템 성능에 대한 쓰기 트래픽의 중요성은 점차 커지고 있다[1].

쓰기 트래픽 문제를 극복하기 위하여 NVRAM(Non-Volatile Random Access Memory)의 사용, 로그 구조 파일 시스템(LFS : Log-structured File System), 그리고 일종의 2단계 디스크 캐쉬로써 작은 용량의 NVRAM을 상위 디스크 캐쉬로 디스크 드라이브내의 여러 개의 플래터(platters)중 일부를 하위 디스크 캐쉬로 사용하는 방법인 DCD(Disk Caching Disk)가 있지만[2], 고비용, 운영체제에 대한 변화 요구, 신뢰성 손실, 혹은 디스크 공간의 낭비 등의 단점이 있다.

본 연구에서는 기존 디스크 시스템의 단점을 없애는 반면 디스크의 성능을 향상시키기 위하여 다음의 세 가지 특성을 이용하고자 한다.

- 최근 디스크 시스템에 흔히 사용되는 다영역 기록 디스크(Multi-Zone Recording Disk)는 각 영역마다 데이터 전송 속도가 다르다. 즉, 가장자리 영역에 위치한 트랙이 내부 영역에 위치한 트랙보다 높은 전송 속도를 가진다[3, 4].
- 디스크 접근이 누적집단(burst)으로 발생함에 따라 많은 유휴 시간이 존재한다.
- 섹터 단위의 데이터 전송보다 트랙 단위의 데이터 전송이 빠르다[5].

본 연구에서는 DCD와 같이 몇 개의 디스크 플래터를 디스크 캐쉬로 이용하는 것이 아니라 모든 디스크 드라이브의 가장자리 트랙을 캐싱(caching)을 위한 공간으로 할당하여 다영역 기록 디스크의 특징인 영역별 전송 속도 차이를 이용함으로써 데이터 전송 시간, 탐색 시간(seek time) 등을 짧게 하여 성능을 개선하고자 한다. 본 연구에서 제안한 디스크 시스템의 성능 향상 여부를 검증하기 위하여 디스크 시뮬레이터로 개발된 DiskSim[6] 프로그램을 수정 및 보완하여 디스크 캐쉬 트랙을 위한 시스템 모델에 적합한 사건 구동 모의 실험을 수행하여 기존의 디스크 시스템과 비교하고자 한다.

본 논문의 구성은 다음과 같다. 2장에서는 기존에 제시되었던 디스크의 성능 향상 방법과 각 방법에 대한 여러 가지 문제점을 열거하며, 3장에서는 기존 디스

크 시스템의 문제점을 극복하기 위한 방법으로 디스크의 가장자리 영역을 디스크 캐쉬로 이용하는 방법을 제안한다. 4장에서는 제안된 방법에 대하여 시뮬레이션을 수행하여 기존 방법들과 성능을 비교 분석한다. 끝으로 5장에서 결론을 맺는다.

2. 관련 연구

자기 디스크(magnetic disk)는 비휘발성이며 주기억 장치로 사용되는 RAM에 비하여 가격이 저렴할 뿐만 아니라 많은 양의 데이터를 저장할 수 있기 때문에 컴퓨터의 저장 장치로써 아주 중요한 부분을 차지하고 있다. 그러나 디스크는 기계적인 특성으로 인하여 프로세서와 메모리에 비하여 성능 향상이 매우 느려 시스템 성능의 병목현상이 되고 있다. 디스크 시스템의 성능은 접근 시간(access time), 용량(capacity), 그리고 신뢰성(reliability) 등으로 나타낸다.

먼저 접근 시간에 대하여 살펴보면 다음과 같이 표현된다[7, 8].

$$T_{\text{access}} = T_{\text{seek}} + T_{\text{latency}} + T_{\text{xfer}} + T_{\text{overhead}}$$

