

스토리지 관련 기술의 새로운 동향

김성운*, 김보관**

*한국전자통신연구원, **충남대학교 전자공학과

I. 서 론

지난 수년간 네트워크 대역폭의 빠른 증가와 컴퓨터 성능의 향상으로 인터넷이 급속히 보급되었다. 이러한 인터넷 환경의 변화로 네트워크를 통한 멀티미디어 데이터 송수신이 급증하게 되었고, 이에 따른 대용량 스토리지 시스템에 대한 수요가 계속 늘어나게 되었다. 최근 들어 전세계 경제가 하락세를 보이는 상황에서도 오히려 스토리지 시스템은 기존의 저장영역뿐 아니라 e-Commerce, 재난 복구 등 기업의 새로운 어플리케이션 영역 확장으로 수요가 상승하고 있다. IDC에 따르면 2006년까지 디스크 시스템 스토리지 시장은 55.3%라는 엄청난 연평균 성장률을 기록할 것으로 전망하고 있다. 이 같은 상황에서 HP, SUN, IBM 등 대표적인 서버 제조회사도 전략 사업 영역을 스토리지 분야로 옮기고 있다.

이러한 IT 환경의 변화로 인한 스토리지 시장의 확대와 경쟁은 스토리지 분야에서의 새로운 기술을 요구하고 있다. 현재 새롭게 떠오르는 많은 스토리지 관련 기술들은 시스템 성능 향상, 시스템 운영의 편리성, 그리고 저장 용량의 확장성을 약속한다.

본 논문은 스토리지 시스템에서 떠오르고 있는 다양한 기술들 가운데서 IP 스토리지, InfiniBand 연결망, DAFS 프로토콜과 최근에 많은 관심을 모으고 있는 OSD에 대하여 알아 본다.

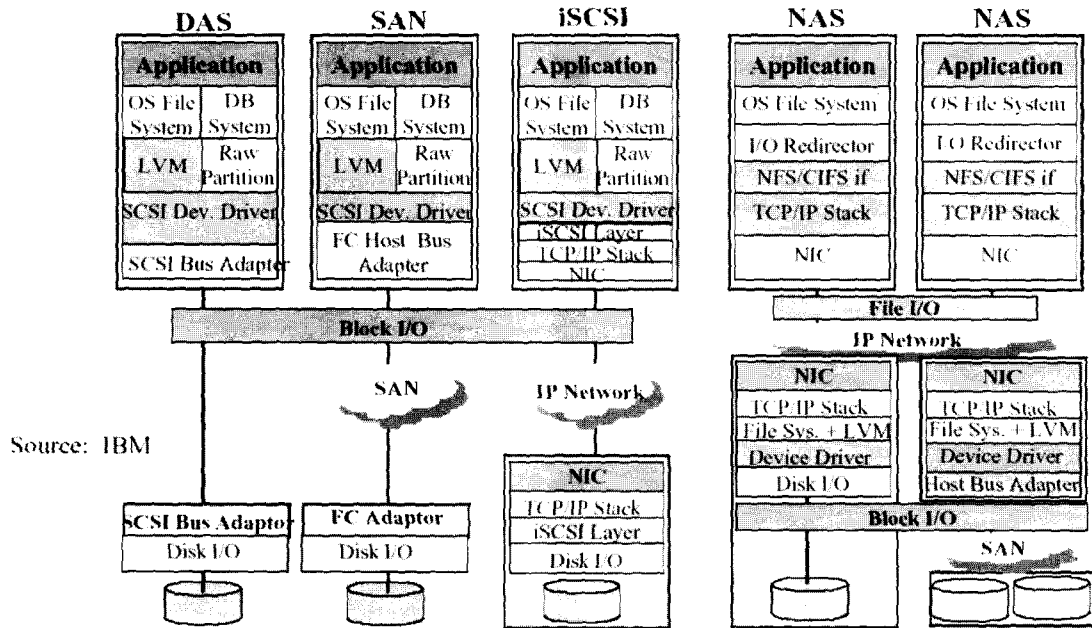
II. IP 스토리지

데이터 집중형 어플리케이션이 늘어남에 따라 저장해야 할 데이터가 기하 급수적으로 증가하게 되었다. 급증하는 데이터의 처리와 관리를 위하여 많은 사용자들은 독립적인 SAN(Storage Area Network) 환경을 구축하였다.

