

고성능 플래시 메모리 SSD(Solid State Disk) 설계 기술

엠트론 | 배영현

1. 서론

플래시 메모리는 비휘발성의 메모리 반도체로서 크기와 소비 전력이 작고 외부 충격에 강하기 때문에 다양한 응용 저장장치로 개발되었다. 최근에는 컴퓨팅의 이동성이 강조되면서 MP3P, PDA, 휴대폰 등의 이동 정보기에 내장되고 있으며, 메모리 카드나 USB 메모리의 형태로 상품화되어 대부분의 이동 저장장치 영역을 독차지하고 있는 상황이다. 또한 플래시 메모리를 이용하는 저장장치는 하드디스크에 비해 기계적인 구성 요소가 없기 때문에 데이터 접근 속도가 매우 빠르며, 외부 환경에 대한 내구성이 매우 강하다. 따라서 소형의 이동 저장장치로서 뿐만 아니라 SSD(Solid State Disk) 형태로 개발되어 일반적인 컴퓨터 시스템에서 하드디스크를 대체할 수 있는 잠재력을 가지고 있다.

플래시 메모리는 단위 용량의 가격이 하드디스크보다 매우 높기 때문에 디스크 드라이브 목적으로 널리 사용되기 위해서는 플래시 메모리 가격의 급격한 하락이 전제 조건이다. 그림 1은 데이터 저장 용도로 사용되는 NAND 플래시 메모리의 용량 및 가격 추세를 보여준다. 플래시 메모리의 용량은 일명 '황'의 법칙에 따라 매년 두 배씩 증가하고 있으며, 가격 또한 급격하게 하락하고 있다. 실제로 2007년 현재 NAND 플래시 메모리의 가격은 그림의 예상 추세를 이미 앞질러서 1GB 당 \$10 안팎에서 형성되고 있다. 따라서 현재도 32GB 플래시 메모리 SSD를 \$400 이하의 가격으로 제조할 수 있는 상황이다. 이와 같은 용량과 가격 추세는 지속적으로 유지될 것이기 때문에 빠른 시간 안에 품질에 대비하여 하드디스크와의 가격 경쟁력을 확보할 것이라고 예상된다.

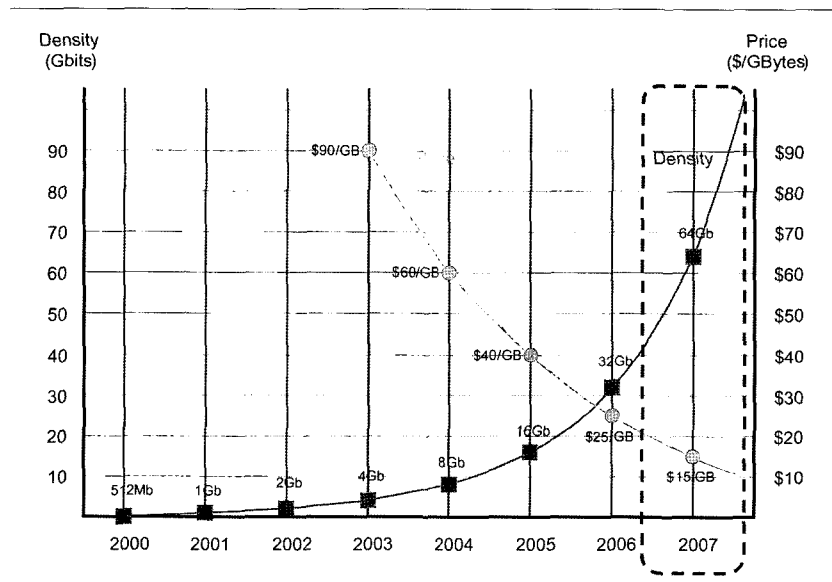


그림 1 NAND 플래시 메모리의 용량과 가격 추세

† 본 연구는 정보통신부 및 정보통신연구진흥원의 IT신성장동력 핵심기술개발사업의 일환으로 수행하였음.
[2006-S-040-01, Flash Memory 기반 임베디드 멀티미디어 소프트웨어 기술 개발]

플래시 메모리 SSD는 기본적으로 안정성 측면에서 하드디스크보다 우수한 특성을 가지고 있다. 그러나 하드디스크와의 경쟁에서 이기기 위해서는 기존의 메모리 카드나 USB 메모리에 비해서 대용량을 지원해야 하고 데이터 전송률이나 접근 속도 면에서 월등히 뛰어난 성능을 보장해야 한다. 샌디스크와 삼성전자, 그리고 애플에서는 이러한 요구에 부합하는 플래시 메모리 컨트롤러를 개발하여 향후 플래시 메모리 SSD 시장을 개척해 나갈 준비를 하고 있다.

본 고에서는 플래시 메모리의 특성과 응용 저장장치의 기본 구조를 설명하고, 기존의 하드디스크를 대체할 수 있는 고성능 대용량의 플래시 메모리 SSD를 구현하기 위한 컨트롤러 설계 기술을 설명한다. 또한 본 고의 설계 기술이 적용된 플래시 메모리 SSD 제품의 성능을 기존 하드디스크 및 타사 SSD 제품과 비교, 평가함으로써 본 설계 기술의 우수함을 보인다.

2. 플래시 메모리와 SSD

데이터 저장용으로 사용되는 NAND 플래시 메모리는 그림 2와 같은 구조를 가지고 있다. 데이터는 페이지(page)라는 공간에 저장되며 보통 64개의 페이지가 모여서 하나의 블록(block)을 구성한다. 현재 주로 사용되는 플래시 메모리에서 페이지는 2KB의 사용자 데이터를 저장하는 영역(main area)과 관련된 메타 정보를 저장하는 부가 영역(spare area)으로 구성된다. 즉, 하나의 블록은 128KB의 데이터를 저장할 수 있으며, 8192개의 블록으로 구성된 플래시 메모리 칩은 1GB의 데이터를 저장할 수 있다[1]. 최근에는 하나의 저장 단위(cell)에 2bit 정보를 저장할 수 있는 MLC(Multi-Level Cell) NAND 플래시 메모리가 개발되어 대용량화와 가격 하락을 촉진하고 있다. 또한 페이지의 크기도 4KB로 확장되면서 플래시 메모리의 용량이 커지고 처리 성능이 향상되고 있는 추세이다.

플래시 메모리에 데이터를 저장하고 읽는 방식은 기존의 RAM이나 하드디스크와는 다른 특성을 가지

고 있다. 읽기/쓰기(read/program) 동작은 페이지 단위로 이루어진다. 기존 데이터를 갱신하여 다시 쓰기 위해서는 해당 페이지를 소거(erase)해야만 한다. 이때 소거 동작의 단위는 블록이기 때문에 특정 페이지의 데이터를 갱신하기 위해서는 동일 블록에 속한 다른 페이지들도 소거되어야하는 제약을 가지고 있다. 플래시 메모리의 읽기, 쓰기, 소거는 각각 서로 다른 동작 시간을 요구한다. 반도체 메모리임에도 불구하고 데이터 전송 시간외에 각 동작을 처리하기 위한 내부 동작 시간이 필요하다. 일반적인 NAND 플래시 메모리에서 페이지의 읽기와 쓰기는 각각 25us, 200us이며, 블록의 소거는 2ms의 시간을 필요로 한다.