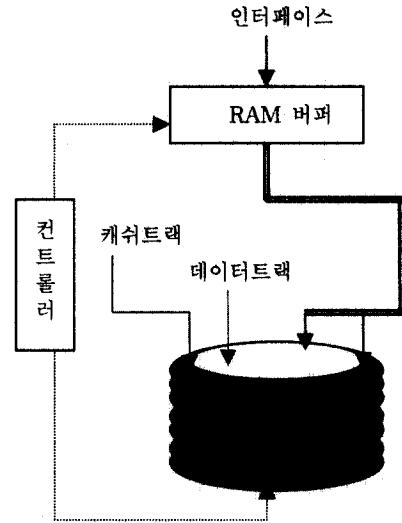
여기서 T_{seek} 은 탐색 시간으로 디스크 헤드를 원하는 트랙에 위치시키는 시간을 의미하며, T_{latency} 는 회전 지연 시간(rotational latency)으로 원하는 섹터를 디스크 헤드 아래에 위치시키는 시간을 의미하며, T_{xfer} 는 전송 시간(transfer time)으로 하나의 데이터 블록을 전송하는데 소요되는 시간을 의미하며, 그리고 T_{overhead} 는 디스크 컨트롤러가 입출력을 수행하는데 부과되는 오버헤드(overhead)를 의미한다. 자기 디스크의 기계적인 특성으로 접근 시간 단축이 매우 느린 속도로 발전했지만 80년 후반 미국 버클리 대학에서 제안한 RAID (Redundant Array of Inexpensive Disks)의 등장은 디스크 시스템 기술의 커다란 전환점이 되었다[8]. RAID는 여러 개의 디스크 배열로 구성되어 있으며 데이터 스트라이핑(striping)에 의한 작업 분배로 접근 시간을 향상하며 패리티(parity) 기반의 데이터 보호 기법을 적용하여 신뢰성에 대하여 기존 디스크 시스템보다 높은 성능을 도모한다. RAID-5의 성능은 한 번의 쓰기에 데이터와 패리티와의 일치성을 위하여 3번 혹은 4번의 디스크 접근이 필요하므로 RAID-5의 성능은 쓰기 트래픽에 크게 달려있다[1]. 현재의 디스크 시스템은 캐쉬를 이용하여 디스크 접근을 빠르게 하고 있지만 RAID-5 시스템에서 나타난 문제점과 같이 디스크

접근 시간에는 쓰기 트래픽의 비중이 커지고 있다. 이와 같은 문제점을 극복하기 위한 두 가지 방법(운영체제의 변화와 하드웨어의 추가)이 존재한다. 전자의 방법으로 LFS의 경우 일련의 파일의 변화를 버퍼에 축적한 후 모든 변화된 내용을 한 번의 순차적인 디스크 쓰기 동작을 수행함으로써 쓰기 성능을 개선한다. 이는 여러 번의 소규모 무작위 쓰기 동작이 하나의 커다란 순차적 전송으로 바꿈으로써 소규모 쓰기 동작에 대한 탐색 시간과 회전 지연 시간의 제거에 기인한다. 그러나 LFS는 운영체제에 대한 변화 요구 및 신뢰도 문제 등이 존재한다[9]. 후자의 방법으로는 NVRAM이나 디스크의 일부를 사용하는 DCD 등이 있다. 고비용의 NVRAM을 디스크의 쓰기 캐쉬로 사용할 경우 좋은 결과가 나타나지만, 작은 용량으로 입출력 동작과 같은 대규모 자료의 국부성을 이용하기 어렵다. DCD는 일종의 2단계 디스크 캐쉬로써 작은 용량의 NVRAM을 상위 디스크 캐쉬로 디스크내의 여러 개의 플랫터 중 일부를 하위 디스크 캐쉬로 사용하는 방법으로 트랙 단위의 데이터 전송 시간이 블록 단위의 전송 시간에 비하여 상대적으로 유리한 점을 이용하는데[2], 하위 디스크 캐쉬는 데이터 디스크와 동일한 것으로 디스크 접근이 누적집단으로 발생함에 따라 생긴 유휴시간(idle time)을 이용하여 블록 단위가 아닌 트랙 단위의 비동기적 데이터 전송을 수행한다. DCD는 플랫터 단위 혹은 별개의 물리적인 디스크를 디스크 캐쉬로 사용하므로 디스크 용량의 낭비가 심하다.