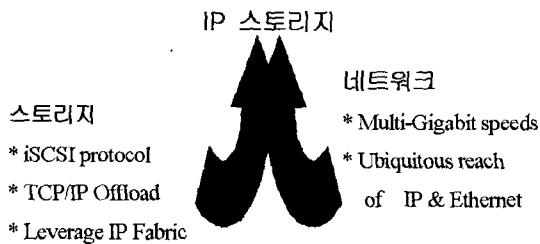
SAN은 고유의 연결망을 가짐으로써 스토리지 트래픽 만을 단독으로 처리할 수 있기 때문에 높은 신뢰성을 제공할 수 있을 뿐만 아니라, 블록 전송 프로토콜을 사용한 파이버 채널로 데이터를 전송하여 I/O 집중적인 스토리지 어플리케이션에 적합하다. 파이버 채널을 이용한 SAN(이하 FC-SAN)은 장점을 가지고 있는 반면에 여러 장치들 간의 상호 접속성 문제와 비용 문제가 있다.

<그림 1>은 DAS, FC-SAN, IP 스토리지, NAS의 계층별 모습을 표시하고 있다. 어플리케이션 계층에서 NAS는 파일 I/O를 가지는 반면에 FC-SAN과 IP 스토리지는 블록 단위로 데이터가 전송된다. 최근 들어 FC-SAN의 문제를 극복하기 위하여 IP 네트워크에서 SAN을 구축하려는 움직임이 나타나고 있다. 이러한 스토리지와 IP 네트워크의 통합에는 iSCSI (Internet Small Compute System Interface) 프로토콜이 사용된다.^[1]

이러한 IP 네트워크를 이용한 SAN을 IP-SAN이라 하고 IP-SAN에 대응한 스토리지를 IP 스토리지라고 한다. IP 스토리지는 <그림 2>와 같이 스토리지와 네트워크의 강점을 끌고루



<그림 1> 여러 스토리지 시스템들의 내부 구성



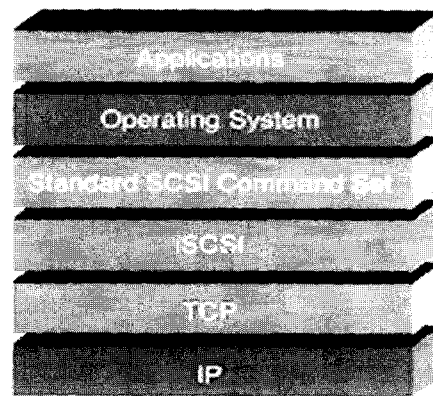
<그림 2> IP 스토리지의 주요 요소들

갖추고 있고, iSCSI 프로토콜이 실리콘 기반 TOE(TCP/IP Offload Engine)를 이용하여 구현되기 때문에 서버의 성능을 감소시키지 않고도 효율적으로 IP 스토리지를 구현할 수 있다.