이와 같이 플래시 메모리의 기본 동작은 기존의 저장 매체와 상이하기 때문에 사용자에게 단순한 동작의 저장장치로 보여주기 위해서는 인터페이스의 변환을 위한 별도의 컨트롤러와 펌웨어(FTL, Flash Translation Layer)가 필수적이다[2,3]. 소거 후 데이터를 기록할 때 결과의 원자성을 보장하기 위해서는 쓰여질 데이터 위치(주소)를 새로 할당하여 기록해야한다. 또한, 소거와 쓰기 단위 크기가 서로 다른 제약 및 각 기본 동작이 불균일한 동작 시간을 가지는 제약을 극복하고 데이터 처리 성능을 높일 수 있는 주소 재사상(address re-mapping) 기법이 필요하다. 플래시 메모리의 각 블록은 제한된 횟수의 소거와 쓰기만 보장한다. 따라서 FTL 펌웨어는 사용자의 데이터 접근이 일정 영역에만 집중되지 않도록 하고, 일부 블록에서 오류가 발생하더라도 저장장치의 기능을 계속 제공하기 위한 블록 관리 기법을 포함하고 있어야 한다. 플래시 메모리 기반의 모든 저장장치에서는 이와 같은 컨트롤러와 FTL 펌웨어가 핵심적인 역할을 수행하고 있으며 용도에 따라 다양한 호스트 인터페이스가 부가되어 구성된다.

플래시 메모리의 새로운 응용 분야로 주목받고 있는 SSD는 기존 하드디스크와 같이 ATA 등의 호스트 인터페이스에 기반한 저장장치로서 하드디스크에 비

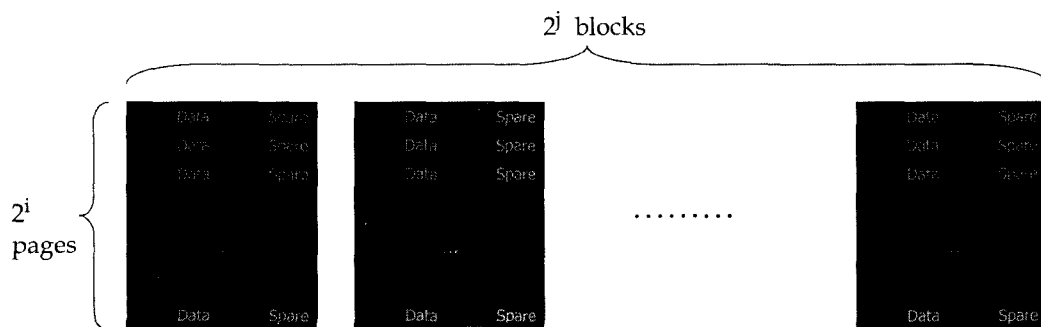


그림 2 NAND 플래시 메모리의 구조

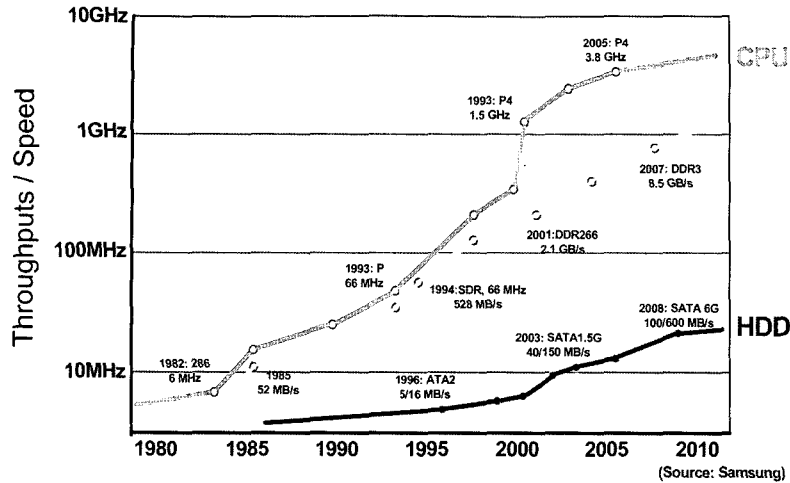


그림 3 컴퓨터 시스템의 성능 발전 추세

해 여러 가지 우수한 특성을 가지고 있다. 기계적인 구동 장치 없이 동작하기 때문에 외부 충격에 대한 안정성이 매우 우수하며, 소비 전력과 데이터 접근 시간이 작다. 따라서 1990년대 후반부터 군사용 혹은 특수 산업용으로서 고도의 안정성을 요구하는 컴퓨터 시스템의 구성요소로 개발되고 사용되어 왔다. 최근에는 NAND 플래시 메모리의 보급이 확대되면서 보다 다양한 분야로 플래시 메모리 SSD의 요구가 늘어가고 있다. 대표적인 요구로는 컴퓨터 시스템의 성능 향상을 들 수 있다. 그림 3에서 볼 수 있듯이 컴퓨터 시스템을 구성하는 주요 부품인 CPU, 메모리는 매년 급속한 발전을 거듭하여 초창기에 비해 1000배에 가까운 성능을 보이고 있다. 이에 비해 하드디스크는 현재 컴퓨터 시스템에서 유일한 기계적 장치로서 발전 속도가 상대적으로 더딘 추세를 보여준다. 따라서 현재의 초고속 컴퓨터 시스템에서는 상대적으로 더 많이 느려진 하드디스크의 병목에 의해 전체 컴퓨터 시스템의 성능이 한계를 보이게 된다.

컴퓨터 시스템 성능 상의 병목 현상을 해소하기 위해서는 하드디스크를 대체하여 보다 높은 데이터 입출력 성능을 보장하는 저장장치가 필요하다. 플래시 메모리 SSD는 반도체로 구성되기 때문에 기계적 동작에 따른 지연 시간을 제거하여 고성능을 구현할 수 있는 저장장치이다. 디스크의 회전과 헤드의 움직임이 필요 없기 때문에 수십배 이상 빠른 접근 시간을 구현할 수 있고, 또한 다양한 메모리 대역폭(bandwidth) 확대 기술을 적용하여 높은 데이터 전송률을 구현할 수 있는 가능성을 가지고 있다. 그러나 현재까지 개발된 다양한 플래시 메모리 컨트롤러와 플래시 메모리 SSD들은 이와 같은 요구를 만족시킬 만한 성능을 구

현하지 못했다. 왜냐하면 플래시 메모리는 다른 반도체 메모리와 다른 동작 특성을 가지고 있어서 메모리 대역폭이나 접근 시간을 원하는 만큼 확대하거나 줄이는 것이 쉽지 않기 때문이다.

그림 4는 대부분의 플래시 메모리 저장장치에 적용되는 컨트롤러의 기본 구조를 보여준다. 특별한 인터페이스를 가진 플래시 메모리에 접근하기 위한 플래시 인터페이스 모듈, 각 저장장치의 용도에 따라 필요한 호스트 인터페이스 모듈, 그리고 FTL 펌웨어를 구동하기 위한 CPU와 SRAM 부분으로 구성된다. 여기서 호스트 인터페이스는 하드디스크에 비해 느린 USB, MMC, SD, CF 등이 사용되었고 기존의 플래시 메모리 SSD의 경우에도 ATA 인터페이스를 사용하기는 했지만 지원하는 전송모드가 낮았기 때문에 호스트와의 데이터 전송에서부터 원하는 성능을 구현하기가 어려웠다. 또한 호스트와 플래시 메모리 사이에 데이터를 주고받기 위해서 일반적인 내장형 시스템에서 주로 채택하듯이 내부 SRAM과 시스템 버스를 이용한 DMA 기법을 사용했다. 이것은 시스템 버스의 점유율을 높여서 CPU의 동작에 영향을 주기 때문에 성능을 저하시킬 뿐만 아니라 SRAM을 거쳐서 가는 지연에 의해 초기 응답 시간을 최소화하는데 문제가 될 수 있다.