디스크 시스템의 용량을 증가 시키기 위하여 선형 기록 밀도(linear recording density)의 개선, 트랙 상호간 거리의 축소, 혹은 디스크 드라이브의 반경 확장 등과 같은 방법이 존재한다. 선형 기록 밀도와 트랙 상호간 거리에 대한 지속적인 발전에 비하여 디스크 반경은 14인치부터 1.9인치까지 다양하지만[8], 디스크의 반경의 증가는 회전 속도의 감소와 전력 소모 및 평균 탐색 시간에 악영향을 미친다. 따라서 현재 대부분의 시스템에서 2.50~5.25 인치의 짧은 반경의 디스크를 사용한다. 최근에는 하드 디스크 제조회사에서 다영역 기록 디스크라는 기법을 사용하여 용량을 증가시킬 수 있는데, 이것은 스피들로부터 거리에 따라 트랙 당 디스크 섹터의 개수를 변화 시키는 것이다. 즉, 예전의 모든 트랙에 동일한 개수의 섹터를 위치시키는 것과는 달리 원주가 긴 외부 트랙에 원주가 짧은 내부 트랙보다 더욱 많은 섹터를 위치시킴으로써 하드 디스

크의 용량을 증가 시킨 것이다[3, 4]. 여기서 동일한 숫자의 섹터를 가진 실린더를 묶어 영역(Zones)이라고 부른다.

3. 가장자리 트랙 캐쉬



(그림 1) 가장자리 트랙 캐쉬 디스크

본 연구에서 제안하는 디스크 시스템 모델은 (그림 1)에 나타난 바와 같이 3단계로 구성된 계층적 구조로써 NVRAM 버퍼, 로그 형식(log format)으로 데이터를 저장하는 디스크 캐쉬, 그리고 기존 디스크와 동일하게 데이터를 저장하는 데이터 디스크로 구성된다. 즉, 소규모 용량의 NVRAM을 1차 캐쉬로 디스크의 가장자리 영역에 위치한 실린더의 모임을 2차 캐쉬로 사용한다. 이와 같은 시스템에서 다수의 소규모 단위 쓰기 동작들을 축적하여 소수의 대규모 단위 쓰기 동작으로의 전환하고 내부 트랙과 가장자리 트랙의 전송률의 차이 등을 이용하여 성능 향상을 도모하고자 한다. 소규모 단위이며 임의적인 쓰기 동작은 먼저 1차 캐쉬인 NVRAM 버퍼에 축적될 것이다. 버퍼에 축적된 쓰기 동작들은 블록 단위의 쓰기 동작들이 모여 트랙 단위의 쓰기 동작이 될 것이다. 최근 연구에 의하면 트랙 단위의 데이터 전송률은 블록 단위의 데이터 전송률에 비하여 8배 이상 빠르다[10]. 전송 단위가 더욱 크면 더욱 빠른 전송률을 성취할 수 있을 것이다. 대부분의 디스크 쓰기 동작은 누적 집단으로 발생하기 때문에

많은 유희 시간이 존재하는데, 이 유희 시간 동안 NVRAM 버퍼에 축적된 모든 데이터가 NVRAM 버퍼와 데이터 디스크의 중간 계층에 위치하는 캐쉬 디스크인 가장자리 영역 트랙에 한 번의 데이터 전송 동작으로 옮겨진다. 그 결과 NVRAM 버퍼는 다시 공백이 되어 추가적인 쓰기 동작 요청에 즉시 유용하게 되므로 프로세서에게는 2차 캐쉬인 가장자리 영역 트랙이 하나의 커다란 NVRAM 버퍼로 보여질 것이다. 만약 NVRAM 버퍼를 빨리 비울 수 없으면 성능에 영향을 미칠 것이다. 왜냐하면 입출력 장치에서의 데이터 전송에서는 작은 용량의 캐쉬만으로는 시간적 국부성(temporal locality)과 공간적 국부성(spatial locality)을 이용할 수 없기 때문이다. 그러나 제안한 시스템에서는 DCD와는 달리 2차 캐쉬로 사용하는 디스크의 가장자리 영역 실린더 공간이 데이터 디스크에 비하여 평균 접근 시간이 좀더 빠르기 때문에 대규모 용량의 쓰기 요청을 빠른 시간에 흡수할 수 있으므로 이와 같은 문제점에 대한 영향이 DCD의 경우와 비교하면 줄어든다. 즉, 최근의 하드 디스크 드라이브에 사용하는 다영역 기록 디스크라는 테크닉에서는 시간 당 읽혀진 섹터의 수, 즉 전송률이 외부 트랙에서 높기 때문이다. 실제 디스크 드라이브에서 외부 트랙의 전송률과 내부 트랙의 전송률을 비교하면 약 1.5~2.0배로부터[3, 11], 최근 제품인 ST-38421A 혹은 ST-36421A 등은 약 2.6배이며 ST-34342A의 경우는 3.0배 이상으로 전송률의 차이는 적지 않으며 조금씩 커지는 경향이다[11].

