

256GB 용량 DRAM기반 SSD의 설계

A Design of 256GB volume DRAM-based SSD(Solid State Drive)

고대식*, 정승국**

Dea-Sik Ko*, Seung-Kook Jeong**

요 약

본 논문에서는 DDR-1 메모리와 PCI-e 인터페이스를 이용하는 256 GB DRAM 기반의 SSD 스토리지를 설계 분석하였다. SSD는 주 저장매체로써 DRAM 이나 NAND Flash 를 사용하는 스토리지로써 메모리칩으로부터 직접 데이터를 처리할 수 있기 때문에 종래의 HDD의 기계적인 처리속도보다 매우 고속인 장점이 있다. 설계된 DRAM 기반 SSD 시스템은 복수 개의 RAM 디스크를 데이터 저장매체로 사용하며, PCI-e 인터페이스 버스를 각 메모리디스크의 통신 경로로 사용하여 고속의 데이터 처리가 가능한 구조이다. 실험을 위하여 UNIX 및 Windows/Linux 서버, SAN Switch, Ethernet Switch를 이용한 실험시스템을 구성하고 IOmeter 를 이용하여 IOPS(Input output Per Second)와 대역폭 성능을 측정하였으며 측정결과에서 DDR-1 SSD는 470,000의 IOPS와 800MB/sec로 HDD 나 Flash-based SSD 에 비하여 높은 대역폭이 나타남을 확인하였다.

Abstract

In this paper, we designed and analyzed 256GB DRAM-based SSD storage using DDR1 memory and PCI-e interface. SSD is a storage system that uses DRAM or NAND Flash as primary storage media. Since the SSD read and write data directly to memory chips, which results in storage speeds far greater than conventional magnetic storage devices, HDD. Architecture of the proposed SSD system has performance of high speed data processing duo to use multiple RAM disks as primary storage and PCI-e interface bus as communication path of RAM disks.

We constructed experimental system with UNIX, Windows/Linux server, SAN Switch, and Ethernet Switch and measured IOPS and bandwidth of proposed SSD using IOmeter. In experimental results, it has been shown that IOPS, 470,000 and bandwidth,800MB/sec of the DDR-1 SSD is better than those of the HDD and Flash-based SSD.

Key words : Solid state storage system, DRAM-SSD, Flash-SSD, HDD, IOPS, IO bottleneck

* 목원대학교 전자공학과 교수

** 한국전자통신연구원 책임연구원

· 제1저자 (First Author) : 고대식

· 투고일자 : 2009년 5월 4일

· 심사(수정)일자 : 2009년 5월 6일 (수정일자 : 2009년 7월 2일)

· 게재일자 : 2009년 8월 30일

I. 서 론

일반 스토리지 장치는 랜 카드 등의 호스트 인터페이스, 데이터를 저장/판독하는 HDD(Hard Disk Drive), 그리고 CPU와 메인 메모리로 구성되어 HDD를 제어하는 제어부로 구성된다. 제어부는 PCI 인터페이스를 통해 HDD로 데이터를 입출력시킨다. 또한, 호스트 인터페이스는 호스트 컴퓨터로부터 전달되는 데이터를 HDD에 저장시키거나, HDD로부터 판독하여 호스트 컴퓨터로 전달하는 역할을 수행한다. 종래의 HDD스토리지 장치는 구조상 기계적인 방식에 의해 데이터를 저장/판독하기 때문에, 데이터 읽기/쓰기 처리 속도가 현저히 떨어지는 문제점이 있다. 또한, 단순히 PCI 인터페이스에 의해 형성되는 데이터 통신 버스를 HDD 전용 통신 경로로 사용하기 때문에, 저장매체를 추가시켜 저장 용량을 확장시키는데 있어서 물리적인 한계가 있다.[1]