기존 플래시 메모리 컨트롤러 구조에서 성능 한계를 유발하는 가장 큰 원인은 플래시 메모리 인터페이스이다. 먼저 플래시 메모리 칩 한 개는 상대적으로 매우 낮은 쓰기 성능을 제공한다. 하나의 페이지를 기록하기 위해서 200us의 내부 동작 시간이 필요하기 때문에 데이터 전송 시간을 제외하더라도 최대 10MB/s (= 2KB/200us)의 데이터 전송률 밖에 보장할 수 없는 것이다. 또한 실제로 데이터 전송 시간 및 블록 소거

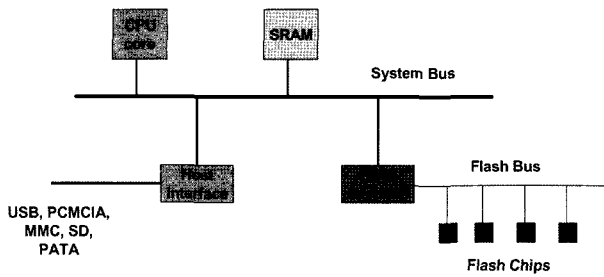


그림 4 플래시 메모리 컨트롤러의 기본 구조

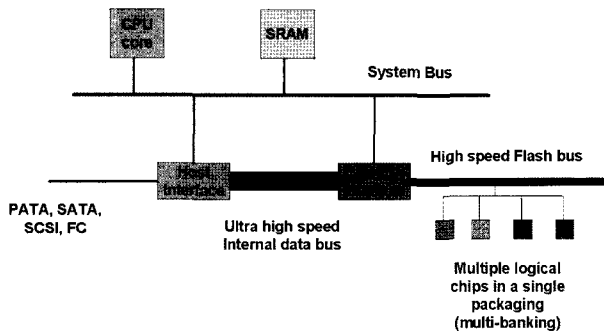


그림 5 고성능 플래시 메모리 컨트롤러의 구조

시간, 그리고 실제 구현된 시스템에서의 오버헤드 등을 고려하면 대략 6MB/s 정도가 최대 전송률이다. 따라서 보다 높은 성능을 구현하기 위해서는 여러 개의 플래시 메모리 칩의 쓰기 동작을 동시에 구동하는 인터리빙(interleaving) 기술이 필수적이다. 또한 현재의 플래시 메모리는 내부 동작 시간을 고려하지 않더라도 데이터 전송 속도가 DRAM 등에 비해 상대적으로 매우 낮다. 현재 양산 중인 플래시 메모리 칩의 최대 데이터 전송률은 25ns/byte(40MB/s)이다. 즉, 다수 개의 플래시 메모리를 인터리빙으로 처리하여 내부 구동 시간의 제약을 없앴다고 하더라도 40MB/s 이상의 데이터 대역폭을 제공할 수 없는 것이다. 따라서 고속의 플래시 메모리 컨트롤러는 이러한 데이터 대역폭을 확대하기 위해 다수 개의 플래시 메모리 버스를 동시에 운용할 수 있어야 한다.

그림 5는 기본적인 플래시 메모리 컨트롤러가 가지는 성능 상의 문제점을 개선하여 고성능의 플래시 메모리 SSD를 구현하기 위한 컨트롤러의 설계 방향을 보여준다. 먼저, 호스트 인터페이스는 고속의 데이터 전송을 지원하는 최신 규격을 채용한다. ATA의 경우, 133MB/s를 지원하는 Ultra DMA 모드 6, 혹은 최근에 각광받는 직렬 통신 방식인 Serial ATA 1, 2 등을 채용한다. 호스트 인터페이스와 플래시 메모리 사이에는 고속의 전용 데이터 버스를 적용함으로써 초기 응답 시간을 최소화하고 시스템 버스에 대한 간섭 없이 데이터 대역폭을 극대화할 수 있다. 플래시 인터페이스

에서는 플래시 메모리에 대해서도 높은 데이터 대역폭을 얻기 위하여 멀티 채널로 확대된 고속의 플래시 메모리 버스를 구현해야 한다. 예를 들어, 8bit 플래시 메모리 버스를 4채널로 확대한 경우에 최대 160MB/s (= 4×40MB/s)의 대역폭을 제공할 수 있다. 또한 플래시 인터페이스는 다수 개의 플래시 메모리 칩을 하나의 논리적인 그룹으로 접근함으로써 연속된 데이터를 여러 개 칩 상의 페이지에 동시에 기록할 수 있다. 예를 들어, 8개의 칩을 하나의 그룹으로 구성하는 경우에는 8개의 2KB 페이지를 동시에 혹은 인터리빙 방식으로 기록함으로써 최대 80MB/s(= 8×10MB/s)의 플래시 메모리 대역폭을 제공할 수 있다.

플래시 메모리의 데이터 대역폭을 증가시키기 위해 플래시 버스나 칩의 인터리빙 수준을 높인다면 동시에 소거되어야 하는 논리적인 처리 단위(논리 블록)가 커진다. 이럴 경우 하나의 논리 블록보다 작은 크기의 쓰기 요청에 대해서는 불필요한 소거와 데이터 복사로 인한 오버헤드가 급증하여 저장장치의 전반적인 성능 특성을 저하시키는 원인이 된다. 예를 들어, 데이터 전송 및 구현 상의 오버헤드를 감안한 실제 시스템에서 80MB/s의 쓰기 성능을 제공하기 위해서는 최소 16개 이상의 플래시 메모리 칩을 4개 이상의 버스에서 인터리빙 방식으로 접근해야 한다. 이때 하나의 논리 블록은 2MB(= 16×128KB)가 되기 때문에 대부분의 쓰기 요청이 전체 블록을 채우지 못하게 되어 비효율적으로 처리될 수밖에 없다. 즉, 논리 블록의 나머지 데이터가 이어지는 쓰기 요청에 의해 채워지지 않으면 해당 데이터들은 기존 플래시 메모리 블록으로부터 복사되어야 하는 것이다. 플래시 메모리 SSD가 일반적인 컴퓨터 시스템에서 우수한 성능 특성을 제공하기 위해서는 데이터 대역폭을 극대화시키면서도 상대적으로 작은 크기의 데이터 접근에 대해서 일정 수준 이상의 효율적인 처리 성능을 보장해야만 한다. 따라서 고성능 플래시 메모리 SSD의 FTL 펌웨어는 논리 블록을 적절한 크기로 제한하여 작은 크기의 접근에 대한 효율성을 제공하면서도 데이터 대역폭을 극대화하는 고도의 설계 기술이 필요하다.