디스크의 가장자리 영역 트랙을 전체 디스크 공간으로부터 캐싱 공간으로 구별함으로써 탐색 시간을 줄일 수 있는데, 이는 두 가지(디스크 시스템이 하나의 물리적 디스크 드라이브로 된 경우와 여러 개의 물리적 디스크 드라이브로 된 경우)로 나누어 살펴볼 수 있다. 전자는 하나의 디스크에 대한 실린더의 수가 데이터 영역과 캐쉬 영역으로 양분되므로 각각의 접근 요청에 대한 탐색 시간이 짧아진다. 후자는 디스크 시스템이 여러 개의 물리적 디스크로 구성되어 있으므로 데이터 영역과 이에 대응하는 캐쉬 영역을 서로 다른 물리적 디스크에 배치함으로써 이용되지 않는 물리적 디스크의 경우 미리 디스크 헤드를 가장자리 트랙에 대기 상태로 위치 시킴으로써 캐쉬 트랙에 대한 탐색 시간을 최소화할 수 있다. 특히 트랙씩 작업 부하가 많은 경우 동일한 디스크 드라이브에 데이터 영역과 캐쉬 영역이 함께 존재하는 시스템에 비하여 성능 향상을 도모할

수 있다. 하지만 빈번한 트랜잭션(transactions)이 동시에 수행되는 경우는 모든 물리적인 디스크가 동시에 사용 중일 확률이 많아 디스크 헤드를 미리 캐쉬 트랙에 대기시켜 놓을 확률이 줄어들지만 대부분의 디스크의 접근이 누적접근으로 이루어지므로 유희시간에 NVRAM 버퍼의 내용을 가장자리 영역 트랙에 옮기므로 디스크 탐색 시간에 영향을 받은 확률이 적다.

본 연구에서 제안한 방법과 DCD를 비교하면 다음과 같다.

- DCD는 디스크 캐쉬 영역으로 할당할 수 있는 단위가 플랫터 혹은 디스크 드라이브 전체이기 때문에 필요 이상의 디스크 공간을 할당하게 되어 디스크 공간의 낭비가 심각하지만, 본 연구에서 제안한 방법은 디스크 플랫터의 가장자리부터 실린더 단위로 할당할 수 있기 때문에 디스크 공간을 더욱 효율적으로 사용할 수 있다. 일반적으로 각 플랫터에 500~2,500개의 트랙이 존재하며 디스크에 1~20개의 플랫터로 구성되는데 [7], 여기서 각 플랫터에 2,000개의 트랙이 존재하고 하나의 물리적 디스크 드라이브가 10개의 플랫터로 구성된다고 가정하자. DCD에서는 캐쉬 디스크를 위한 용량이 최소한 디스크 공간의 1/10 (10개의 플랫터중 하나) 단위씩 할당해야 하므로 디스크의 공간 낭비가 크다. 특히 별도의 디스크 드라이브를 디스크 캐쉬로 할당한다면 데이터 디스크에 비하여 적은 용량의 디스크를 디스크 캐쉬로 이용한다 하더라도 하나의 디스크 드라이브 공간 전체를 캐쉬 디스크로 할당해야 하기 때문에 디스크 공간의 낭비가 심할 수 밖에 없다. 그러나 본 연구에서 제안한 방법은 디스크 전체 공간의 1/200(1/2,000 트랙×10 플랫터) 단위씩 할당하므로 디스크의 공간에 낭비가 현저히 줄어든다.

- DCD에서 별도의 고성능 디스크를 캐쉬 디스크로 할당하는 경우에는 데이터 디스크보다 평균 접근 시간이 빠르므로 일반적인 캐쉬의 의미를 부여할 수 있지만, 동일한 디스크에서 플랫터 단위로 데이터 디스크 혹은 캐쉬 디스크로 구별하는 경우 모든 플랫터의 특성이 동일하므로 물리적인 속도의 의미로 볼 때 캐쉬라는 의미를 부여하기 곤란하다. 그러나, 본 연구에서 제안한 방법은 DCD처럼 블록 단위가 아닌 트랙 단위의 접근으로 인한 성능 향상을 도모할 뿐만 아니라 디스크의 영역마다 특성이 다른 다영역 기록 디스크의 가장자리를 사용하므로 물리적인 속도가 달라 캐쉬로 불릴 수 있다.