한편, SSD(Solid State Drive)는 HDD의 기계적 동작부분이 전혀 없기 때문에 HDD보다 높은 데이터 읽기/쓰기 액세스 속도를 가질 수 있다. 최근까지도 SSD는 가격이 비싸기 때문에 항공/우주/국방 산업분야와 같은 특수 분야에서 제한적으로 사용되어왔으나 메모리 가격의 하락과 고객의 고성능 스토리지 요구에 의하여 SSD는 플래시 기반 SSD를 중심으로 상품화가 가속화 되고 있다. SSD는 주저장 매체로 무엇을 사용하는가에 따라서 DRAM 기반 SSD와 플래시 기반 SSD 두 가지 유형이 있다. 플래시 기반 SSD는 주저장 매체로 플래시 메모리를 사용하는 유형이며 비휘발성, 고속의 랜덤 읽기, 저 전력 소모, DRAM보다 낮은 가격 등의 장점이 있기 때문에 이미 휴대용 단말기나 노트북 시장에서 각광을 받고 있다. 반면에 DRAM 기반 SSD는 주 저장 매체로 DRAM을 사용하는 유형이며 플래시 기반 SSD 보다도 고속의 읽기/쓰기 그리고 내구성 부분에서 장점이 있지만 GB당 가격이 높고 휘발성 매체를 사용하기 때문에 신뢰성 확보를 위한 백업장치 확보가 필요하며 휘발성과 고 가격 때문에 개인용보다는 고성능을 필요로 하고 24시간 무정전운영을 기본으로 하고 있는 기업용 시장을 목표로 시장을 개척하고 있다[2]-[5].

본 논문에서는 DDR-1 메모리와 PCI-e 인터페이스

를 이용하는 DRAM 기반의 SSD 스토리지를 설계하였다. SSD는 주 저장매체로써 DRAM 이나 NAND 플래시를 사용하는 스토리지로써 메모리칩으로부터 전자적으로 직접 데이터를 처리할 수 있기 때문에 종래의 HDD의 기계적인 처리속도보다 매우 고속인 장점이 있다. 본 논문에서 설계된 SSD 시스템은 복수 개의 DRAM를 데이터 저장매체로 사용하며, PCI-e 인터페이스 버스를 각 DRAM의 통신 경로로 사용하기 때문에 고속의 데이터 처리가 가능한 구조이다. 실험을 위하여 UNIX 및 Windows/Linux 서버, SAN Switch, Ethernet Switch를 이용한 실험시스템을 구성하고 Iometer 를 이용하여 데이터 블록의 크기에 따라서 HDD, Flash SSD, 그리고 DRAM기반 SSD의 IOPS와 대역폭 성능을 측정하여 비교 분석하였다.

II. PCI-e 인터페이스를 이용한 DDR-1 DRAM기반 SSD의 설계

제안된 DDR-1 DRAM 기반 SSD는 기존 HDD기반의 스토리지 시스템과 달리 데이터 입출력 제어 부분과 데이터 제어 부분을 분리하여 실제 데이터 입출력 처리를 CPU가 수행하도록 하여 CPU 성능이 높아지면 높아질수록 성능이 향상되는 특징을 가지며 대역폭 확보를 위하여 외부로는 4Gbps급 FC(fiber Channel)와 10Gbps NIC 인터페이스를 채택하였고 최대 8-12개의 인터페이스를 제공하도록 설계하였다. 또한, DRAM의 휘발성을 해결하기 위하여 이중화 전력공급기와 배터리를 제공하며, 데이터 손실 방지를 위한 자체 백업기능을 추가하여 설계하였다. 스토리지 시스템의 운용관리를 위한 SRM(Storage Resource Management) 시스템을 별도로 제공하여 이기종 스토리지까지 관리할 수 있도록 하였으며 시스템의 용량은 단위시스템 당 최대 256GB까지 가능하며, 가상화 솔루션을 수십 TB까지 확장이 가능하도록 설계하였다.

