

# 고속 하이브리드 저장장치의 입출력 성능개선을 위한 헤테로-미러링 기법

변시우\*

<sup>1</sup>안양대학교 디지털미디어공학과

## A Hetero-Mirroring Scheme to Improve I/O Performance of High-Speed Hybrid Storage

Siwoo Byun<sup>1\*</sup>

<sup>1</sup>Dept. of Digital Media, Anyang University

**요약** 플래시 메모리 기반의 SSD는 데스크톱 및 이동형 컴퓨터의 저장 장치를 지원하는 우수한 미디어이다. 플래시 메모리는 비휘발성, 낮은 전력소모, 빠른 데이터 읽기속도 등의 특징으로 데스크톱 및 서버용 데이터베이스의 핵심 저장 요소가 되었다. 하지만, 빠른 읽기 연산에 비하여 상대적으로 느리거나 프리징이 있는 SSD의 쓰기 연산 특성을 고려하여 HDD와 RAID에 기반을 둔 기존의 전통적인 스토리지 관리 기법을 개선할 필요가 있다. 이를 위하여, 본 논문은 전통적인 HDD 미러링에 기반을 둔 헤테로-미러링이라는 새로운 저장 장치 관리 기법을 제안한다. 헤테로-미러링 기반 기법은 SSD에서 발생 가능한 프리징 현상을 피하기 위한 쓰기-부하 밸런싱과 쓰기 지연 연산을 통하여 RAID-1 처리 성능을 개선한다. 테스트 결과 제안한 기법은 쓰기 연산 부하와 프리징 부하를 크게 감소시키고, 기존의 SSD-RAID-1 기법의 성능을 18%, 응답 시간을 38% 개선함을 확인하였다.

**Abstract** A flash-memory-based SSDs(Solid State Disks) are one of the best media to support portable and desktop computers' storage devices. Their features include non-volatility, low power consumption, and fast access time for read operations, which are sufficient to present flash memories as major database storage components for desktop and server computers.

However, we need to improve traditional storage management schemes based on HDD(Hard Disk Drive) and RAID(Redundant array of independent disks) due to the relatively slow or freezing characteristics of write operations of SSDs, as compared to fast read operations. In order to achieve this goal, we propose a new storage management scheme called Hetero-Mirroring based on traditional HDD mirroring scheme. Hetero-Mirroring-based scheme improves RAID-1 operation performance by balancing write-workloads and delaying write operations to avoid SSD freezing. Our test results show that our scheme significantly reduces the write operation overheads and freezing overheads, and improves the performance of traditional SSD-RAID-1 scheme by 18 percent, and the response time of the scheme by 38 percent.

**Key Words** : SSD, HDD, RAID, Flash Memory Storages, Mirroring

### 1. 서론

하드 디스크(HDD: Hard Disk Drive)는 지난 30년 전부터 대부분의 데이터 서버에서 사용되는 절대적인 저장

미디어이다. 그러나 최근 중대형 데이터 서버에서의 저장 시스템으로서 Solid State Drive(SSD)[1,2]가 주목 받기 시작하였다.

이미 1990년대에 등장한 SSD는 하드 디스크에 비하

이 논문은 2010년 한국연구재단의 지원을 받아 수행된 기초연구사업 연구임(NRF-2010-0016705)

\*교신저자 : 변시우(swbyun@anyang.ac.kr)

접수일 10년 01월 11일

수정일 (1차 10년 05월 11일, 2차 10년 09월 06일)

게재확정일 10년 12월 17일

여 많은 장점에도 불구하고, 그동안 가격 경쟁력이 매우 부족하여, 특수한 분야(군사, 항공, 우주, 선박 등)에서만 주로 사용되었다. 그러나 최근에는 플래시 메모리[3]의 급속한 가격 하락에 힘입어, 서버나 휴대용 노트북등 고급 제품을 중심으로 기업 및 일반 소비자 시장으로 확대되고 있다.

특히 기존의 휴대용 USB 저장장치의 용도를 뛰어넘어 급속하게 하드디스크의 새로운 대체품으로 급부상하고 있다. 상대적으로 비교하자면, 용량과 가격측면에서는 하드 디스크가 아직은 훨씬 더 우세하나, 접근 성능 및 저전력성 측면에서는 SSD가 절대적으로 우세하다[4].

또한, 과거부터 하드 디스크 저장 시스템 업계에서는 저장 성능과 응답 성능의 문제해결을 위하여 여러 개의 디스크를 연동한 RAID 기술을 사용하여 왔었다. 마찬가지로 SSD도 성능과 안정성을 모두 개선하기 위해서는 이 RAID 기술의 활용이 필요하다. 이러한 하드 디스크 기반의 RAID 기술을 SSD에 그대로 적용해도 어느 정도의 성능 개선과 응답 성능 향상을 얻을 수 있다. 그러나 읽기 및 쓰기 성능이 대칭이 되는 하드디스크에 비하여, 쓰기가 매우 느려서 비대칭인 SSD를 하드디스크와 동일한 기법으로 저장 및 검색하면 심각한 I/O 문제가 발생할 수 있다.

본 연구에서는 이러한 고속 SSD에 대하여, 기존의 저장 장치로서의 약점은 상호 보완하여, 입출력 성능을 극대화할 수 있는 융합형 데이터 저장 기술인 헤테로-미러링(Hetero-Mirroring) 기법을 제안한다. 본 기법은 SSD의 신뢰성 측면에서 가장 중요한 레벨인 RAID-1에 대한 응답시간 및 처리 성능을 개선한다.

## 2. 기존 저장 장치 관련 기술

### 2.1 기존 HDD의 특성 및 최근 기술 동향

하드디스크[5]는 원판 회전형으로 분당 5400 RPM의 회전 속도로 기계적인 모터에 전력을 공급하여 스피들을 회전하여 데이터를 저장하는 방식이다. 하드디스크는 주로 자성체로 코팅된 알루미늄 원판에 자료를 저장할 수 있게 만든 데이터 저장장치로서 데이터를 저장하는 원반 모양의 금속판 플래터(Platter), 플래터 위를 움직이며 데이터를 기록, 수정, 삭제하는 헤드(Head), 플래터를 회전시키는 스피들 모터(Spindle Motor) 등으로 구성되어 있다.