다음 장에서는 이와 같은 접근 방법에 기초한 고성능 플래시 메모리 SSD용 컨트롤러 설계 기술을 설명한다. HYDRA 아키텍처라고 불리는 본 설계는 다수 개의 플래시 메모리를 그룹화하여 동시에 접근할 수 있고, 효과적인 쓰기 버퍼를 활용하여 여러 개의 그룹을 동시에 처리함으로써 연속적이지 않은 데이터에 대해서도 플래시 메모리 인터리빙을 수행할 수 있다. 또한, 연속된 플래시 메모리 작업에 대해 FTL 펌웨어

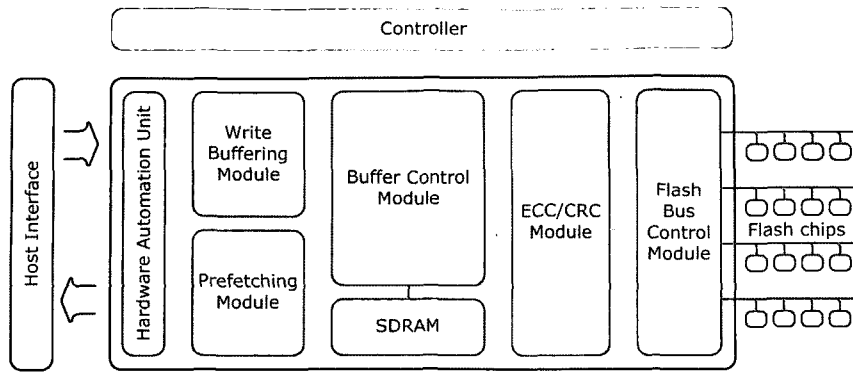


그림 6 HYDRA 컨트롤러 아키텍처

의 개입을 최소화하여 자동적으로 수행하는 지능적인 플래시 인터페이스를 포함하고 있다. 플래시 메모리에 대한 높은 수준의 병렬처리 기능과 밀접하게 결합되어 설계된 FTL 펌웨어를 통해 HYDRA 아키텍처는 기존 하드디스크가 가지는 성능 문제를 효과적으로 해결하는 플래시 메모리 SSD를 구현할 수 있도록 한다.

3. 고성능 플래시 메모리 SSD 설계 기술

3.1 HYDRA 컨트롤러 아키텍처

HYDRA 아키텍처의 핵심 접근 방법은 각 구성 모듈에 대해 다수 개의 제어 채널을 이용하여 동시에 처리하는 병렬성을 극대화하는 것이다. 플래시 메모리의 경우에는 다수 개의 버스와 다수 개의 플래시 메모리를 인터리빙하는 기능을 포함하며, 버퍼 데이터의 처리에 있어서도 다수 개의 데이터 스트림을 동시에 처리하도록 한다. 또한 플래시 메모리에 대한 제어 채널을 이중으로 구현하여 FTL 펌웨어의 제어에 의해 플래시 메모리를 접근하는 전위(foreground) 작업과 버퍼의 데이터를 처리하는 후위(background) 작업을 효율적인 스케줄링에 따라 동시에 처리할 수 있도록 구현하였다.

그림 6은 HYDRA 컨트롤러의 개략적인 아키텍처를 보여주는 블록도이다. FTL 펌웨어는 그림의 “Controller” 부분에서 수행되면서 호스트 명령의 해석과 처리, 플래시 메모리 상의 주소 변환 계산, 컨트롤러의 각 모듈에 대한 최상위 제어 등의 주요한 역할을 수행한다. HYDRA 컨트롤러의 모듈 구성은 크게 3가지 부분으로 나뉜다. 첫 번째로 호스트 인터페이스 부분은 ATA-7 규격에 부합하는 저장장치 프로토콜을 처리하는 모듈로서 플래시 메모리 SSD의 빠른 응답 시간을 구현하기 위한 자동화 기능을 포함하고 있다. 이것은 ATA 명령을 인식한 다음에 일정한 수준까지 펌웨어의 지시 없이 자동적으로 데이터를 처리하는 기능으로서 순차

적인 데이터 접근뿐만 아니라 임의의 주소에 대한 데이터 읽기도 매우 빠르게 처리할 수 있다. 기계적인 동작이 필수적인 하드디스크에 비해 100배 이상의 접근 속도 성능을 보일 수 있는 요인이다.

두 번째는 SDRAM으로 구현되는 버퍼를 관리하는 부분이다. HYDRA 아키텍처에서 버퍼는 기존 하드디스크에서의 버퍼와 같이 호스트의 순간적인 쓰기 요청을 휘발성의 SDRAM 메모리에 빠르게 저장한 다음에 일정한 재구성 절차를 거쳐 플래시 메모리에 보다 효율적으로 기록하기 위한 목적으로 사용된다. 또한 읽기 요청에 대해서는 연속된 데이터를 버퍼에 미리 읽어둠(prefetch)으로써 다음 읽기 요청에 대한 처리를 SDRAM에서 빠르게 처리하여 응답 시간을 최소화하는 용도로 사용된다. 버퍼에 저장된 데이터를 플래시 메모리에 내보내어 기록하는 작업(flush)은 지능적인 버퍼 컨트롤 모듈에 의해 현재 FTL 펌웨어의 작업에 대한 후위 작업으로 수행된다. 따라서 버퍼 데이터의 기록 작업이 진행되는 도중에도 호스트에서 읽기 요청이 발생하면 명령의 해석이 동시에 이루어지고 최소한의 응답 시간으로 처리할 수 있다. 버퍼의 데이터를 비워야 할 필요가 있을 때 버퍼 컨트롤 모듈은 버퍼에 저장된 데이터들의 논리 블록 정보를 재구성하여 해당하는 데이터를 플래시 버스 컨트롤 모듈로 전송한다. 이때, 하나의 연속된 데이터를 순차적으로 처리하는 대신에 4개의 독립된 논리 블록 정보를 동시에 처리함으로써 플래시 버스 컨트롤 모듈의 인터리빙 기능을 최대한으로 활용할 수 있다. 또한 논리 블록의 크기를 비교적 작게 설정하더라도 4개의 서로 다른 논리 블록을 동시에 처리할 수 있기 때문에 실질적인 인터리빙의 수준은 논리 블록의 4배 크기로 확대되는 것이다. 결국, 실질적인 플래시 메모리 인터리빙 수준을 최대화하여 높은 데이터 대역폭을 구현하면서도 FTL의 주소 사상은 비교적 작은 크기의 논리 블록을 바탕으로

처리되기 때문에 작은 크기의 데이터 접근을 효율적으로 처리할 수 있다.

세 번째로는 플래시 메모리에 대해 최대화된 병렬적 혹은 인터리빙 접근을 구현한 플래시 버스 컨트롤 모듈이 있다. 플래시 버스 컨트롤 모듈은 4개의 8bit 플래시 버스를 동시에 제어할 수 있으며 각 버스에서는 최대 16개의 플래시 메모리 칩을 인터리빙 방식으로 접근할 수 있다. 이러한 높은 수준의 병렬성은 기존의 어떠한 플래시 메모리 컨트롤러에서도 제공하지 못한 기능으로서 단일 플래시 메모리 컨트롤러로서 높은 확장성(scalability)을 보장하는 것이다. 이에 따라 HYDRA 아키텍처에서는 실질적인 플래시 메모리의 쓰기 대역폭으로 80MB/s, 읽기 대역폭으로는 100MB/s 이상의 성능을 제공할 수 있다. HYDRA 아키텍처의 플래시 버스 컨트롤 모듈은 매우 지능화된 제어 기능을 가지고 있다. 다수 개의 플래시 메모리에 대한 인터리빙 접근뿐만 아니라 동일한 형태의 연속된 접근에 대해서도 하드웨어 수준에서 자동화하여 처리함으로써 성능을 최대화하였다. 즉, FTL 펌웨어는 최상위 수준에서 제어 인자를 설정하고 수행 결과를 확인할 뿐 개개의 플래시 메모리 접근 과정에는 전혀 개입하지 않기 때문에 CPU의 코드 수행에 따른 지연 시간을 최소화하고 데이터 전송 과정에서도 중단됨이 없이 처리할 수 있다.