본 논문에서 설계된 DDR-1 DRAM 기반 SSD는 고속의 읽기와 쓰기 속도를 처리하기 위하여 Intel 4-Core 2CPU를 사용한 상용 마더보드를 사용하였으며, 그림 1의 구조를 갖는다. 내부 버스는 PCI-e

Gen-II를 사용하였고 데이터 손실 방지를 위한 자체 백업매체는 HDD를 PCI-x 인터페이스로 연결되도록 하였으며 그림 1 우측의 M-SSD 1부터 M-SSD 8로 표현된 메모리 어레이의 디스크 인식을 위한 디바이스 드라이버는 SCSI 에뮬레이션 기능을 구현하여 기존 인터페이스와의 호환성을 갖도록 하였다. 사용된 OS는 Linux Open Source를 이용하여 범용성을 갖도록 설계하였다.

2-1 DDR1 DRAM SSD Memory

그림 1에서 DDR 1 DRAM SSD 메모리는 시스템 Backplane으로부터 연결된 PCI-e 버스의 고속 데이터 I/O 가 처리되는 부분이며, 시스템컨트롤 보드의 PCI-e 클럭과의 동기화가 이루어지도록 설계되었다. DDR1 메모리 모듈에 사용되는 PCI-e 커넥터 (Connector) 는 DDR1 메모리 모듈이 처리하는 데이터 전송 속도를 처리 할 수 있으며, 시스템의 전원공급 장치의 공급전원에 이상이 있을 시에도, DDR 1 메모리의 데이터를 보존하기 위해 메인칩셋과 DDR 1 메모리 모듈에 인가되는 전원의 입력은 이중화 구조로 설계되었다. SSD 드라이버는 DDR 메모리를 시스템에서 인식할 수 있는 SCSI 장비의 형태로 보여

주는 역할을 수행한다. 상위 계층의 시스템 OS에서는 디바이스 드라이버가 SSD를 SCSI 장비로 보여줌으로 기존의 SCSI 장비를 사용하는 인터페이스와 동일한 인터페이스를 사용할 수 있도록 하였다.

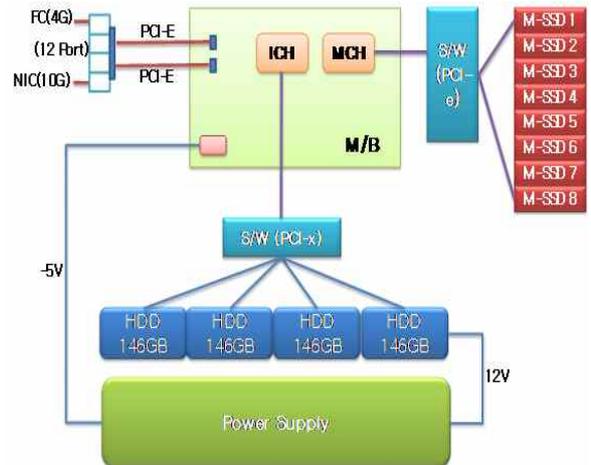


그림 1 DDR-1 DRAM SSD 구조도
Fig.1. Architecture of DDR-1 DRAM SSD

2-2 시스템 컨트롤러

DDR-1 DRAM SSD 시스템 컨트롤러는 시스템의 가장 상위에 위치하여 각 모듈을 직접 컨트롤 하는 기능을 수행하고, 관리 소프트웨어와 통신을 통해 시