- 디스크 공간을 실린더 단위로 캐쉬와 데이터 영역으로 구분하므로 각 영역에 속한 실질적 실린더 수의 감소로 탐색 시간이 짧아진다. 특히, 하나 이상의 물리적인 디스크 드라이브를 사용하는 시스템인 경우 DCD와는 달리 데이터 영역에 대응하는 캐쉬 영역 트랙을 서로 다른 디스크 드라이브에 위치시킴으로써 대부분의 탐색 시간을 없앨 수 있다.
- 적은 비용으로 높은 성능을 성취하기 위하여 작은 용량의 NVRAM 버퍼의 확장으로 DCD는 디스크의 데이터 디스크와 동일한 일부 플래터를 이용하므로 기하학적으로 디스크의 수평적 이용이라 한다면 가장자리 영역 캐쉬는 디스크의 가장자리 영역만을 이용하므로 디스크의 수직적 이용이라고 볼 수 있다.

4. 모의 실험 및 성능 분석

본 연구에서 제안한 가장자리 영역 트랙 캐쉬의 성능을 평가하기 위하여 미시건 대학에서 C언어로 개발한 디스크 시뮬레이터인 DiskSim[6]을 사용하여 본 연구에서 제안한 캐쉬 기능을 추가하였다. DiskSim는 저장 장치 시스템 구조를 위한 시뮬레이터로써 매우 정확하고 효율적이며 융통성 있게 구성할 수 있어 디스크 시스템을 연구하는 데 널리 사용되고 있다. 모의 실험을 위한 트레이스는 단일 사용자 워크스테이션에서 1주일 동안 모은 디스크 접근의 내용을 수집한 hplajw[12]를 사용했으며, 현재 알려져 있는 소수의 입출력 트레이스로는 다양한 쓰기 트랙픽이나 누적집단의 특성을 나타낼 수 없기 때문에 DiskSim에서 제공하는 매개변수 예를 들면 I/O Trace Time Scale 를 변화함으로써 상호 도착 시간(inter-arrival time)을 여러 가지 비율로 압축하거나 읽기 트랙픽의 일부를 쓰기 트랙픽으로 전환하는 등 작업부하를 다양하게 변형함으로써 제안된 시스템의 성능을 파악하고자 한다.

시스템을 간단히 하기 위하여 다영역 기록 디스크를 여러 개의 영역으로 나누지 않고 2개의 영역(가장자리 영역과 내부 영역)으로만 나누었으며, 가장자리 영역의 밀도를 다양하게 변화 시켜 사용하였다. 또한 본 연구에서 제안하는 가장자리 영역 트랙 캐쉬를 DCD와 비교하기 위하여 두 가지의 DCD 방법 중 논리적 DCD(하나의 디스크 드라이브에서 캐쉬 디스크와 데이터 디스크로 구별하여 사용)의 경우만 사용했으며, 물리적 DCD의 경우는 별개의 디스크 드라이브 요구로 인하여

비용면에서 비효율적이므로 모의 실험에서 제외하였다. 아울러 DCD에서 쓰기 요청에 대한 2가지 방법(버퍼에 쓰기가 완료되면 즉시 CPU에 통보하는 경우와 버퍼와 디스크 캐쉬에 모두 쓰기를 마쳤을 때 CPU에 통보하는 경우)을 사용하는데, 본 연구에서는 후자의 경우에 대한 모의 실험 결과만을 사용하였다. 그 이유는 전자의 경우 대부분의 요청이 버퍼에서 흡수되므로 버퍼의 크기에 성능이 좌우되며 DCD와 본 연구의 결과가 모두 유사한 결과를 나타내기 때문이다. 유휴 시간을 DCD에서와 동일하게 50ms동안 어떤 새로운 디스크 접근 요구가 없을 경우로 가정하였다. 모의실험에서 사용한 디스크 매개변수는 <표 1>에서 나타난 바와 같이 HP C2200A를 기초로 하였다.