하드디스크는 크기와 용도에 따라 크게 2.5인치 이하 모바일용 하드디스크 제품, 2.5인치 모바일용 제품, 3.5인치 데스크톱용 HDD제품, 2.5/3.5인치 기업용 하드디스크

제품시장으로 나누어진다. 이러한 하드 디스크는 기계적인 특성상 헤드의 이동 시간, 디스크 회전 지연시간 등의 많은 오버헤드가 발생하므로, 향후 플래시 메모리 저장장치(SSD)로 대체될 수밖에 없다[9].

### 2.2 최신 SSD의 특성 및 최근 기술 동향

하드 디스크 저장 용량이 급속도로 증가하고는 있을 지라도, 디스크의 기계적인 특성(헤드의 이동 시간, 디스크 회전 시간)에 의한 지연시간은 매년 15%정도의 미약한 수준으로 개선되고 있다. 반면에 플래시 메모리 처리 속도는 매년 50% 정도의 빠른 속도로 발전되고 있다.

이러한 기계적 특성에 따른 속도 차이에 의하여 하드 디스크는 점차로 한계 성능에 도달하게 된다. 또한, 이에 따른 전력 소모가 평상시에도 전체 시스템 전력의 20% 정도로 매우 큰 편이며, 사용 대기 중에도 전력이 계속 소모된다. 이에 대한 대안으로 최근에 주목받고 있는 저장장치가 바로 그림 1과 같이 플래시 메모리를 장착한 SSD이다.



【그림 1】 32GB 플래시 메모리가 내장된 SSD

또한 2GB~8GB 정도의 휴대용 USB에 사용되던 플래시 메모리가, 이제는 저용량 백업 장치의 용도를 뛰어넘어, 기존의 노트북, 넷북, UMPC나 운영 서버에 탑재된 데이터 저장장치(HDD)의 새로운 대체품(현재,256GB급)으로 급부상하고 있다.

현재 SSD의 가격과 기술적 측면을 볼 때, 현재의 모든 데이터베이스 스토리지에 적용하기는 어렵다. 특히, 속도가 중요하지 않은 배치 작업이나, 백업은 기존의 하드디스크와 RAID기술로 충분히 지원가능하다. 하지만, 높은 데이터베이스 성능과 랜덤 읽기 및 쓰기가 중요하다면 SSD가 좋으며, 특히, 높은 트랜잭션 처리, 순간적인 피크로드가 걸리는 경우, 실시간 데이터베이스 등의 분야에서 효과적이다 [6,7,9].

### 2.3 HDD 기반의 RAID 저장 시스템

기존의 하드디스크를 기반으로 하는 RAID (redundant array of independent)[8,10]저장 시스템에는 10가지 이상

의 변형된 형태가 있다. 그 중 잘 사용되는 주요 레벨은 다음과 같으며, SSD의 특징을 고려한 개선이 필요하다.

- RAID-0: 스트라이프를 가지고는 있지만 데이터를 중복해서 기록하지 않는다. 따라서 가장 높은 읽기 및 쓰기 성능을 기대할 수 있지만, 고장대비 능력이 전혀 없으므로 이 방식은 진정한 RAID라고 하기 어렵다.

- RAID-1: 흔히 디스크 미러링이라고도 하는데, 중복 저장된 데이터를 가진 적어도 두 개의 드라이브로 구성된다. 각 드라이브를 동시에 읽을 수 있으므로 읽기 성능은 향상된다. 쓰기 성능은 단일 디스크 드라이브의 경우와 비슷하다. RAID-1은 다중 사용자 시스템에서 최고의 성능과 고장대비 능력을 발휘한다.

- RAID-3: 이 형식은 스트라이프를 사용하며, 패리티 정보를 저장하기 위해 별도의 드라이브 한 개를 쓴다. 내장된 ECC 정보가 에러를 감지하는데 사용된다. 데이터 복구는 다른 드라이브에 기록된 정보의 XOR를 계산하여 수행된다. 입출력 작업이 동시에 모든 드라이브에 대해 이루어지므로, RAID-3은 입출력을 겹치게 할 수 없다. 이런 이유로 대형 레코드가 많이 사용되는 업무에서 단일 사용자시스템에 적합하다.

- RAID-5: 회전식 패리티 어레이를 포함한다. 그러므로 모든 읽기/쓰기 동작은 중첩될 수 있다. RAID-5는 패리티 정보를 저장하지만 데이터를 중복저장하지는 않는다. RAID-5는 보통 3~5개의 디스크를 어레이를 요구한다.

### 3. SSD의 성능 개선을 위한 헤테로-미러링 저장 기법

#### 3.1 SSD에 기존 RAID 기술의 적용

SSD 저장 시스템에 성능개선을 목적으로 기존의 RAID 기술을 적용시에는, 다음과 같이 SSD의 특징들을 고려하여 설계하면, 효율과 안정성을 향상시킬 수 있다.

- RAID-0의 적용: SSD에서는 read에 비하여 write가 매우 느리므로, 이를 병렬로 분산 write를 수행하여야만, 전체적인 I/O의 밸런스가 이루어진다. 특히, write에 수반되는 매우 느린 소거 연산이 수반될 경우, 병목현상이 자주 발생한다. 따라서 SSD에서는 속도개선상 필요하다.

- RAID-1의 적용: SSD에서는 특히 더 필요하다. SSD는 저장 수명이 100만 번으로 제한되어 있으며, 제품의 보증수명도 10년으로 되어 있으므로, 장기간 사용할 경우의 신뢰성을 보장할 수 없다. 따라서 미러링이 필수적이며, 기존의 RAID 기법에서 SSD에 적합하도록 효율적인 성능 개선이 필요하다. 가장 중요한 적용 기술이다.

- RAID-3의 적용: 디스크가 파괴되면, 복구시 상당한 read/write 연산이 수행되게 된다. 이 경우 패리티 디스크에서는 상당한 read가 발생하고, 복구 디스크에서는 상당한 write가 발생하게 된다. 결국, 느린 write 속도로 인하여 전체 복구가 더욱 느리게 진행된다. 따라서 패리티 디스크를 별도로 두는 것은 SSD에서는 비효율적이다. 패리티를 여러 SSD로 분산하면, 복구 속도를 높일 수 있다.

- RAID-5의 적용: 패리티가 분산되어, 복구시 read 연산이 분산되어 로드 밸런싱이 잘 된다. 다만, 읽기 연산은 SSD에서는 고속으로 수행되므로, 읽기 연산이 많은 데이터 패턴의 경우나 파일 크기가 작은 경우는, 오히려 분산된 다중 read가 더 느리고, 비효율적이다.