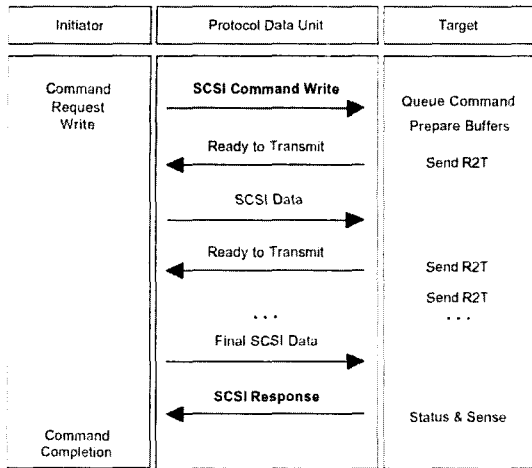
IP 스토리지의 가장 큰 장점은 이더넷을 사용하는 것이다. 이더넷은 기업 내에서 오랫동안 사용하고 있는 환경이므로 이기종들 간의 상호 접속성이 좋고, 가격 면에서 유리할 뿐 아니라, IP 스토리지의 대응 접속이 TCP/IP 레벨에서 이루어지기 때문에 기기의 가격 면에서도 규모 확대에 의한 이점을 살릴 수 있다. 또한, IP-SAN 환경에서는 이론상 거리의 제약이 없다. 데이터 센터 내부가 IP 네트워크로 일원화 된다. 클라이언

트-서버간, 서버-스토리지간(SAN) 모두 IP 네트워크를 이용할 수 있다. 데이터베이스나 파일 시스템 등과의 호환성이 유지 된다. FC-SAN과 마찬가지로, 서버 직결 스토리지와 같은 SCSI 커맨드를 사용할 수 있다. 따라서, IP-SAN은 FC-SAN을 보완하는 기술로서 현재 가장 주목을 받고 있다.

iSCSI는 Fibre Channel 기반의 SAN을 대체하기 위해 Cisco, IBM이 주도로 개발된 TCP/



<그림 3> iSCSI 프로토콜 스택

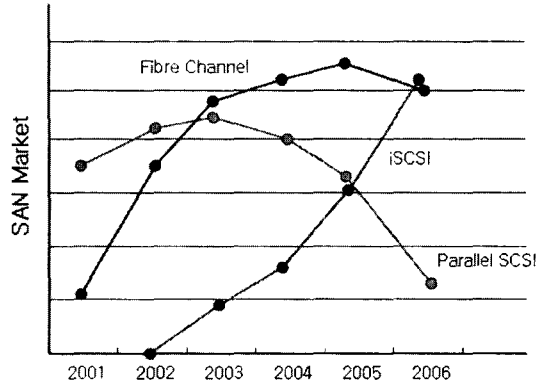


〈그림 4〉 iSCSI 쓰기 예

IP 기반의 블록 레벨 데이터 전송 스토리지 프로토콜이다(〈그림 3〉 참조). SCSI 명령을 TCP/IP 패킷으로 Encapsulation 하여 일반적인 IP 네트워크로 SAN을 구축할 수 있다(〈그림 4〉 참조).

과도기적인 iSCSI는 문제점이 없는 것은 아니다. 우선 기존의 FC-SAN에서 사용하는 프로토콜을 10G 이더넷 인터페이스에서 사용가능 하도록 해야 한다. 또한 고객들이 기존 스토리지 방식을 바꾸지 않고도 10G 네트워크에 적용할 수 있어야 하며, 10G 인터페이스에서 그냥 NAS를 사용하는 것에 비해 효율이 높아야 한다.

2006년 FC-SAN 시장이 24억 달러로 예상하는데 비해 IP 스토리지는 10억 달러의 시장이 예측되고 있다. 비교적 저렴한 가격대를 형성할 iSCSI 서버는 2006년 140만대를 육박해서 FC-SAN 보다 2배 이상 많이 팔릴 것으로 예상된다. 초기 비용 때문에 SAN 전환을 꺼리는 중소기업을 대상으로 IP 스토리지 시장은 형성될 것으로 보이며 초기에 MS 윈도우 기반의 스토리지 시스템에 큰 영향을 미칠 것으로 보인다. 따라서 iSCSI 장비들은 스토리지 관리자를 따로 두기 힘든 중소 규모의 기업을 위한 보급형 장비로 사용될 것이다. 즉, IP 스토리지는 기존의 SAN 시장을 위협하기 보다는 기존의 DAS 환경을 스토리지 네트워크로 인도하는 역할을 수행할 것으로



〈그림 5〉 스토리지 예상 시장

보인다.