마지막으로 버퍼 컨트롤 모듈과 플래시 버스 컨트롤 모듈 사이의 데이터 전송 경로에는 데이터 에러를 감지하고 정정하기 위한 CRC 및 ECC 모듈이 포함되어 있다. NAND 플래시 메모리는 구조적으로 읽기 동작 시에 불규칙적으로 에러가 발생할 수 있다. 따라서 NAND 플래시 메모리가 적용된 저장장치에서 원활한 데이터 저장 기능을 제공하기 위해서는 별도의 에러 정정 기능을 구현해야 한다. Denali 사의 발표에 따르면 512bytes 데이터 당 SLC NAND를 위해서는 1bit의 에러 정정이 가능한 ECC가 필요하며, MLC NAND를 위해서는 최소한 4bit 이상의 에러 정정이 가능한 ECC가 필요하다[4]. HYDRA 아키텍처의 ECC 모듈은 Binary BCH 기반으로 7bit 에러 정정 능력을 제공하기 때문에 향후의 MLC NAND 플래시 메모리에 대해서도 수정 없이 적용될 수 있다. 또한, Micron 사의 발표에 의하면 Binary BCH 기반의 ECC 기법은 Reed-Solomon 기반의 기법에 비해 코드 부가 영역의 크기가 39% 절약되며, NAND 플래시 메모리에서 에러 위치가 서로 무관하게 발생하기 때문에 보다 효과적인 방법으로 알려져 있다[5]. HYDRA 아키텍처에는 ECC 기능뿐만 아니라 보다 신속한 에

러 감지를 위해 CRC 기능이 포함되어 있다. 즉, 데이터의 실제 전송 속도에 거의 영향을 주지 않는 속도로 CRC 및 ECC 코드를 생성하여 부가 영역에 저장하고, 읽기 동작에는 CRC 만을 먼저 검사함으로써 에러가 발생하지 않은 정상적인 상황에서는 원래의 데이터 대역폭을 그대로 유지하면서 동작한다.

3.2 HYDRA FTL(Flash Translation Layer)

HYDRA 아키텍처에서 FTL의 역할은 기존 플래시 메모리 응용 제품에서와 같이 사용자의 데이터 읽기 및 쓰기 요청을 플래시 메모리의 기본 동작에 맞게 변환하여 처리하는 기능이 핵심이다. 이를 위해 호스트 인터페이스의 프로토콜 규격에 따라 명령을 해석하고 처리해야 하며, 사용자 데이터의 논리적인 주소와 플래시 메모리의 물리적인 주소 사이의 변환 관계를 관리해야 한다. 또한 HYDRA 하드웨어의 각 모듈을 제어하여 가장 최적화된 방식으로 데이터가 전송되고 플래시 메모리에 저장되거나 읽히도록 해야 한다.

HYDRA FTL은 하드웨어가 제공하는 높은 수준의 병렬성과 자동성을 적극 활용하여 동작함으로써 플래시 메모리 SSD의 성능을 극대화하도록 설계되었다. 작은 크기의 접근을 처리하기 위한 오버헤드가 비교적 낮은 수준에서 플래시 메모리 버스 및 칩 인터리빙 수준을 최대화하였고, 하드웨어의 자동화된 플래시 메모리 처리 기능에 따라 펌웨어 코드 수행 부분을 최소화하였다. 또한, FTL에서 가장 많은 계산을 필요로 하는 주소 사상의 복잡성을 최소화하기 위하여 기본적인 블록 수준 사상을 따르면서도 블록의 병합(merge)은 SDRAM 버퍼의 데이터가 기록되는 즉시 수행되도록 하였다. 즉, 플래시 메모리에 저장된 데이터는 임시로 저장된 위치를 가지지 않고 언제나 하나의 일관성 있는 주소 사상 정보로만 관리되게 하였다. 따라서 데이터 읽기 시에 위치를 찾는 과정이 최소화되었고 전원 단절 등의 오류가 발생했을 때 데이터의 일관성을 복구하는 과정이 매우 단순해 졌다. HYDRA FTL에서는 버퍼로부터의 데이터 기록 시 마다 수행되는 즉각적인 블록 병합에 의해 주소 사상 정보의 변경이 매우 빈번하게 발생한다. 그러나 전체 주소 사상 정보 내에서 변경된 정보를 정해진 위치에 일일이 기록하지 않고 변경된 정보만을 모아서 점진적으로 기록하는 방식(incremental check-pointing)을 채택함으로써 주소 정보 갱신에 따른 오버헤드도 최소화하였다.

플래시 메모리 버스와 칩의 인터리빙 수준을 최대화하였을 때의 문제점을 해소하기 위하여 HYDRA FTL은 SDRAM 버퍼와 버퍼 컨트롤 모듈의 병렬 처리 기

능을 활용한다. 버퍼 컨트롤 모듈이 SDRAM 버퍼에 저장된 서로 다른 논리 블록을 동시에 처리할 수 있기 때문에, FTL은 주소 사상의 단위가 되는 논리 블록의 크기를 비교적 작게 설정할 수 있다. 즉, 4개 버스에서 각각 1개 혹은 2개씩의 플래시 메모리 칩을 그룹으로 묶어서 하나의 논리 블록을 만들더라도 SDRAM 버퍼에서 재구성된 여러 개의 논리 블록을 동시에 처리하여 각 버스 당 4개 혹은 8개의 플래시 메모리 칩을 인터리빙 방식으로 접근할 수 있는 것이다. 이렇게 함으로써 최대 데이터 대역폭을 극대화함과 동시에 작은 크기의 접근에 대해서도 비교적 효율적인 성능을 제공할 수 있다.

이 외에도 HYDRA FTL은 간단하고 명료한 배드 블록(bad block) 관리 기법을 포함하여 플래시 메모리의 오류 상황에서도 SSD의 안정성을 유지한다. 또한 SSD의 동작 중에 동적으로 수행되는 블록 소거 횡수 평준화(wear-leveling) 기법을 포함하고 있어서 사용자의 데이터 사용이 특정한 영역에만 집중되더라도 실제 플래시 메모리 상에서는 칩 내의 모든 블록에 그 영향을 골고루 분산시킬 수 있다. 따라서 특정 영역이 파괴되어 플래시 메모리 SSD의 수명이 단축되는 것을 효과적으로 방지할 수 있다.

4. 성능 비교 평가

본 고에서 소개한 HYDRA 아키텍처는 실질적인 플래시 메모리 SSD 컨트롤러 개발에 적용되어 제품화 되어있다. 그림 7의 SSD는 PATA 인터페이스를 가지며 기존의 3.5" 하드디스크와 동일한 규격으로 제품화 된 예이다. 구현된 HYDRA 컨트롤러는 최대 64개의 플래시 메모리를 장착할 수 있고, 호스트 인터페이스는 ATA-7 규격의 UltraDMA mode 6인 133MB/s의 데이터 전송률을 지원한다. 호스트 시스템에 따라 차이가 있지만 각종 벤치마크 테스트 결과에서 본 제품은 최대 데이터 대역폭으로 읽기 100MB/s, 쓰기 80MB/s 성능을 보였다. 또한 저장장치의 명령 처리 성능을 표시하는 IOPS(Input Output Operations Per Seconds) 수치에서도 최대 76,000이라는 매우 높은 결과를 보였고, 실제 컴퓨터 시스템에서의 일반적인 저장장치 사용 형태를 테스트하는 PCMark04에서도 매우 높은 수치로 하드디스크 및 타사 SSD와의 월등한 성능 차이를 보여 주었다.