#### 3.2 SSD의 불안정한 I/O 특성의 문제점

최근 급부상중인 SSD는 첨단 플래시 메모리를 사용하기 때문에, 매우 빠르다고 알려져 있지만, 실제 사용해보면 아직은 다음 문제점이 존재한다.

[표 1] SSD의 불안정한 I/O 특성

No	불안정한 I/O 특성	발생 현상
1	읽기/쓰기 속도의 비대칭성	읽기에 비하여 쓰기 속도가 너무 느림
2	자료 누적에 의한 성능 열화	파일시스템에 일정용량을 넘어서면 저장 성능이 떨어짐
3	저장연산 집중시의 프리징 현상	저장부하가 급격히 높아지면, 시스템 홀딩이 수 초간 발생

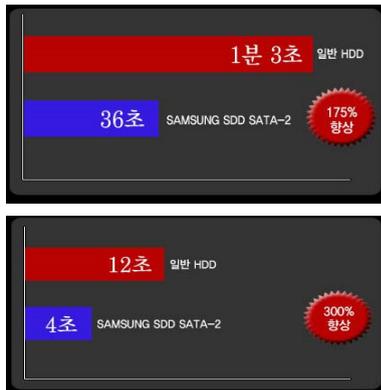
#### 1) 읽기/쓰기 속도의 비대칭성 문제

그림 2의 그래프를 보면, 윈도우 부팅시 1분3초인 일반 하드 디스크보다 175% 속도가 향상되어 36초 만에 부팅이 완료되고, PDF 문서 읽기를 수행하면 무려 3배 가까운 속도의 향상이 나타난다고 한다. 실제 사용자가 직접 테스트해보아도 상당히 빠른 접근 속도가 나타난다.

이렇게 높은 접근 성능의 이유는 SSD는 HDD의 기계적 동작부분이 전혀 없기 때문이다. 그러나 대부분의 SSD의 실제 체감속도는 아주 고가의 SSD를 제외하면, 기대만큼 빠르지는 않다. 물론, 읽기 속도는 매우 개선되어 충분한 속도감을 느낄 수 있다. 즉, 이것은 읽기 위주의 부팅이나, 문서 로딩에 국한되는 것이다.

이러한 느린 체감 성능의 가장 큰 이유는, 일반 사용자 환경에서는 읽기만 하는 것이 아니고, 쓰기 연산도 병행되기 때문이다. 따라서 쓰기 비중이 크다면 이전에 장착된 하드 디스크와 비슷하거나 오히려 낮은 성능이 나타난

다. 즉, 사용자의 작업은 읽기와 쓰기 연산이 임의로 혼합되어 있으므로, 전체 성능에서는 불만족스럽게 나타난다.



[그림 2] 64GB-SSD 성능 자료 (2009, 모델: MMDOE28G5MPP-OVA)

## 2) 자료 누적에 의한 성능 열화

SSD는 데이터가 별로 없는 제품구입 초기에는 저장 속도가 빠르나, 이후에 점점 더 데이터가 많이 쌓여갈수록 더 느려지는 성능 열화현상이 나타난다. 이는 SSD에 데이터가 일정용량이상 누적되게 되면, 느린 쓰기 특성과 관리 모듈의 overhead가 급속히 증가하기 때문에 저장 성

능이 서서히 떨어진다. 이러한 열화현상은 SSD뿐 아니라 USB용 플래시 메모리에서도 나타난다. 실험결과 전체 SSD 영역에서 60%를 채우면 가비지콜렉션(Garbage Collection) 오버헤드가 10%증가하고, 90%정도를 채우면 오버헤드가 50%가 증가한다고 한다[11].

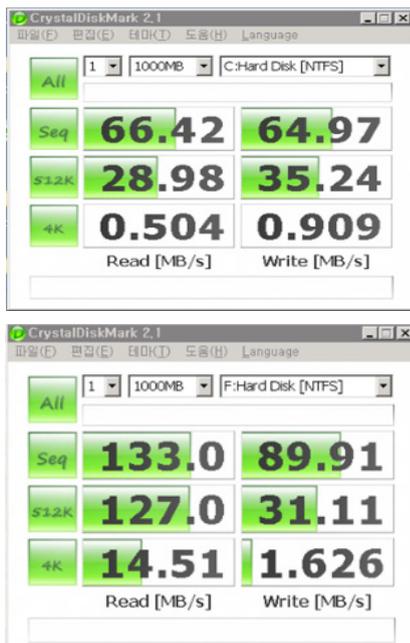
이러한 읽기와 쓰기 속도의 체감 성능의 차이를 미리 확인하기 위하여, 일반적으로 많이 사용하는 SSD와 HDD를 비교 실험하였다[9]. 그 결과 그림 3과 같이 SSD는 읽기 속도는 순차모드에서 약 2배, 랜덤 모드에서 4배 이상의 좋은 성과를 보인 반면, 쓰기 속도에서는 일반 파일 크기인 512K 랜덤 모드에서 오히려 하드 디스크 (35MB/s)보다 13%정도 더 나쁜 결과(31MB/s)를 보이는 문제가 발생하였다.

## 3) 저장연산 집중시의 프리징 현상

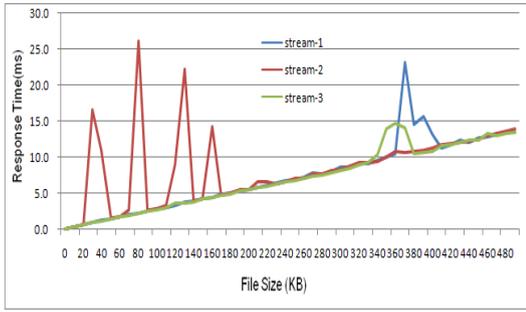
SSD는 사용자의 저장 연산이 집중되는 경우, 평소 속도보다 수 배 이상의 느린 속도로 저장되는 현상이 발생한다. 물론 가끔씩 발생하지만, 하드 디스크에 비하여 더 빈번히 발생하고, 홀딩 시간도 훨씬 더 길다. 물론, 이러한 홀딩 현상은 읽기 연산에서는 거의 발생하지 않는다. 그러나 큰 소프트웨어를 설치하거나 큰 콘텐츠를 복사하는 도중에 매우 심한 경우 수 초 씩 정지(흔히 프리징 'freezing'이라고 함) 되는데, 이러한 체험 사례는 인터넷에서도 여러 번 보고되어 있다.