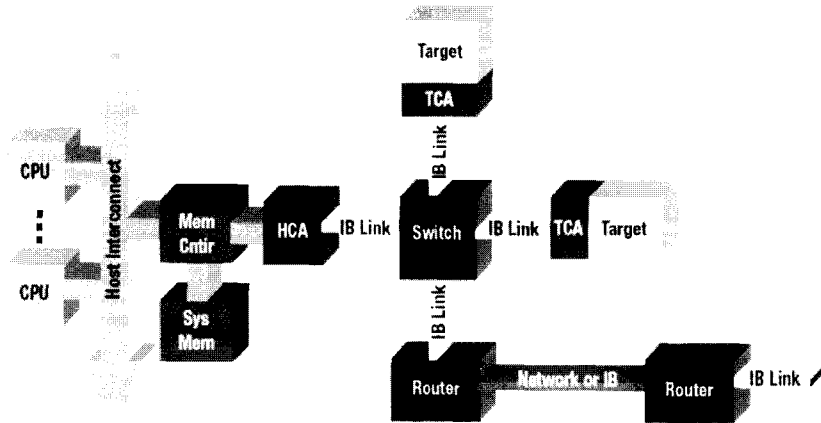
SAN 전송률이 10Gbps에 도달할 때까지 iSCSI와 Fibre 채널은 같이 공존할 것이지만, 10Gbps에 들어서면, IP 기반 솔루션이 네트워크 데이터 스토리지를 지배할 것으로 보인다.

II. InfiniBand 연결망

온라인 상의 정보의 양이 기하 급수적으로 증가함에 따라 이러한 정보를 처리하기 위해서는 연산능력 중심으로 하는 컴퓨팅 서버에서 연산 능력은 물론 입출력 처리 능력에서도 높은 성능을 요구하고 있다.

I/O는 컴퓨터 시스템의 중요한 자원인 동시에 전통적인 시스템 구조에서 병목으로 지적되고 있는 부분이다. 이는 CPU와 메모리 작동 속도에 비해 I/O 부분의 대역폭이 상대적으로 작기 때문이다. 따라서, PCI(Peripheral Components Interconnect) 버스^[8]로 대표되는 I/O 부분에 변화를 요구하게 되었다. 즉, 버스의 구조적인 단점인 고가용도와 확장성의 한계로 고성능 컴퓨터 시스템을 구현하는데 있어서 성능 한계를 맞게 되었다.

InfiniBand는 무한(Infinite)과 대역폭(Bandwidth)의 합성어로 차세대 서버 연결망을 위한 업계 표준으로, 기업의 전산망을 구성하는 서버,



Source: InfiniBand Trade Association

〈그림 6〉 InfiniBand 구조 모델

스토리지, 네트워크 장비 등을 점대점 (Point-to-Point) 방식의 스위치 기반의 시리얼(serial) 네트워크로 구성함으로써 고성능 연결망을 통한 동시전송을 지원하고, 입출력장치와 프로세서간의 처리속도 차이에 따른 데이터 병목현상을 제거하기 위해 제안된 연결망 기술이다.

인텔과 썬을 주축으로 하는 NGIO (Next Generation IO) 그룹과 HP (Hewlett Packard)와 IBM, Compaq을 주축으로 하는 Future IO 그룹이 차세대 연결망을 목표로 구성되었다. 양 진영은 모두 스위치 기반형 점대점 연결망이라는 유사한 연결망 구조와 이를 이용한 서버 구조를 검토, 제안하고 있었으며 지난 1999년 8월 두 진영이 통합하면서 InfiniBand라는 차세대 서버 연결망 구조를 제안하였다.^{[2][4]} 현재 InfiniBand는 표준 규격서가 발표되었으며 2002년 부터 시제품이 나왔으며, 올해는 상용화 제품들이 출시되고 있다.