현재 시중에서 판매되고 있는 SSD는 크게 세 가지 그룹으로 나누어 볼 수 있다. 먼저, 수년 전부터 개발되어 군사 및 산업체의 특수한 용도 혹은 고속의 컴퓨

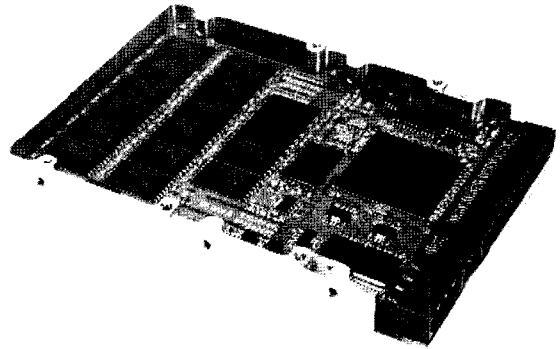


그림 7 HYDRA 아키텍처를 적용한 플래시 메모리 SSD 제품
터 시스템을 위한 용도로 사용된 SSD이다. M-Systems, Adtron, Bitmicro 등의 제품이 대표적이며 비교적 우수한 성능의 데이터 대역폭과 접근 속도를 제공한다. 또한 군사용 요구 사항에도 적합한 매우 견고한 특성을 가지고 있다. 현재 1GB 당 \$100 이상의 매우 고가로 판매되고 있다. 그러나 이들의 성능이 기존 HDD를 완전히 대체할 만큼 높은 수준은 아니며(30~60MB/s), 특히 대부분의 제품은 특정한 사용자 요청 형태에 대해서만 일정한 수준의 성능을 보장하는 약점을 가지고 있다.

두 번째는 ATA 인터페이스를 가지고 있기는 하지만 기존의 플래시 메모리 카드용 컨트롤러와 유사한 용량과 성능을 제공하는 컨트롤러를 채용하여 제작된 SSD이다. 이러한 제품의 데이터 대역폭은 20~30MB/s 수준이며, 임의 쓰기 성능이 매우 낮기 때문에 성능적인 면에서 하드디스크를 대체할 만한 특성을 보여주지 못한다. 단, SSD로서의 내구성과 소비 전력 등의 특성은 동일하기 때문에 성능에 민감하지 않는 분야의 산업용 컴퓨터에 적용되고 있다.

마지막으로 최근에 일반적인 컴퓨터 시스템에서 하드디스크를 대체할 성능과 내구성을 가진 저장장치로 개발이 완료된 본격적인 플래시 메모리 SSD 제품이 있다. 가격적인 측면에서 플래시 메모리 가격의 하락세를 가장 민감하게 반영하여 저가화를 지향하고 있으며, 성능적인 측면에서도 기존 하드디스크를 능가하도록 설계되고 있다. HYDRA 아키텍처가 적용된 제품이 대표적인 예이며, 삼성과 샌디스크에서도 이러한 특성의 SSD를 개발하여 새로운 SSD 시장에 진출하고 있다. 그러나 후자의 제품들은 공개된 스펙을 볼 때 성능적인 측면에서 HYDRA 아키텍처가 적용된 제품에는 많이 미치지 못하는 상태이다.

본 장에서는 HYDRA 아키텍처를 구현한 SSD와 타사의 SSD, 그리고 일반적인 하드디스크의 성능을 비교하

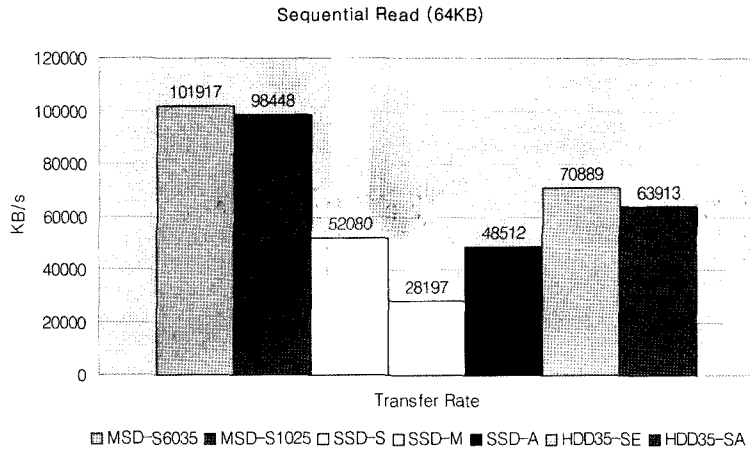


그림 8 ATTO Disk Benchmark 순차 읽기 성능(읽기 단위: 64KB)

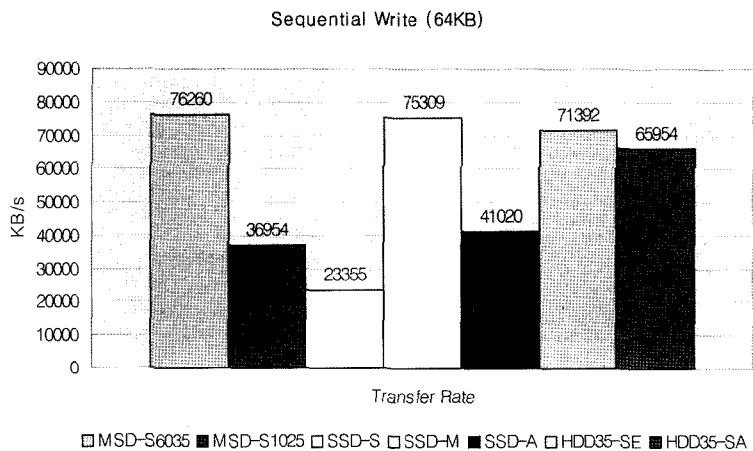


그림 9 ATTO Disk Benchmark 순차 쓰기 성능(쓰기 단위: 64KB)

여 평가한다. 알려진 몇 가지 벤치마크를 이용한 테스트 결과를 제시함으로써 HYDRA 아키텍처의 우수한 성능 특성을 객관적으로 검증한다. 본 테스트에 사용한 호스트 PC에는 ASUS 사의 M2N-E 메인보드가 내장되어 있고, AMD Athlon 64-X2 4800+ CPU와 PC2-5300의 1GB DDR2 메모리가 장착되어 있다. 그리고 디스크 장치와 인터페이스하기 위한 칩셋으로는 nVidia 사의 nForce 570 Ultra MCP(UDMA6, SATA2)이 사용되었다.