본 연구에서도 이러한 프리징 현상을 확인하기 위하여 성능 실험을 하였다. 이러한 불안정한 응답성은, 그림 4에 의 결과 그래프에서 확인 가능하다. 이 그래프는 10KB~500KB 까지 10KB 단위로 50개의 저장 스트림을 발생시켜서 응답시간(Response Time)을 측정된 결과이다. 파일 크기를 이 범위로 설정한 이유는, 일반적인 시스템 환경에서는 하루 동안 수정이나 읽기가 수행된 파일의 96%이상이 500KB 미만이기 때문이다. 실험결과 다수의 응답 지연 피크가 발생하는데, 스트림-2의 경우처럼 한번 홀딩이 되면 연이어서 계속 프리징이 발생하기도 한다.

이러한 프리징의 이유는, 쓰기 연산의 집중으로 인하여, SRAM 등의 버퍼가 포화되어, 더 이상 버퍼링하지 못하고, 버퍼에서 밀려 있던 플래시 메모리의 소거 작업과 이에 수반되는 마모 평준화(웨어레벨링) 연산, 복사 오버헤드, 느린 구동 모듈 등의 복합적인 요인 때문이다. 또한, SSD에 데이터가 별로 없는 제품구입 초기 보다는, 데이터가 점점 더 쌓여갈수록 저장 시스템 전체의 관리 오버헤드와 가비지-콜렉션 오버헤드 등이 증가하여, 결국 프리징현상이 더 빈번해진다.



[그림 3] HDD와 SSD의 성능 테스트 결과 (상:HDD, 하:SSD, CrystalDiskMark2.1)



[그림 4] 자체 테스트 결과 (SSD Write Stream)

### 3.3 헤테로-미러링 기법의 제안

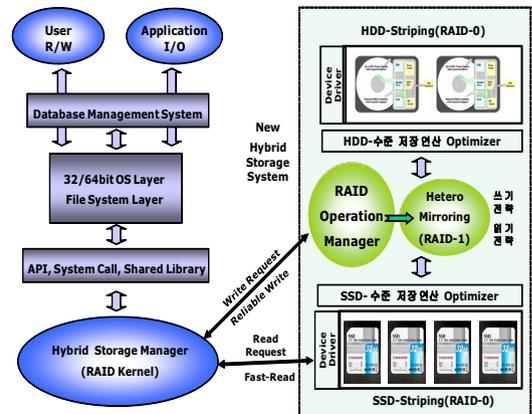
일반적으로 기존 RAID 기술을 SSD에 그대로 적용해도 어느 정도의 효과를 얻을 수 있다. 그러나 단일 매체로 구성된 RAID는 아무리 그 구성 방법을 바꾸어도, 그 단일 매체의 자체의 본질적인 약점을 제대로 극복하기는 어렵다. 예를 들어, HDD만으로 혁신적인 RAID기법을 개발하더라도, HDD의 구성만 달라질 뿐, 본질적인 기계적인 약점은 그 속에 그대로 내재되어 있기 때문이다. SSD의 경우도 마찬가지이다. 따라서 반드시 서로의 약점을 보완해줄 성질이 다른 이질형(Heterogeneous)-RAID 기술이 필요하다. 즉, 한 매체의 약점이 다른 매체에서는 강점인 반대 특성을 활용하여, 그 약점을 흡수해야 한다.

본 연구에서는 이러한 관점에서 기존의 RAID-1 기술을 개선하고, 고속 검색용 SSD의 약점을 보완하기 위하여, 대용량 HDD를 융합하여, 전체적인 시너지 효과를 낼 수 있는 새로운 헤테로 미러링(Hetero Mirroring) 기술을 제안 한다. 본 논문에서는 미러링된 스토리지 제어에서 읽기 연산은 어느 한쪽의 빠른 드라이브를 통하여 수행된다. SSD의 읽기 속도는 매우 빠르며 검색에서는 병목 현상 문제가 없으므로 간략히 서술하며, 본 논문의 이슈인 저장 즉 쓰기 연산에 대하여 설명하고자 한다.

헤테로 미러링 기법을 활용하면 전술한 SSD 프리징 현상을 감소시킬 수 있다. 즉, 일반 부하시의 SSD의 응답 시간이 피크치로 급상승할 때에도, HDD의 응답시간은 정상적인 경우가 많다. 반대로 헤드 이동 등으로 HDD도 피크치로 응답지연이 길어질 수 있는데, 이때는 SSD의 응답시간은 정상적인 경우가 많다. 즉, 동일한 저장 연산을 수행하더라도 두 매체가 모두 비정상적으로 오랜 시간이 지연된다거나, 프리징에 걸리는 경우는 없었다. 이는 4장의 성능 실험을 통해서도 알 수 있었다. 이러한 상호 보완적인 특징을 잘 활용하여, 특성이 상반된 이질형 두 매체(SSD와 HDD)를 미러링하면, 어느 한 쪽 드라이브의 급격한 응답지연을 충분히 흡수할 수 있다. 기존의 하드 디스크 기반 RAID-1(미러링) 기술은 안정성과 처리

성능에 일대 혁신을 이룬 상용화된 매우 중요한 기술이다. 즉, 원래 기존 미러링의 목적이 동일한 매체를 동시에 복제하여 안정성을 높이는 것이라면, 헤테로 미러링은 좀 더 나아가서, 프리징의 충격까지 흡수시켜서 최종 성능까지 높이는 새로운 이질형 RAID 기술의 핵심이다.

또한, 전술한 데로, SSD를 빠른 입출력을 위하여 기존의 RAID에 그대로 적용하게 되면, 많은 성능과 안정성에 심각한 문제가 발생하게 된다. 그 주요 원인은 read/write 처리 시간이 비슷한 하드 디스크에 비하여, 플래시 메모리 SSD는 read/write 처리 시간의 매우 차이가 나는 극심한 비대칭적 구조를 가지고 있으므로, I/O의 밸런스가 제대로 맞지 않기 때문이다. 어떤 시스템도 전체적인 밸런스가 맞아야만 제대로 성능을 발휘할 수 있는 것과 같은 이치이다. 이러한 수배에 달하는 극심한 읽기/쓰기의 밸런스의 차이는 여러 개의 드라이브를 길게 일렬 연결(RAID-0,스트라이핑)하여 어느 정도 극복할 수 있다. 즉, 일렬 배열 드라이브에 비례하여 성능은 점차 높아질 수 있다. 이러한 이질형 상호-보완 미러링과 더불어, 밸런스가 맞춰져서 성능을 위하여 설계한 본 연구의 시스템 구성도는 그림 5와 같다.