서버의 내부 연결망(PCI 버스)에서 발생하는 병목 현상과 공유 버스 구조의 낮은 확장성 및 가용성을 극복하기 위하여 InfiniBand는 다음과 3가지 구조적 해결책을 제시하고 있다^[2, 3, 4, 5]

〈그림 6〉은 InfiniBand의 구조를 간략히 모델화하여 나타낸 것이다. InfiniBand 연결망은 스위치와 라우터로 구성된다. InfiniBand 연결망은 글로벌넷과 서버넷 두 단계로 구성되며, 각

〈표 1〉 PCI 버스와 InfiniBand의 비교

| | PCI 2.2 | PCI-X 1.0 | InfiniBand |
|-------------|----------------|------------------|----------------|
| 동작속도 | 66MHz | 133MHz/ 1slot | 2.5Gbps |
| 데이터 신호 폭 | 32/64비트 | 32/64 비트 | 1x, 4x, 12x |
| 최대 대역폭 | 533 Mbyte/s | 1064 Mbyte/s | 2.5Gbps/ 포트 |
| 연결 방식 | 공유 버스 | 공유 버스 | 스위치 기반 점대점 |
| 최대 연결 노드 | 10개 이하 | 1개/ 133MHz | 216개/ 서브넷 |
| 통신방식 | 공유메모 리 매핑 | 공유메모 리 매핑 | 채널기반 메시지 패싱 |

단계에 따라서 경로 설정에 사용되는 식별자를 구분하고 있다. 스위치는 서버넷 연결망에 연결된 각 노드 간의 경로 설정과 연결망을 통해 패킷 데이터 전송을 담당하고, 라우터는 글로벌넷 연결망을 구성하는 서버넷 연결망간의 경로 설정을 담당한다. 링크 속도가 2.5~30Gbps인 InfiniBand는 이더넷 및 파이버 채널 인프라와의 효율적인 통합을 가능하게 한다. InfiniBand의 높은 대역폭과 낮은 지연시간 패브릭으로 애플리케이션 성능이 크게 개선할 수 있다.

InfiniBand는 채널 기반형 통신 방식을 사용하며, 채널은 중앙처리 장치와 입출력 장치 사이에 구성된 독립적인 메시지 패싱 인터페이스로 주기억장치와 입출력 장치간의 직접적인 데이터 전송을 보장하여 공유 메모리 매핑 방식에서 발생하는 데이터 복제 오버헤드를 제거하고, 각 채널은 독립적으로 유지 관리된다. 채널 기반형 메시지 패싱은 중형 클러스터 컴퓨터에서 사용되던 VIA (Virtual Interface Architecture)^[7]를 InfiniBand에 적용한 것으로 중앙처리 장치 상호간 또는 중앙처리 장치와 입출력 장치간에 독립적인 사용자 수준 메시지 패싱 인터페이스를 제공하여 운영체제의 간섭 없이 통신을 가능하게 한다. 채널은 점대점 연결망 사용을 위한 사용자 프로그래머블 레지스터 레벨 인터페이스로 정의할 수 있으며, 채널을 제어하는 하드웨어는 DMA (Direct Memory Access) 엔진을 사용하여 주기억 장치와 통신하고자 하는 상대방과 직접 데이터를 송수신할 수 있게 한다. 따라서 중앙처리 장치는 입출력 장치 제어 및 접근과 데이터 전송에 관련된 작업 부하를 줄임으로서 시스템 성능을 향상시킬 수 있다^[3, 4, 5].

InfiniBand는 PCI 버스와 달리 고속 직렬 전송 방식을 사용한다. InfiniBand 연결망 구조가 스위치 기반형 점대점 구조를 가짐에 따라 스위치를 포함하여 연결망 구성에 있어서 비용을 절감하고 신호 전달 속도를 고속화 하기 위해서이다. 현재 2.5Gbps급 직렬 전송을 기본으로 4개, 12개의 직렬 전송로를 하나의 포트에 사용하는 것도 가능하다.