<표 1> 디스크 드라이브 매개변수

포맷된 용량	335 MB
트랙 버퍼	없음
실린더 개수	1449
데이터 헤드 수	8
내부영역 트랙당 데이터 섹터 수	113
섹터당 바이트 수	256 B
회전 속도	4002 rpm
컨트롤러 오버헤드 (읽기)	1.1 ms
컨트롤러 오버헤드 (쓰기)	5.1 ms
평균 8 KB 접근	33.6 ms

RAM 버퍼 크기는 512KB이며 가장자리 영역 실린더의 수는 고정시키지 않고 최소한 20MB의 크기를 가질 수 있도록 하였다. 예를 들어 가장자리 영역이 내부영역과 동일한 밀도를 가질 경우 80,000개의 섹터가 필요하므로 90개의 실린더로 정했으며, 내부영역의 3배의 밀도를 가질 경우 30개의 실린더로 정하여 가장자리 영역 캐쉬가 최소한 20MB가 되도록 유지하였다. 제안한 디스크 시스템을 DCD 혹은 기존의 디스크 시스템과 비교하기 위하여 기존 디스크 시스템의 응답시간에 대한 속도 향상(Speedup), S를 사용하였다. 즉,

$$S = \frac{\text{기존 시스템의 평균 응답 시간}}{\text{제안한 시스템 혹은 DCD의 평균 응답 시간}}$$

먼저 영역별 밀도 차이에 따른 모의 실험 결과를 살펴보면 <표 2>와 같다. 현재까지 구현된 다영역 기록 디스크 시스템에서 영역별 최대 비율이 4를 초과하지 않으므로 4까지 만의 결과를 나타내었다. 결과에 의하

면 DCD는 기존 디스크 시스템에 비하여 약 2.48배의 성능향상을 가졌지만, 제한한 시스템은 영역별 밀도 차이에 따라 기존 시스템보다 2.54에서 2.59배의 성능향상을 가진다. 이는 DCD와 마찬가지로 탐색 시간과 회전지연시간의 단축으로 인한 결과이며, 특히 디스크의 가장자리 영역의 높은 밀도로 얻어진 빠른 전송 속도로 인한 추가적인 성능 향상이 있기 때문에 DCD의 성능을 능가한다. 트레이스 hplajw에서 나타난 평균 디스크 접근율이 0.076/초로 매우 낮기 때문에 DCD에 대한 제한한 시스템의 장점에 크게 영향을 미치지 못하여 성능 향상 폭이 그리 크지는 않다. 또한, 영역별 밀도 차이가 증가할수록 속도 향상의 폭은 조금씩 작아지는데 이는 전체 평균 응답시간에 대한 전송 시간이 차지하는 비율이 작아지기 때문이다.

<표 2> 영역별 밀도 차이에 따른 성능 향상

영역별 밀도 비율	2	3	4
DCD	2.59	2.70	2.82
제한한 디스크 시스템	2.54	2.57	2.59

<표 3>은 가장자리 영역의 밀도를 내부영역보다 3배로 고정한 상태로 세 가지의 쓰기 트래픽 비율에 대하여 기존 디스크 시스템에 대한 DCD와 제한한 시스템의 성능 향상을 나타낸다. 쓰기 트래픽 비율이 2인 경우에는 최초의 트레이스에서의 쓰기 요청을 2배로 확장하여 모의 실험한 결과이다. 쓰기 트래픽 비율이 2배인 경우는 누적집단으로 디스크의 접근을 요청할 확률이 더욱 많아지며 누적집단의 규모가 더욱 커져 기존의 디스크 시스템에서는 평균 응답시간이 매우 커지게 된다. 반면에 DCD 혹은 제한한 시스템에서는 누적집단으로 인한 디스크 접근 요청을 디스크 캐쉬를 이용하여 처리하기 때문에 상대적인 성능 향상이 <표 2>의 경우보다 훨씬 커진다. 쓰기 부하가 높을 땐 NVRAM 버퍼 공간이 부족하여 빈번히 버퍼의 내용을 디스크 캐쉬로 옮겨야 하는데 다음과 같은 이유로 제한한 시스템이 DCD의 경우보다 버퍼를 비우는 시간이 빨라 <표 3>에 나타난 것처럼 쓰기 부하가 높을수록 성능 향상의 폭이 커진다. 첫째로 제한한 시스템이 3배의 가장자리 영역 밀도를 가지므로 3배의 빠른 전송물을 가진다. 둘째로 누적집단의 규모가 크므로 DCD의 경우에는 캐쉬 디스크의 가장자리에서 중앙의 모든 트랙에 캐싱된 데이터가 산재하며 제한한 시스템의 가장자리

영역에 제한된 경우에 비하여 탐색 시간이 길어진다.