[그림 5] Hetero-RAID 기반 하이브리드 스토리지

먼저, 사용자나 응용프로그램의 입출력 요청이 데이터 베이스 시스템에 접수되면, 중간의 파일 시스템 레이어를 거쳐서 하부의 하이브리드 저장 저장 관리자로 전달된다. 이 하이브리드 저장 관리자는 헤테로-레이드 저장 기능을 주관하여, 물리적 레이드 장치에 미러링된 입출력 요청으로 변환하여 전달하고, 추후 처리된 결과를 리턴한다.

사용자의 Read연산은 이를 통과한 후, 프리징 예방을 위하여 헤테로-미러링된 이중화 Write 연산으로 변환된다.

또한, 사용자의 Write 연산은 성능 향상을 위하여 RAID-0모드로 스트라이핑되어서 고속화된 Read 연산으

로 변환된다. 따라서, 하이브리드 스토리지의 RAID 기술과 더불어 프리징 장애를 최소화하고, I/O성능을 가속시키려면, 프리징을 빨리 파악하고 즉각 부하를 밸런싱할 수 있는 효율적인 연산관리자(Operation Manager)가 필요하며, 표 2와 같은 항목이 설계시 충분히 고려되어야한다.

[표 2] 연산관리자의 설계 목적 및 고려사항

설계 목적	고려 사항
프리징 최소화	- 저장 연산 과부하 판단 - 부하 조절
성능 향상 수명 연장	- 지연 저장 - 일괄 순차 저장

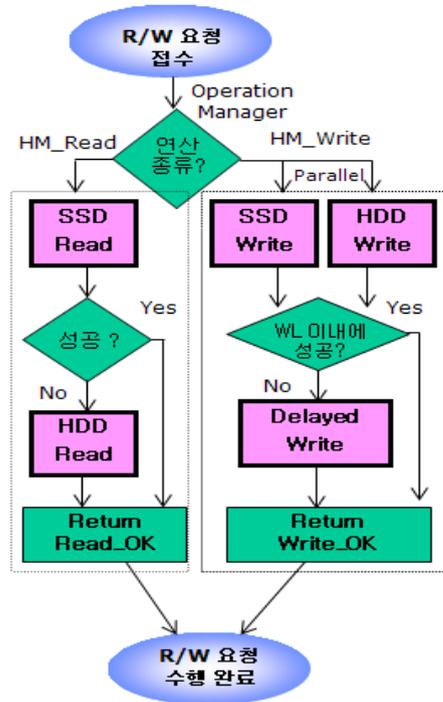
1) 프리징 최소화 측면

주로 SSD에서 많은 양을 저장하거나, 소규모 데이터를 계속하여 랜덤 쓰기를 하면, 프리징 현상이 발생한다. 이러한 과밀 적체를 판단한 후 바로 부하를 조절하여 피해갈 수 있도록 설계하여야 한다. 예를 들어 적체가 임계점을 초과하면서 응답시간이 급속히 느려지기 시작하는 프리징이 시작된다면, HDD에 우선적으로 저장하고, SSD의 저장을 지연하여 일단 과부하를 줄이는 것이 효율적이다.

2) 성능 향상 측면

먼저, 안정된 성능을 위하여 과부하 상태의 저장 연산은 지연시키는 편이 성능 향상에 유리하다. 또한 이렇게 지연된 SSD 큐잉으로 쓰기 연산이 큐에 모아지므로, 추후 블록클리닝을 위한 쓰기 횟수가 줄어들어서 결과적으로 수명 연장의 효과도 볼 수 있다. 즉, 랜덤 쓰기가 일정 시간 모여져서 추후에 순차쓰기로 유도된다. 당연히 플래시 메모리에 기반하는 SSD도 순차쓰기가 빠르며, 특히 작은 크기의 파일일 경우 더 효과적이다. 이는 특히 중대형 데이터베이스 서버나 멀티태스킹 작업이 많은 사용자 환경에서도 예측가능하고 안정된 응답시간을 더 높은 수준으로 보장해 줄 수 있다. 그림 6은 제안된 Hetero-Mirroring기법의 수행 흐름도이며, 그 알고리즘의 설명은 다음과 같다.

- ① 사용자 프로그램이 하이브리드 스토리지에 대한 읽기 또는 쓰기 연산(operation)을 요청한다.
- ② 사용자 입출력 요청이 접수되면, 요청된 연산은 헤테로-미러링 연산으로 변환된다. 이를 관리하는 Operation Manager는 입출력 요청연산을 자신의 연산-큐의 마지막에 삽입한다. 추후에 지금 요청한 연산이 완료되면 최종적으로 사용자 프로그램에 결과 값을 리턴 한다.



[그림 6] 헤테로-미러링 기법의 수행 흐름도

- ③ 연산-큐에서 한 연산씩 차례로 꺼낸 후 연산의 종류를 분류한다. 읽기 연산이면, ④번의HM\_READ을 수행하며, 쓰기 연산이면 ⑤번의 HM\_WRITE를 수행한다.
- ④ HM\_READ(Hetero-Mirroring\_READ):
  - 1) 먼저 읽기 속도가 빠른 SSD에 읽기연산을 포워드한다. 기존의 미러링의 경우 두 디스크에 반반씩 나누어서 읽기가 가능하지만, SSD가 HDD보다 3배 이상 읽기속도가 빠르므로, SSD가 전체를 읽는 것이 유리하며, 상대적으로 전력절감 효과도 크다.
  - 2) SSD로부터 읽기 작업이 성공하면, 판독한 값을 Operation Manager에게 리턴 한다.
  - 3) 만일 SSD로부터 읽기 오류가 발생하면, HDD에 동일하게 읽기 요청을 한 후, 판독한 값을 Operation Manager에게 리턴 한다.
- ⑤ HM\_WRITE(Hetero-Mirroring\_WRITE):
  - 1) 미러링은 SSD와 HDD 모두에 동일한 데이터를 저장해야하므로, 두 드라이브 모두에게 저장 명령을 병렬로 수행한 후 기다린다.
  - 2) 두 드라이브 중에 하나라도 저장 연산의 수행이 정상적으로 완료되었다는 이벤트가 나오면, 처리 시간(wt)을 기록하고, Operation Manager에게 저장 완료 메시지를 리턴 한다.