그림 8과 9는 각 디스크의 최대 데이터 대역폭을 테스트하기 위한 ATTO Disk Benchmark 프로그램의 수행 결과이다. 테스트 파일 크기는 32MB로 설정하고 테스트의 단위 데이터 크기는 64KB로 설정하였다. 그림에서 보듯이 HYDRA 아키텍처가 적용된 MSD 시리즈의 SSD 제품은 읽기 100MB/s, 쓰기 76MB/s의 데이터 대역폭을 보여준다. 단위 데이터의 크기에 따라 최대 데이터 대역폭에는 차이가 발생한다. MSD 시리즈 SSD 제품 중에서 MSD-S1025는 성능 특성이 좋지 않은 MLC

NAND를 적용한 제품이다. 이 경우에도 HYDRA 아키텍처의 제품은 읽기 98MB/s, 쓰기 37MB/s로 비교적 우수한 성능을 보인다. 이에 반해 S, M, A사의 SSD는 모두 SLC NAND로 구현되었지만 50MB/s 이하로서 상대적으로 높지 않은 성능이다. M사 SSD 제품의 경우, 쓰기 성능은 75MB/s로 오히려 읽기보다 우수하게 나타나는데 이것은 SSD 제품 내에 128MB의 SDRAM 버퍼를 가지고 있고, 테스트 파일의 크기가 32MB 밖에 되지 않기 때문에 버퍼에 의한 쓰기 대역폭 향상으로 분석된다. 최대 대역폭의 경우에는 하드디스크의 성능도 비교적 우수하다. 왜냐하면 최신의 하드디스크는 매우 빠른 디스크 회전 속도로 동작하기 때문에 헤드가 고정된 다음 연속적으로 데이터를 읽거나 쓰는 최대 대역폭 테스트에서는 비교적 우수한 성능을 보이는 것이다. 따라서 그림에서 보듯이 HYDRA 아키텍처가 적용된 SSD를 제외하고는 현재의 타사 SSD들은 하드디스크의 데이터 대역폭을 능가하지 못하고 있는 상황이다.

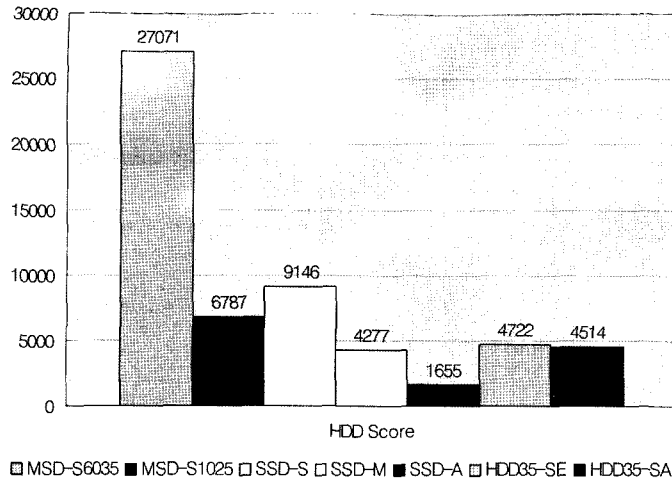


그림 10 PCMark04 하드디스크 테스트 종합 점수

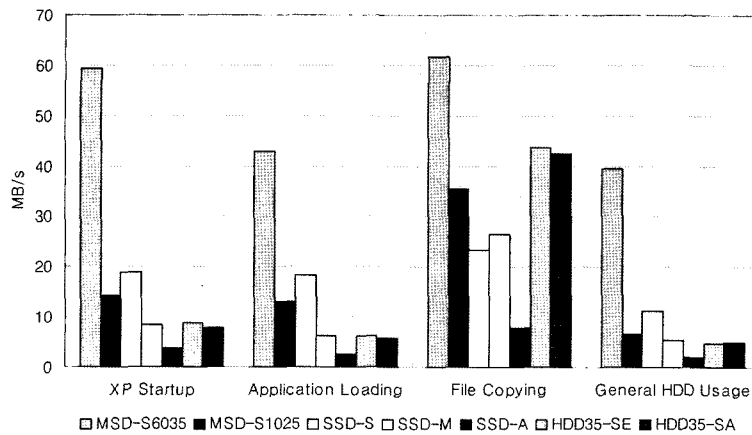


그림 11 PCMark04 하드디스크 테스트 각 항목별 성능

그림 10과 11의 그래프는 PCMark04 테스트 결과를 보여준다. PCMark04는 최대 대역폭을 테스트하기 위하여 지속적으로 순차적인 접근을 하는 등의 극단적인 사용 형태가 아니라 일반적인 컴퓨터 시스템에 흔히 나타나는 디스크 사용 형태를 이용하여 성능을 테스트한다. 따라서 사용자가 실제로 해당 디스크를 사용할 때의 체감 성능을 보여주는 것일 뿐만 아니라 저장 장치의 다양한 측면에서의 종합 성능을 평가할 수 있는 좋은 도구 중의 하나이다. PCMark04의 디스크 테스트는 4가지 항목으로 구성된다. 이 중에서 “File Copy”를 제외하고는 데이터의 임의 접근 비율이 큰 사용 형태이다. 따라서 그림 11에서 보는 바와 같이 SSD와 하드디스크의 성능 차이가 훨씬 많이 나타난다. 실제로 HYDRA 아키텍처와 SLC NAND로 구현된 MSD-S6035 제품은 하드디스크에 비해 10배 가량의 성능 차이를 보여준다. 타사 SSD의 경우에는 하드디스크와 유사하거나 최대 2배 정도의 성능 차이만 보일 뿐이다. 그리고 순차 접근 비율이 대부분인 “File Copy” 테스

트에서는 타사 SSD는 모두 하드디스크보다 낮은 성능을 보이고 있다. 특히, A 사의 SSD 제품은 최대 데이터 대역폭 성능은 비교적 우수하지만 임의 접근이 많은 PCMark04 테스트에서는 하드디스크보다도 매우 낮은 성능을 보여준다. 즉, 높은 병렬성으로 대역폭은 증가시켰지만 그 이상의 플래시 메모리 컨트롤러 기술을 개발하지 못하여 매우 불균형한 성능 특성을 보이고 만 것이다. 그림 10은 이러한 4가지 테스트의 결과를 가중 평균하여 계산한 디스크 점수를 보여준다. MSD-S6035 SSD 제품의 경우, 하드디스크에 비해 5배 이상, 타사의 SSD에 비해서도 3배 이상의 우수한 성능을 보여준다. MLC NAND에 기반한 MSD-S1025 SSD 제품의 경우에는 상대적으로 낮은 점수를 보여주지만 타사 SSD와는 유사한 수준의 성능이다. MLC NAND의 기본적인 동작 성능이 SLC NAND에 비해 4배 이상 낮을 것을 감안하면 이것 역시 HYDRA 아키텍처의 고성능 설계를 입증하는 부분이다. 또한 HYDRA 아키텍처는 하나의 구현으로 SLC와 MLC NAND를 모두 지원

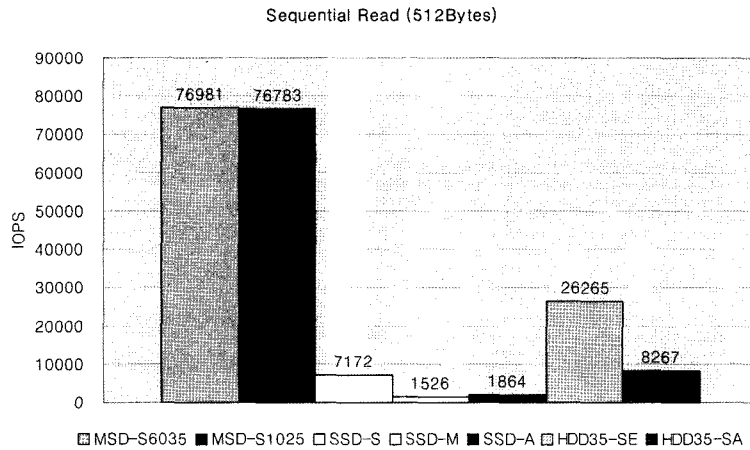


그림 12 순차 읽기 IOPS(읽기 단위: 512Bytes)