<표 3> 쓰기 트래픽 비율에 따른 성능 향상

쓰기 트래픽 비율	2	3	4
DCD	2.59	2.70	2.82
제한한 디스크 시스템	2.65	2.86	3.11

<표 4>는 상호 도착 시간을 여러 가지 비율로 압축한 경우 기존 시스템에 대한 DCD와 제한한 시스템의 성능 향상을 나타낸다. <표 3>과 동일하게 <표 4>의 결과도 제한한 시스템에 대하여 가장자리 영역의 밀도를 내부영역보다 3배인 경우를 가정하였다. 이 실험도 쓰기 트래픽 비율에 따른 결과와 유사한 결과를 나타내지만 전반적으로 성능향상에 대한 수치가 낮다. 이는 읽기 및 쓰기에 대한 모든 디스크 요청에 대하여 빈도수는 높아졌지만 읽기 요청에 대한 쓰기 요청의 상대적인 비율은 동일하며 많은 경우가 압축 전 트레이스에서의 두 개의 쓰기 요청 누적 집단의 융합 즉, 앞의 요청을 완전히 처리하기 전에 후의 요청이 시작되어 두 개의 누적집단 쓰기 요청이 하나의 누적집단으로 됨이 많지 않아 <표 3>의 경우에 비하여 성능 향상이 조금 작다.

<표 4> 상호 도착 시간 압축 비율에 따른 성능 향상

상호 도착시간 압축 비율	2	3	4
DCD	2.54	2.66	2.71
제한한 디스크 시스템	2.62	2.79	3.00

끝으로 기존 디스크 시스템에 대한 DCD와 제한한 시스템의 읽기 연산에 대한 성능 향상은 각각 2.41와 2.46 이상으로 기존 디스크 시스템에 비하여 모두 성능이 우수하다. 이와 같은 읽기 동작에 대한 성능 향상은 주로 데이터 디스크에서의 쓰기 트래픽의 감소로 기인한 것이라 할 수 있다. 제한한 시스템이 DCD보다 약간의 성능 향상이 있는 것은 DCD에 비하여 실질적 트랙 수의 감소로 인한 탐색 시간이 짧아지기 때문이다.

<표 2>에서 <표 4>까지의 결과에 나타난 것처럼 제한한 디스크 시스템은 기존의 디스크 시스템에 비하여 최소 2.5배 이상의 성능향상을 나타내며 모든 경우에서 DCD의 경우보다 성능이 우수할 뿐만 아니라 3장에서 언급한 것처럼 디스크 공간을 효율적으로 사용한다. 기존 디스크 시스템에 대한 경비 측면을 살펴보면, 1MB

의 용량에 대하여 디스크는 약 \$0.25 그리고 NVRAM은 약 \$120.00 정도이며[13], 하드웨어의 발전으로 지속적인 가격하락을 고려하면 추가적인 20~100 MB의 디스크 용량과 512KB의 NVRAM은 현재 \$100.00 미만의 투자로 기존 디스크 시스템에 비하여 2배 이상의 성능향상을 도모할 수 있으므로 비용면에서도 효율적인 방법이라 주장할 수 있다.

5. 결 론

디스크 시스템의 쓰기 트래픽에 대한 성능 개선을 위하여 본 연구에서 디스크의 가장자리 영역 트랙을 디스크 캐쉬로 사용하기를 제안하였다. 가장자리 영역 디스크 캐쉬는 다영역 기록 디스크의 전체 공간을 두 개의 영역으로 분할하여 데이터를 위한 디스크 공간으로 할당한 내부 영역을 제외한 나머지 영역으로 기존 디스크 시스템에서 나타난 쓰기 트래픽에 의한 성능 감소를 줄이고자 하는 것이다.