InfiniBand는 스위치 기반형 점대점 연결 구조를 사용한다. 스위치 기반형 점대점 연결 구조는 공유 버스 구조의 가장 큰 단점인 확장성과 고가용성 보장을 위해서이다. 스위치 기반의 점대점 연결 구조를 가짐으로서 연결 포트의 확장 및 제거가 자유로우며, 특히 공유 버스 구조가 내포한 결함 가능성을 제거할 수 있다. 또한 점대점 연결 방식을 사용하므로 hot-plug 기능을 하드웨어적으로 완벽하게 지원할 수 있으며, 신규 장치의 장착 및 오류 장치의 분리를 손쉽게 구현할

수 있다.

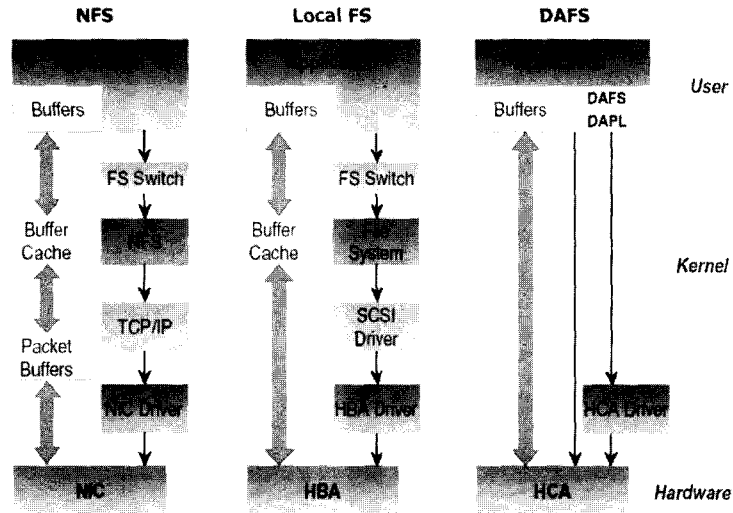
InfiniBand는 서버 시장의 변화뿐 아니라 스토리지 시장에서도 변화를 일으키고 있다. 인터넷의 폭발적인 성장과 온라인 상의 정보유지의 필요성으로 대용량 스토리지에 대한 요구로 인해, 기존의 서버와 외장형 스토리지를 전용선으로 연결하는 DAS에서, LAN을 통해 서버와 스토리지를 접속하는 NAS, 그리고 새로운 엔터프라이즈 환경을 겨냥해 스토리지에 대한 고속 액세스, 일원적인 관리, 그리고 확장성과 고장 감내성을 높이기 위한 SAN이 스토리지 시스템의 주류로 등장하고 있기 때문이다.

따라서, 스토리지 시스템에서도 Infiniband 기술을 활용한 InfiniBand 스토리지가 점점 증가할 것으로 전망하고 있는데, Yankee Group은 2006년 전세계 스토리지 시스템 시장에서 InfiniBand 스토리지 시스템 시장이 차지하는 비중이 2.47%에 이를 것으로 전망하고 있다^[8].

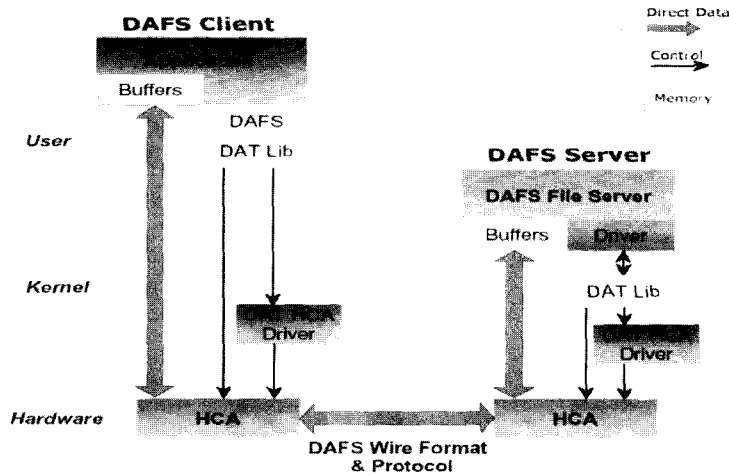
III. DAFS(Direct Access File System)