● NGS PCI-e Signal Connection Block Diagram

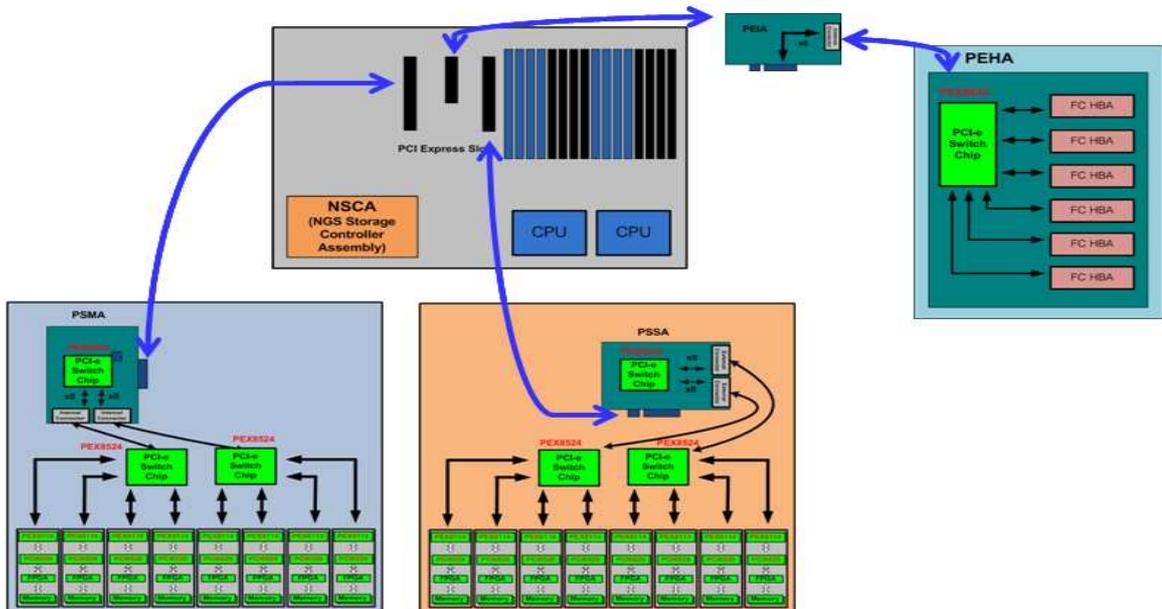


그림 2 PCI-e 신호 연결 도식도
Fig.2. Block diagram of the PCI-e signal connection

시스템을 관리/운영하는 중요한 부분이다. 본 논문의 시스템 컨트롤러는 시스템 컨트롤러 기능을 유지하기 위하여, 최대 2개의 이중 코어 인텔 Zeon 5100/5200시리즈 프로세서 또는 4중 코어 인텔 Zeon 5300/5400 시리즈 프로세서를 사용하고, 1066/1333/1600 MHz FSB를 지원 가능하고, 최대 128GB FBDIMM-533/667/800 메모리를 지원할 수 있으며, 2개의 64-bit/133MHz PCI-x 슬롯, 2개의 PCI-e x16 슬롯, 1개의 PCI-e x8 슬롯을 가지고 있다. 또한 DDR-1 DRAM SSD 시스템 컨트롤러는 인텔 82563을 채용하여 2 포트의 기가비트 이더넷을 지원하며, IDE, 플러피, 6개의 SATA2 포트 8개의 SAS포트를 갖도록 하여 DDR1 DRAM SSD 모듈의 데이터 대역폭과 IOPS 성능을 최대로 끌어낼 수 있도록 설계하였다.

2-3 Backplane 설계

백플레인(backplane)은 시스템컨트롤러와 DDR DRAM 메모리와의 연결을 위한 보드로서 시스템컨트롤러 보드의 PCI-e 신호는 PCI-e 신호 인터페이스 보드를 통하여 백플레인으로 연결 된다. 이때, 연결 방식은 pcisig 의 PCI-e 스펙에 준하며, 디스크 스토리지 모듈에서와 마찬가지로 클럭 신호의 동기화가 이루어져야 한다. 또한, PCI-e 의 스위칭에 의한 병목현상이 최소화 될 수 있도록 설계되었으며, DDR1 DRAM SSD 메모리 모듈로의 전원 공급이 원활히 이루어질 수 있도록 설계하였다. 그림 2는 PCI-e 신호의 연결을 보여준다.

2-4 DDR 메모리인식을 위한 SSD 드라이버

SSD 드라이버는 핵심 구성요소인 SSD를 동작시키는 디바이스 드라이버로서 DDR 메모리를 시스템에서 SCSI 에뮬레이션(Emulation) 을 통해 시스템이 인식할 수 있는 SCSI 장비의 형태로 보여 주는 역할을 수행한다. 상위 계층의 시스템 OS에서는 디바이스 드라이버가 SSD를 SCSI 장비로 보여줌으로 기존의 SCSI 장비를 사용하는 인터페이스와 동일한 인터페이스를 사용할 수 있도록 설계하였다.