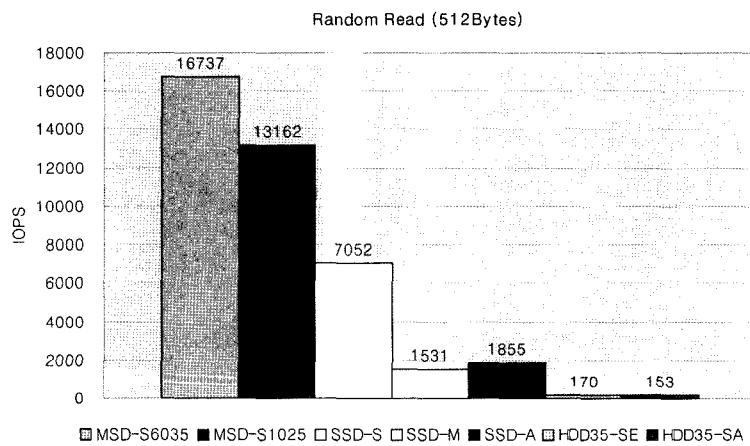


그림 13 임의 읽기 IOPS(읽기 단위: 512Bytes)

하는 매우 유연한 구조의 설계이다.

그림 12와 13은 저장장치의 최대 데이터 요청 처리 성능을 평가하기 위한 벤치마크 결과이다. IOMeter 2004.07.30 프로그램으로 테스트한 것이며 입출력 단위 크기는 512Bytes로 가장 작은 크기를 사용함으로써 IOPS의 최대값을 얻을 수 있었다. 그림에서 볼 수 있듯이 읽기의 경우에 HYDRA 아키텍처가 적용된 SSD 제품은 타사 SSD에 비해 10배 이상의 성능 차이를 보이며, 특히 기계적인 움직임에 따라 성능이 더욱 나빠지는 임의 읽기에서는 기존 하드디스크에 비해 100배 가량의 성능 차이를 보여준다. 이것은 기존 SSD에 비해서도 월등히 우수한 성능이며 컴퓨터 시스템에서 하드디스크를 HYDRA 아키텍처 기반의 SSD로 대체했을 때 현저한 성능 향상 효과를 얻을 수 있음을 보여주는 것이다.

현재 플래시 메모리 기반의 저장장치들은 임의 쓰기 데이터 접근에 대해 상대적으로 매우 좋지 않은 성능을 제공한다. 특히, 매우 높은 수준의 데이터 대역

폭을 구현하는 SSD 제품의 경우에는 논리 블록의 크기가 커지면서 상대적인 성능 저하가 더 두드러진다. HYDRA 아키텍처를 적용한 SSD 제품의 경우에도 대부분의 경우 기존 하드디스크보다 몇배에서 수백배까지의 성능 향상을 보이지만 작은 크기의 임의 쓰기는 유사하거나 조금 낮은 성능에 머물고 있다. 이런 성향은 타사의 SSD에서는 더 크기 때문에 상대적으로는 HYDRA 아키텍처가 여전히 우수한 설계라고 할 수 있지만 성능 면에서 개선해야할 최우선 목표이다. HYDRA 아키텍처에서 데이터 쓰기는 SDRAM에 버퍼링되었다가 필요시 전체 블록이 직접 플래시 메모리에 쓰여진다. 따라서 작은 크기의 쓰기가 임의 위치에 산재되어 접근될 때에는 기존 데이터를 복사하여 즉각적인 블록 병합을 완성시키는 오버헤드가 극대화되는 것이다. 개선된 아키텍처에서는 플래시 메모리 상에도 SDRAM 버퍼의 내용을 임시로 저장하는 영역을 두어 작은 크기의 데이터에 대해서는 이중으로 버퍼링을 하게 함으로써 이러한 급격한 성능 저하 현상을 개선할 예정이다.

5. 결론

플래시 메모리 SSD는 새롭게 전개되는 플래시 메모리 응용의 주목받는 제품으로서 기존의 하드디스크가 가지는 많은 문제점을 해결하는 우수한 특성의 저장 장치이다. 기계적인 동작이 전혀 없기 때문에 어떠한 환경에서도 고장 발생률이 매우 낮고, 반도체 메모리의 특성으로 데이터 처리 속도를 매우 빠르게 할 수 있다. 따라서 현재 컴퓨터 시스템에서 유일한 기계 장치인 하드디스크를 대체하여 매우 안정적인 새로운 개념의 컴퓨터 시스템을 추구할 수 있고, 하드디스크가 가지는 성능 상의 한계를 극복하여 전체 시스템의 성능을 획기적으로 향상시킬 수 있다. 그러나 여전히 상대적으로 가격이 높고, 현재까지의 알려진 제품 중에는 본격적인 SSD로서의 원하는 성능을 제공하는 제품이 존재하지 않기 때문에 이러한 교체의 움직임이 크지 않았다. 본 고에서 설명한 HYDRA 아키텍처는 병렬 처리의 가능성을 극대화하고 컨트롤러 각 모듈과 펌웨어의 설계를 유기적으로 최적화하여 매우 우수한 데이터 처리 성능을 제공한다. 이 기술이 적용된 SSD 제품의 성능 비교 평가 결과를 제시함으로써 우수한 성능 특성을 검증하였다. 지속적으로 유지되는 플래시 메모리의 가격 하락과 이와 같은 우수한 플래시 메모리 SSD 컨트롤러 설계 기술이 결합되어 멀지 않아 다양한 컴퓨터 시스템에서 하드디스크를 대체하는 플래시 메모리 SSD 제품을 볼 것으로 예상된다.

참고문헌

- [1] Samsung Electronics, Co., "1G x 8 Bit NAND Flash Memory(K9K8G08U0A) Data Sheets," <http://www.samsungelectronics.com/>, 2006.
- [2] Intel Corporation, "Understanding the Flash Translation Layer(FTL) Specification," Application Note 648, 1998.
- [3] J. Kim, J. M. Kim, S. H. Noh, S. L. Min, and Y. Cho, "A Space-efficient Flash Translation Layer for CompactFlash Systems," IEEE Transactions on Consumer Electronics, vol. 28, no. 2, pp. 366-375, 2002.
- [4] J. Cooke, "MLC NAND Technology," Denali Webcasts MLC NAND Flash Series, Part 2/3, Feb. 2007.
- [5] P. Feeley, "NAND Error Correction Code Choices," Flash Memory Summit 2006 Presentations, Aug. 2006.



배영현

1993 서울대학교 컴퓨터공학과(학사)
1995 서울대학교 컴퓨터공학과(석사)
2006 서울대학교 컴퓨터공학과(박사)
1999~2006 지인정보기술(주) 기술이사
2006~현재 (주)엠트론 연구소장
관심분야 : Embedded System, Flash Memory, Storage System
E-mail : yhbae@mttron.net

KCC 2007(한국컴퓨터종합학술대회)

- 일 자 : 2007년 6월 25~27일
- 장 소 : 무주리조트
- 내 용 : 논문발표 등
- 주 최 : 한국정보과학회
- 상세안내 : <http://www.kiss.or.kr/conference02>