연결망의 급속한 발전은 스토리지 기술 분야에서도 상당한 변화를 예고하고 있다. RAID 기술이 기존 스토리지의 물리적인 한계를 극복한 반면, iSCSI는 기존 SCSI 기술을 IP망에서 사용할 수 있도록 개발이 진행되고 있다. iSCSI는 기가비트 네트워크 상에서 사용할 수 있도록 개발을 하고 있다. 하지만 본격적인 10GB 시대가 도래하면 iSCSI는 도태할 것이라는 전망도 조심스럽게 개진되고 있다. 네트워크 프로토콜 위에 SCSI라는 추가 오버헤드를 안고 있으며, 10GB의 충분한 대역폭이 이러한 오버헤드 없이 바로 IP 전송으로도 성능이 구현될 수 있다고 보기 때문이다. 이러한 상황에 DAFS(Direct Access File System)의 등장으로 현재의 스토리지 네트워크의 성능 및 관리 수준을 한 단계 올릴 수 있을 것으로 기대되고 있기 때문이다^[9].

DAFS는 InfiniBand, VI, iWARP와 같은



<그림 7> NFS, System, DAFS 접근 모델



<그림 8> DAFS Client/Server 간의 데이터 전송

연결망 기술상에서 현재 떠오르고 있는 RDMA (Remote Direct Memory Access)의 이점을 얻기 위해 설계된 새로운 파일 접근 프로토콜이다. 이 프로토콜은 IDC 환경의 Web Farm, E-commerce, 데이터베이스 응용에서 성능, 안정성, 확장성을 현격히 증가시킨다. <그림 7>은 NFS, 지역 파일 시스템, DAFS의 차이를 간략히 나타낸 것이다. DAFS는 그림과 같이 기존의 파일 시스템에서 진행되는 여러 단계의 파일 처

리 과정을 3~4단계의 단순한 과정으로 파일 처리를 가능하게 하는 프로토콜이다. 직접 파일 접근 방법의 이점은 사용자와 커널 간의 데이터 복사 일어나지 않는 것이다. 즉 데이터는 응용 프로그램의 버퍼로 직접 전송이 된다. 또한 데이터 패킷의 Fragmentation이나 Reassembly가 필요 없다. 이것은 IP Jumbo 프레임과 유사한 이점을 가지는데, 적은 전송 오버헤드로 적은 인터럽트를 발생시킨다. 프로토콜 헤더와 데이터 버

퍼는 각각 나뉘어 전송된다(<그림 8> 참조).

DAFS를 구현하기 위한 기술력 있는 회사들의 협의체(DAFS Collaborator)는 85개사로 구성되어 있으며 이 구성원들 중에는 IBM, HP, 썬, 오라클 등 시장 지배력을 갖고 있는 유수의 업체들이 대거 참여하고 있다. 2001년 9월 DAFS Specification Version 1.0이 이미 완성되어 IETF(Internet Engineering Task Force)에 제출되었으며 DAFS API Specification Version 1.0은 2001년 11월 완성되었다. DAFS Collaborators는 크게 두 가지 분야에 초점을 두고 진행되고 있는데 그 중 하나는 DAFS라는 프로토콜 자체를 만드는 분야이며, 나머지 하나는 RDMA(Remote Direct Memory Access) 기술을 구현하고 그 기술을 DAFS에 접목시키고자 하는 분야이다.