III. 실험 및 고찰

3-1 실험시스템의 구성

HDD와 달리 새로운 동작구조를 갖는 DDR1 DRAM SSD 스토리지 의 성능을 확인하기 위하여 그림 3과 같은 스펙의 UNIX 및 Windows/Linux 서버, SAN Switch, Ethernet Switch를 이용한 실험시스템을 구성하였다.

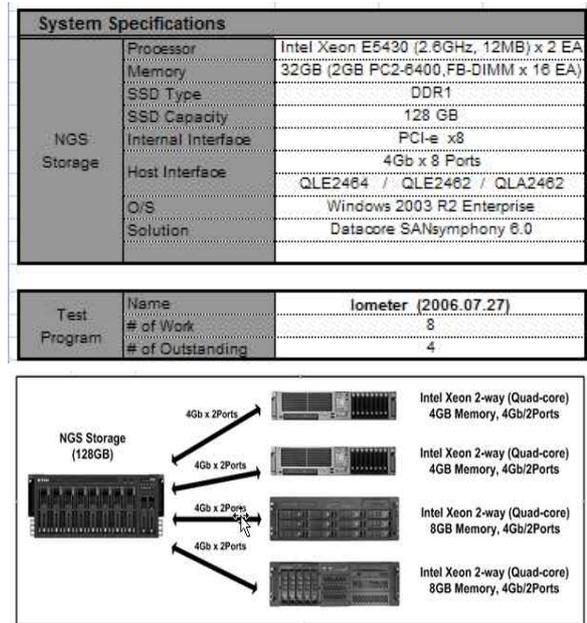


그림 3 실험시스템의 스펙
Fig. 3. Specification of the experimental system

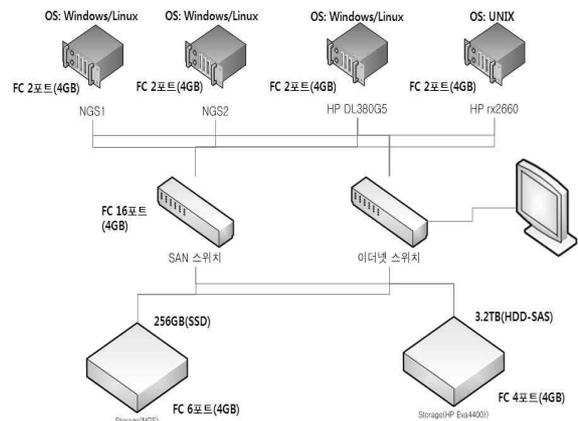


그림 4 실험시스템의 구성도
Fig. 4. Block diagram of the experimental system

측정틀로 선택한 IOmeter 는 시스템에 장치된 디

스크와 같은 저장장치의 성능을 측정하는 대표적인 툴로서 디스크 성능을 확인하기 위한 입출력성능 인 IOPS와 전송능력인 대역폭을 측정하는 프로그램이다[7][8].

그림 4는 그림 3의 장비들을 이용하여 구성된 실험시스템 구성도이다. 좌 상단에 NGS 로 표시된 것이 본 논문에서 설계 개발된 DDR-1 SSD 스토리지이다.