- 3) 나머지 한 드라이브의 완료 이벤트를 설정된 시간 WL(wait\_limit: wt\*1.3로 기본 값으로 함)만큼 더 기다린다. wait\_limit은 먼저 완료한 저장 속도에 대한 상대적인 시간 값으로, 클수록 동기화 간격은 짧아지나, 응답속도는 더욱더 느려진다. 이를 설정하지 않으면, 홀딩현상이 발생할 경우 저장 연산에 대한 응답 대기 시간이 심각하게 길어진다. 실험값으로 wt\*1.3으로 설정하면, 확률상 전체연산중 80% 이상이 프리징 징후인 WL시간이상은 홀딩하지 않게 된다. 이는 자체 튜닝과정에서 통계적으로 찾은 설정치이다.
  - 4) 만일 해당 드라이브가 WL값을 초과하면, 이 드라이브에는 일시적으로는 더 이상 저장 연산을 큐잉하지 않고, 대신 지연 저장을 위한 Delay\_Queue에 쌓이게 된다.
- ⑥ HM\_DELAY\_WRITE(Hetero-Mirroring\_DELAY\_WRITE):
- 1) 해당 드라이브가 비지 상태이면 계속 대기하고, 아니면 Delay\_Queue에 큐의 맨 앞에 있는 연산을 선택하여, 해당 드라이브에서 지연된 쓰기 연산을 가져 온다.
  - 2) 해당 드라이브에 지연된 쓰기 연산을 수행시킨다.
  - 3) 쓰기 연산 작업이 성공하면, 완료 메시지를 Operation Manager에 리턴 한다.
  - 4) 큐에서 지연된 연산이 있으면, 큐의 내용이 없을 때까지 1)번 이하를 반복한다. 결과적으로, 지연 연산 큐에 쓰기 연산들이 모이므로, 유희시 한꺼번에 연속해서 저장할 수 있다. 따라서 SSD의 경우 모아쓰기를 유도하는 부수적인 효과가 발생하여, SSD이 저장효율과 전력절감 효과가 발생한다. HDD의 경우에도 랜덤한 임의쓰기 연산이 모아지므로 한꺼번에 순차쓰기할 효과가 있다.

## 4. 성능 평가

본 연구에서 제안된 헤테로 미러링 기법의 저장 성능을 검증하기 위하여 실제 환경에서 작업 부하를 주고, 그 실험 결과를 분석해 보았다. 본 실험 수행시 제안한 헤테로 미러링기법(Hete-Mr)과 비교된 저장 관리 기법은 SSD를 1개만 사용하는 단일디스크 저장 기법(SSD-1)과, 가장 보편적이고 안정적인 SSD RAID 구성인 미러링 기법(SSD-Mr)이다.

### 4.1 실험 환경

#### 1) 실험 시스템 구축

실험이 수행된 하드웨어 환경은 최대한 일반 사용자 환경에 가깝게 구성하였다. 아래 표 3과 같이 보편적인 성능의 DELL 펜티엄4 CPU와 메인 메모리 2G, 운영체제는 현재 가장 많이 사용하고 있는 윈도우 XP를 사용하였다. 언어는 비주얼 C++를 사용하였으며, 부수적으로 작업 부하 생성모듈과 물리적인 실험후 통계 자료 수집의 편의상 CSIM[12] 시뮬레이션 라이브러리를 사용하였다. 성능 측정 부분은 모두 SATA로 연결된 SSD와 HDD를 통하여 검출되었다.

[표 3] 실험 환경의 구성

항 목	설 정	항 목	설 정
System	DELL OPTIPLEX 745	CPU	INTEL 2.1 MHz
OS	Windows XP	TOOL	VISUAL C++ Crystal Mark 2.1
HDD	SATA 250G Read:100MB/s Write:100MB/s	SSD	SATA 32G*2 Read: 150MB/s Write: 100MB/s
Disk-Cache	Disable	I/O	Random (10kB ~500KB)

#### 2) 저장 데이터의 구성

본 실험에 앞서서, 실험용 샘플이 되는 데이터가 필요 한데, 실제 사용자의 윈도우 환경에서 사용 중인 모든 파일을 검색한 후, 사용 빈도가 가장 높은 순으로 나열해 보면, 96%이상이 500KB 미만의 작은 크기의 스몰 파일들이다. 이를 근거로 하여, 10KB부터 500KB 사이의 10KB 단위로의 랜덤 파일을 생성하여 미리 저장하여 두었다.

#### 3) 워크로드 모델 및 평가 지표

[표 4] 성능 평가 지표

지 표	설 명	단 위
operation throughput	초당 수행 완료된 저장 연산의 수	operations sec.
response time	저장 연산이 발생한 후 수행완료까지의 소요 시간	millisec

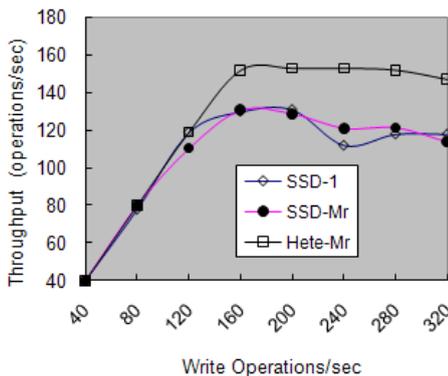
실험을 위한 저장 워크로드 생성 및 결과 분석 모델이 필요한데, 이는 CSIM에서 제공되는 폐쇄형 큐잉 모델(Closed Queuing Model)과 통계 라이브러리를 사용하였다. 이 모듈들을 통하여 초당 일정한 수의 쓰기 오퍼레이션을 생성하고, 이와 연관된 처리 시간과 성능을 측정하

면 된다. 본 실험의 주요 성능 평가 지표는 표 4와 같다.

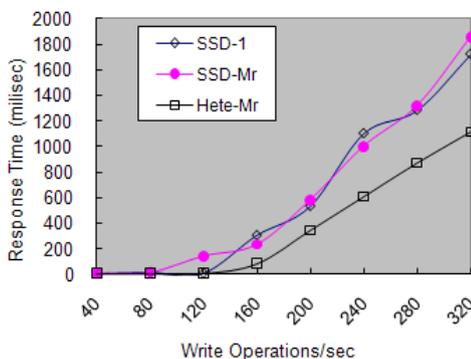
본 실험에서 생성되는 저장 오퍼레이션의 수는 초당 40개에서부터 320개까지 40개 단위로 8단계로 변화시켜 보았으며, 이는 저장 시스템에 가해지는 작업 부하를 의미한다.