가장자리 영역 디스크 캐쉬에 대한 기본 아이디어는 DCD와 유사하게 3단계로 구성된 계층적 구조로써 RAM 버퍼, 로그 형식으로 데이터를 저장하는 디스크 캐쉬, 그리고 기존 디스크와 동일하게 데이터를 저장하는 데이터 디스크로 구성된다. 이와 같은 구조를 사용함으로써 디스크 접근의 국부성, 블록 단위와 트랙 단위의 전송률의 현저한 차이를 이용할 수 있다. 뿐만 아니라, 다영역 기록 디스크의 영역별 밀도 차이에 의한 데이터 전송률의 이용, 디스크 공간을 두 개의 영역으로 분할함에 따른 탐색 시간의 축소 등을 이용하는 것이다. 아울러 실린더 단위로 캐쉬 공간을 할당하므로 DCD에 비하여 디스크 공간의 낭비가 거의 없다. 모의 실험 결과에 의하면 기존 디스크 시스템에 비하여 hplajw 트레이스를 여러 가지로 변형하여 사용한 결과 2.54에서 3.11배에 이르는 성능 향상을 보였으며 모든 경우에 대하여 추가적인 부담이 별로 없이 DCD에 비하여 성능이 우수하였다. 향후 가장자리 영역 디스크 캐쉬를 RAID 구조에 적용하여 처리율과 응답 시간을 최적화하는 가능성에 대한 연구를 수행할 계획이다.

참 고 문 헌

- [1] D.A. Patterson, G.Gibson, and R.H. Katz, "A Case for Redundant Arrays of Inexpensive Disks (RAID)," Proceedings of GIGMOD, June, 1988.
- [2] Yiming Hu and Qing Yang, "DCD Disk Caching Disk : A New Approach for Boosting I/O Performance," ISCA '96, pp.169-177, May, 1996.
- [3] Rodney Van Meter, "Observing the Effects of Multi-Zone Disks," USENIX Technical Conference, pp.19-30, 1997/.
- [4] C. Ruemmler and J. Wilkes, "An Introduction to Disk Drive Modeling," IEEE Computer, pp.17-36, March, 1994.
- [5] M. Rosenblum and J. Ousterhout, "Design and Implementation of a Log-Structured File System," ACM Transactions on Computer Systems, pp.26-52, Feb. 1992.
- [6] Greg Ganger, Bruce Worthington, and Yale Patt, The DiskSim Simulation Environment, <http://www.ece.cmu.edu/~ganger/disksim/>
- [7] J.L. Hennessy and D.A. Patterson, Computer Architecture A Quantitative Approach(2nd Ed.), Morgan Kaufmann Publishers, 1996.
- [8] D.A. Patterson and J.L. Hennessy, Computer Organization & Design The Hardware/Software Interface(2nd Ed.), Morgan Kaufmann Publishers, 1998.
- [9] M. Baker, S. Asami, etc, "Non-volatile Memory for Fast, Reliable File Systems," Proceedings of the 5th Int'l Symposium on ASPLOS, pp.10-22, Oct. 1992.
- [10] D. Stodolsky, M. Holland, etc., "Parity Logging Disk Arrays," ACM Transactions of Computer Systems, pp.206-235, Aug. 1994.
- [11] Seagate World Wide Web Page, <http://www.seagate.com>.
- [12] John Wilkes, HP Laboratories Disk I/O Traces World Wide Web Page, http://www.hpl.hp.com/personal/John_Wilkes/traces/ajw.1week.srt.Z.
- [13] M. Wu and W. Zwaenepoel, "eNVy, a Non-volatile, Main Memory Storage System," Proceedings of the 6th Int'l Symposium on ASPLOS, pp.86-97, Oct. 1994.



우 종 정

e-mail : jwoo@cs.sungshin.ac.kr
1982년 경북대학교 전자공학과
(학사)
1982년~1988년 산업연구원 책임
연구원
1988년~1993년 Univ. of Texas
at Austin 전기 및 컴퓨터
공학과(박사)

1998년~1999년 Univ. of Texas at Austin 전기·컴퓨터 공학과 교환교수

1993년~현재 성신여자대학교 컴퓨터정보학부 부교수
관심분야 : 병렬처리, 컴퓨터구조, 멀티미디어시스템



홍 춘 표

e-mail : cphong@biho.taegu.ac.kr
1978년 경북대학교 전자공학과
(학사)
1985년~1991년 Georgia Inst. of
Technology 전기·컴퓨터
공학과(박사)

1978년~1985년 국방과학연구소 연구원

1990년~1991년 Atlanta Signal Processors, Inc. Consultant

1992년~현재 대구대학교 컴퓨터정보공학부 부교수

관심분야 : DSP 하드웨어 및 소프트웨어, 컴퓨터구조,
VLSI 신호처리