3-2 실험결과

그림 5는 전송되는 블록크기변화에 따른 제안된 DDR-1 DRAM SSD 시스템의 입출력 성능인 IOPS와 대역폭을 측정 한 결과이다. 그림 5의 상단은 DRAM SSD 그리고 하단은 HDD 의 IOPS와 대역폭을 측정 한 것이다. HDD의 IOPS가 50,000 인데 반하여 DRAM SSD의 IOPS 는 300,000 이상인 것을 볼 수 있다. 한편 블록크기에 따른 대역폭 또한 HDD는 1200MB/sec 인데 반하여 DRAM SSD 는 2700MB/sec 로 높은 것을 확인할 수 있다

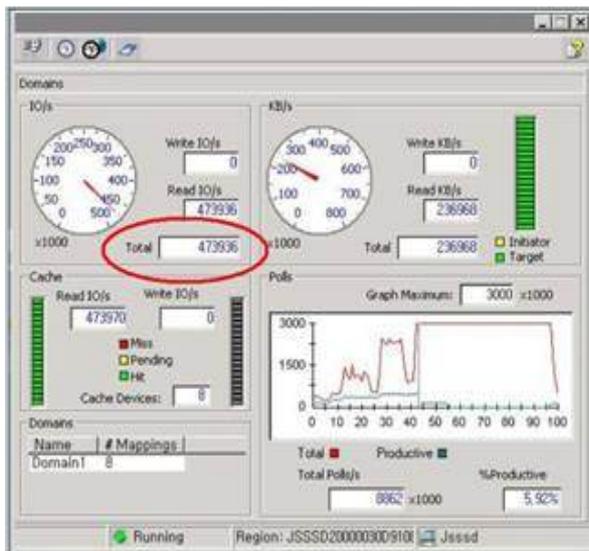


그림 5. 블록크기 변화에 따른 HDD, DDR 1 DRAM SSD의 대역폭 및 IOPS 측정

Fig. 5. IOPS and Bandwidth measurement results of HDD, DDR 1 DRAM SSD by variation block size

그림 5 실험결과에서 DDR 1 DRAM SSD 가 읽기 및 쓰기에서 대역폭이 높게 나타남을 확인할 수 있다. 플래시 SSD 의 경우 읽기속도는 상대적으로

HDD에 비하여 고속이지만 쓰기속도에서는 HDD와 유사한 저속으로 발표되고 있지만[10,12] 제안된 DDR-1 DRAM SSD는 읽기/쓰기 모두 고속으로 측정되었다. 표 1은 제안된 DDR-1 SSD 시스템의 성능을 정리한 것이다.

Bandwidth (Read)

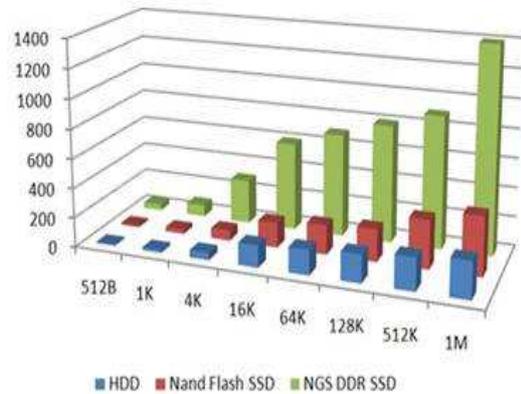


그림 6. Device별 Bandwidth(Read)측정결과
Fig 6. Each Device Bandwidth(Read) measurement result

Bandwidth (Write)

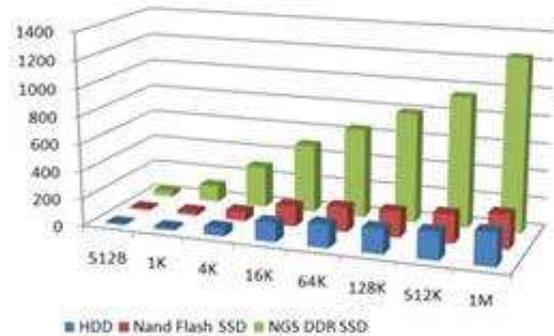


그림 7. Device별 Bandwidth(Write)측정결과
Fig 7. Each Device Bandwidth(Write) measurement result

IV. 결 론

본 논문에서는 전기적인 방식으로 데이터를 처리

하는 DDR 메모리를 데이터 저장매체로 사용하고 PCI-e를 이용하는 RAM 기반의 대용량 스토리지 장치를 설계하고 그 성능을 분석하였다. 설계된 SSD는 RAM을 주 저장 장치로 이용하고 PCI-e버스를 이용하여 256GB 이상의 대용량 스토리지를 구성할 수 있으며 고속의 데이터가 가능하였다.