#### 4.2 실험 결과 및 분석

이 실험은 기본적으로 하드 디스크나 SSD를 탑재한 일반적인 사용자 시스템에서 부과한 저장 워크로드, 즉 초당 발생된 저장 연산의 수가 시스템 성능에 어떤 영향을 미치는지를 분석하기 위함이다. 그림 7은 초당 발생된 저장 연산의 수의 증가에 따른 저장 연산의 처리치를 표시한 그래프이다. 발생된 초당 저장 연산의 수가 늘어날수록 점차로 연산의 처리 결과치가 증가함을 알 수 있다. 또한, 전반적인 저장 연산의 처리 성능을 측정할 결과, 제안기법인 Hete-Mr가 SSD-1과 SSD-Mr보다 높게 나타났다.



[그림 7] 초당 저장 연산 처리치의 비교



[그림 8] 저장 연산의 응답 시간의 비교

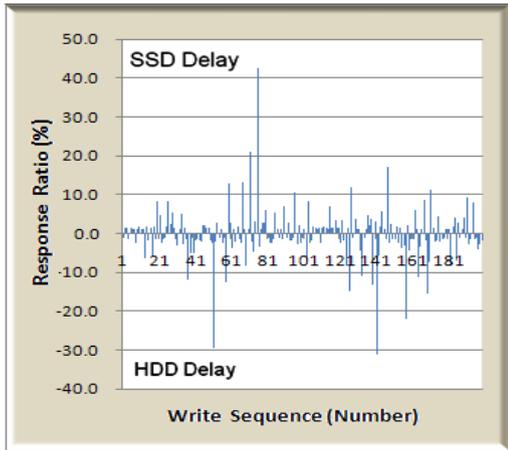
그림 7과 그림 8의 가로축은 저장 시스템에 가해지는 저장 연산의 부하를 의미하며, 그림 7의 세로축은 그 부하에서의 처리된 저장 연산수를 의미하며, 그림 8의 세로축은 저장연산에 걸린 응답시간을 의미한다.

그림 7에서 보면, 초당 저장 연산의 수가 대략 160개를 넘으면서 각 저장 기법의 성능이 점점 낮아진다. 이는 초당 저장 연산수의 증가에 의한 플래시 메모리 기반 SSD 스토리지에서의 I/O 연산 집중화가 성능에 영향을 크게 미치는 주요 요소임을 의미하며, 이 수치 이상으로 워크로드를 활성화시키는 것이 성능 향상에 도움이 되지 않음을 의미한다. 특히 SSD를 구성하는 플래시 메모리의 쓰기 연산의 느린 속도와 수반 되는 소거 연산이 주요 요인이다. 그래프에서 보면 저장 연산의 부하가 적은 시작 구간에서는 SSD의 느린 쓰기와 소거 연산이 시스템 안에서 수용되어 성능 저하를 일으키지 않는다. 하지만, 시작 구간을 지나면서 생성되는 저장 연산의 수가 증가하므로, 플래시 메모리에 하위 레벨의 쓰기 연산의 수도 증가하게 된다. 따라서 중간 구간 이후에서는 저장 연산들이 SSD 컨트롤러의 작업 큐에 과도하게 쌓이면서, 저장 연산 처리가 지체되며 결과적으로 성능이 서서히 저하되게 된다.

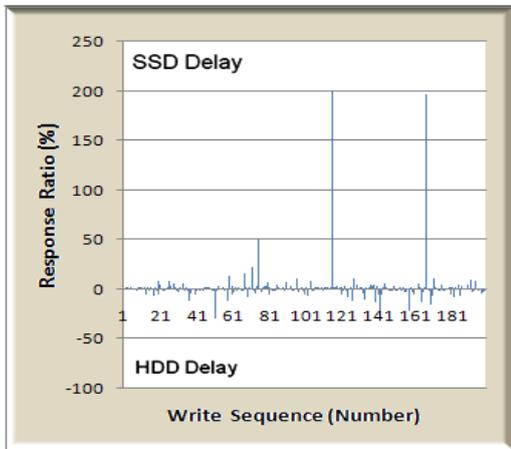
하지만, 동일한 워크로드의 조건에서도 제안한 Hete-Mr 저장 기법이 SSD-1나 SSD-Mr 기법에 비하여 성능이 상대적으로 더 높다. 전체 구간에서 평균하면, Hete-Mr가 기존 기법보다 17.8% 정도 처리 성능이 높다. 특히, 최대 저장 연산 처리치는 Hete-Mr가 147개로 SSD-Mr의 114개 보다 28.9%정도 더 높다. 그 이유는 위의 저장 성능 저하의 주요 원인인 SSD의 쓰기 연산의 부담에 의한 일시적인 홀딩현상을 헤테로 미러링기법은 일정 시간 비율을 초과하여 심각하게 길어지는 매체의 경우에는 일단 한 매체에 저장후 리턴하고, 늦어지는 매체는 지연 저장을 수행함으로써 전체적인 대기시간을 감소시킬 수 있었기 때문이다. 또한 SSD의 플래시 메모리 영역에 대하여, 쓰기 연산이 뒤로 지연되어 모아쓰기를 유도하는 효과도 발생하므로, 플래시 메모리의 수명이 조금 더 연장되며, 소거 연산을 위한 전력소모도 좀 더 줄이는 부수적인 효과도 기대할 수 있다.

그림 8에서 사용자 프로그램이 요청한 저장 연산에 대한 각 스토리지 관리 기법의 응답시간을 비교해보면, Hete-Mr, SSD-1, SSD-Mr 순으로 우수하게 나타났다. 단일 SSD를 사용하는 SSD-1 기법에 비하여, SSD를 미러링한 SSD-Mr가 응답시간이 5% 더 늦게 나타났는데, 이는 SSD-Mr는 두 개의 SSD에 모두 즉, 2번의 저장이 완료되어야 하기 때문이다. Hete-Mr는 SSD-1에 비하여, 38% 정도 응답성능이 빠른데, 이는 전술한 데로, Hete-Mr는

HDD나 SSD중에 어느 한쪽이 먼저 저장되면 바로 저장 연산을 완료시키므로 SSD의 심각한 홀딩현상을 회피가능하기 때문이다. 이러한 홀딩현상이 풀리면, 지연되었던 저장연산이 최종적으로 완료되게 된다.



[그림 9] 응답 시간의 비교(초당 100개 부하)



[그림 10] 응답 시간의 비교(초당 200개 부하)

그림 9와 그림 10은 제안 기법이 저장 연산의 처리치와 응답시간 측면에 대해서는 좋은 성능을 보이는 결과의 기본 원인이라 할 수 있는 헤테로-미러링 저장의 효과 분석을 위한 그래프이다. 이 실험에서는 500KB이하의 스몰 파일을 계속 저장하여, 부하를 지속시켜 프리징 상황을 만든다.

그림 9와 그림 10의 가로축은 이 저장연산의 순차 번호를 의미하며, 세로축은 이 저장연산을 수행하는데 걸린 HDD와 SSD의 응답시간의 비율을 의미한다. 즉, 저장 연산을 수행하였을 경우, 수평축의 윗부분은 SSD 응답시간

비율이며, 아랫부분은 HDD의 응답시간 비율이다. 예를 들면, 두 매체에 저장을 요청하였을 경우, SSD의 저장 시간이 2초이고, HDD의 저장 시간이 1초이면, 상단으로 y값으로 2를 표시하게 된다. 반대의 경우는 하단으로 y축값으로 -2를 표시하게 된다. SSD는 HDD에 비하여 읽기는 빠르지만, 저장 성능은 비슷하거나 약간 더 좋은 수준이므로, 그림 9에서 아래위로 비슷한 막대그래프가 나타난다. 여기서 주목할 점은 여러 지점에 아래위로 높은 막대그래프가 나타나는데, 그 저장 속도의 상대 비율은 무려 30-40배 까지도 높아진다.

특히,그림 10에서는 그림 9에서 보다 두 배의 워크로드를 가한 경우인데, 200배까지 저장 속도의 차이가 나타난다. 특이점은 이러한 과도한 저장 속도의 차이는 대부분 SSD의 저장 속도의 심각한 저하에 의한 것이다. Hete-Mr기법은 바로 이러한 심각한 저장 속도의 차이를 고려하여, 일시적으로 빠른 저장 매체를 우선 저장하고, 느린 저장 매체는 지연 저장을 수행하게 하여 전체적인 성능을 개선시킨다.

## 5. 결론

본 연구에서는 기존의 하드 디스크 기반 RAID 기술을 주요 레벨들에 대하여 비교 분석한 후, 새로운 SSD의 매체 특성에 맞게 보완하여 개선된 설계 방안을 제시하였다. 또한, 전통적인 RAID-1 에 기반하여 SSD의 특성을 고려한 헤테로-미러링이라는 하는 새로운 저장 장치 관리 기법을 제안하였다. 제안한 헤테로-미러링 기법은 SSD에서 발생하는 느린 쓰기 속도와 프리징 현상을 회피하기 위하여, 저장 매체간의 심각한 저장 속도의 차이를 고려하여, 일시적으로 빠른 저장 매체를 우선 저장하고, 느린 저장 매체는 지연 저장을 수행하게 하여 전체적인 성능을 개선하게 된다. 실험 결과를 통하여 제안 기법이 쓰기 부하와 프리징을 감소시켜서 기존의 SSD-미러링 기반의 저장 관리 기법보다 연산 처리 측면에서 18%, 응답 시간 측면에서 38% 더 우수함을 확인하였다.

## 참고문헌

- [1] Samsung, Samsung SSD is NOW, <http://www.samsung.com/global/business/semiconductor/products/flash/ssd/2008/home/home.html>, 2009.
- [2] 변시우, “플래시 메모리-SSD의 인덱스 연산 성능 향상을 위한 압축된 핫-콜드 클러스터링 기법”, 한

국산학기술학회논문지, Vol. 11, No. 1, pp. 166-174, 2010.

- [3] Samsung, Samsung, what is NAND Flash based SSD?, [http://www.samsung.com/global/business/semiconductorproducts/flash/Products\\_FlashSSD.html](http://www.samsung.com/global/business/semiconductorproducts/flash/Products_FlashSSD.html), 2009.
- [4] 배영현, “고성능 플래시 메모리 SSD 설계 기술”, 정보과학회지 vol 25, No. 6, pp. 18-28, 2007.
- [5] Samsung, SpinPoint F Series, <http://www.samsung.com/global/business/hdd/productSubType.do?group=72&type=61&subtype=63>, 2007.
- [6] 변시우, “하이브리드 하드디스크 시스템을 위한 플래시 노드 캐싱 기법”, 한국산학기술학회논문지, Vol. 9, No. 6, pp. 1696-1704, 2009.
- [7] Solid Data Systems, "Comparison of Drives Technologies for High-Transaction Databases", Solid Data Systems, Inc. White paper, 2007.
- [8] 김호진,황인철,맹승렬,윤현수, “SDIO에서 RAID 레벨 5의 효율적인 구현,” 정보과학회 학술발표대회, 제31권 제1호, pp. 64-66, 2004.
- [9] 변시우,허문행, “플래시 메모리를 사용하는 효과적인 RAID 스토리지에 대한 연구”, 대한전자공학회 정보 및 제어 학술대회, pp. 240-242, 2009.
- [10] 최귀열, 박계원, “지능형 I/O 구조를 갖는 RAID 시스템의 성능향상을 위한 연구,” 정보통신학회지 제10권 11호, pp. 1989-1995, 2007.
- [11] David Roberts, Taeho Kgil, and Trevor Mudge, "Integratin NAND Flash Devices onto Servers", Communications of the ACM, Vol52. No.4, pp. 98-106, 2009.
- [12] Mesquite, "CSIM2.0", [http://www.mesquite.com/documentation/documents/CSIM20\\_User\\_Guide-C.pdf](http://www.mesquite.com/documentation/documents/CSIM20_User_Guide-C.pdf), 2008.

**변 시 우(Siwoo Byun)**

[정회원]



- 1989년 2월 : 연세대학교 이과대학 전산학과 (공학사)
- 1991년 2월 : 한국과학기술원 전산학과 (공학석사)
- 1999년 2월 : 한국과학기술원 전산학과 (공학박사)
- 2000년 3월 ~ 현재 : 안양대학교 디지털미디어학부 부교수

<관심분야>  
데이터베이스, 저장장치, 임베디드 시스템 등