실험을 위하여 UNIX 및 Windows/Linux 서버, SAN Switch, Ethernet Switch를 이용한 실험시스템을 구성하고 IOMeter 를 이용하여 데이터 블록 크기에 따른 랜덤 읽기/쓰기에 대한 성능을 측정하였다. 실험결과, DDR-1 DRAM SSD 는 HDD 에 비하여 매우 높은 IOPS 입출력 성능과 대역폭을 보였다.

현재 SSD의 가격이나 기술 성숙도를 볼 때, SSD를 모든 DB에 적용하는 것을 권하기에는 무리가 있다. 예를 들어 속도가 중요하지 않은 배치작업, 백업 같은 낮은 트랜잭션 DB는 이미 HDD와 RAID 시스템으로 지원되기 때문에 별 이슈가 아니다. 하지만 DB가 높은 트랜잭션 성능이 요구되고 정기적으로 튜닝을 요구하고 있다면 SSD는 아주 적절한 솔루션이 될 것이다. SSD의 랜덤 읽기/쓰기 성능이 좋은 특징을 고려한다면 SSD는 DB가 고 트랜잭션 율(high transaction rate), 결정되지 않았거나 변화하고 있는 DB 사용 패턴, 수용 불가능한 쿼리 지연시간이 있을 때, 순간적인 피크로드가 걸리는 작업(high peak load), 실시간 DB, 24시간 온라인 동작하는 DB, 유지 보수 시간이 제한된 DB에 적합할 것이다.

참 고 문 헌

- [1] 정승국, 고대식, “차세대 스토리지 SSD 기술 동향”, *IITA, 주간기술동향 1369호*, 2008.10.22
- [2] Solid Data systems, "Impact of Solid-state disk on high-transaction rate databases", *Solid data systems, Inc. White paper*, 2005 Feb.
- [3] Greg Schulz, "Achieving Energy Efficiency using FLASH SSD", *The Sotrage IO group white paper*, 2007.Dec.
- [4] Dae-Sik Ko, Seung-Kook Cheong, “Web Performance Enhancement of E-business System using the SSD”, *UNESST2008*

- [5] Jamon Bowen, "Flash in the Enterprize", *TMS white paper*, 2007. Sep.
- [6] TMS, "Increase Application Performance with Solid State Disks", *TMS white paper*, 2008.Feb
- [7] Solid Data systems, "Comparison of Drives Technologies for High-Transaction Databases", *Solid Data systems, Inc. White paper*, 2007 August.
- [8] Solid Data systems, "Impact of Solid-state disk on high-transaction rate databases", *Solid data systems, Inc. White paper*, 2005 Feb
- [9] TMS, "Increase Application Performance with Solid State Disks", *TMS white paper*, 2008.Feb
- [10] Greg Schulz, "Achieving Energy Efficiency using FLASH SSD", *The Storage IO group white paper*, 2007.Dec.
- [11] Solid Data systems, "Comparison of Drives Technologies for High-Transaction Databases", *Solid Data systems, Inc. White paper*, 2007 August.

고 대 식 (高大植)



1991.2: 경희대학교 공학박사
 1989 - 현재: 목원대학교 전자공학과 교수
 2001 - 2003 : 목원대학 학술정보처장
 2008 - 현재: 한국정보기술학회 수석부회장
 2009 - 현재 : KSN 포럼 부의장
 관심분야 : 멀티미디어 통신, Healthcare

정 승 국 (鄭承國)



2004년 2월 : 한남대학교 전자정보통신 공학과(박사)
 1985년 ~ 현재 : 한국전자통신연구원 (책임연구원)
 관심분야 : Grid Computing, Utility Computing, Solid State Disk, Storage&Server Virtualization