현재 DAFS 프로토콜을 만드는 분야는 SNIA(Storage Networking Industry Association)의 한 부문으로써 DAFS Implementers Forum 그룹을 중심으로 그 노력이 경주되고 있다.^[10]

이 포럼에 참가하고 있는 업체들이 초창기부터 구현하고자 했던 것은 dDAFS(driver DAFS)였다. 그러나 진정한 DAFS의 실현은 uDAFS(user DAFS)의 완성이라고 할 수 있다. uDAFS 방식은 응용 프로그램은 파일 접근을 위해 uDAFS 라이브러리를 호출한다. 이 방식은 획기적인 성능 향상을 이룬다. 또한, 어떠한 OS 플랫폼에서도 비동기 파일 I/O를 실현할 수 있다. 사용자들은 DAFS의 특징인 OS에 독립적인 API의 이점을 가지기 위해 응용 프로그램들은 변경하여야 한다. Fencing이나 locking 기법을 사용하여 지역 파일을 공유할 수 있다.

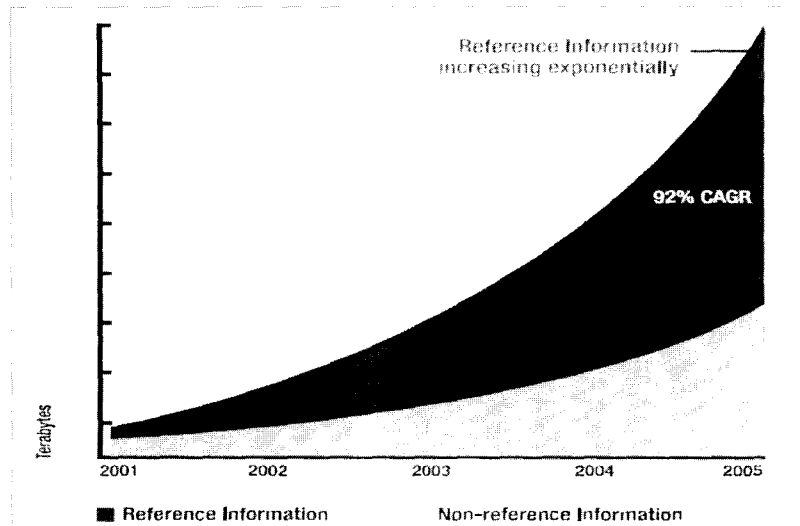
fDAFS는 IFS/VFS처럼 구현되는 DAFS이다. 이 방식은 파일 I/O 인터페이스를 가지고 있어 기존의 응용 프로그램을 변경하지 않아도 된다. 그러나, 커널내의 변환 부담을 제거하는 장점을 살리지 못하고, OS 구조에 한정되는 비동기 I/O 기능만을 가진다. 또한 지역 파일을 공유할 수 없다.

IV. OSD(Object Based Storage Devices)

인터넷의 발전에 따라 e-Commerce, 네트워크 기반의 응용들에 의하여 생성되는 데이터의 양이 계속 증가하고 있다. 이러한 상황에서 기업들은 더욱 경쟁력을 증가시키고, 기업의 공동사업자, 공급자들과 더욱 효과적으로 일하고, 피고용자의 생산성을 향상시키고, 기업의 고객들에게 더욱 양질의 서비스를 제공하기 위하여 현재 급격하게 축적되고 있는 데이터를 이용하는 것을 시도하고 있다. 따라서, 기업은 처리되고 저장되어야 할 정보의 홍수 속에 직면하게 되었고, ESG(Enterprise Storage Group)은 이러한 정보를 Reference Information(RI)이라 부른다.

RI는 ESG에 의하여 “digital asset retained for active reference and value”(활발한 참조와 가치를 위하여 유지되는 디지털 자산)라고 정의되었다. RI는 다음과 같은 것을 포함하지만, 그들에 국한되지는 않는다. 계약서, 전자우편, 전자우편 첨부물, 발표자료, CAD/CAM 설계, 소스 코드, 웹 콘텐츠와 같은 전자문서, 그리고 음성 데이터와 모든 종류의 영상자료. 영상자료들은 생물 정보학에서 발생한 검사 영상들, 비디오들, 역사적인 문서들, 의료 영상들, 사진들을 포함한다. ESG는 RI의 양이 모든 다른 형태의 정보보다 빨리 증가하고 있으며, 2004년까지는 모든 다른 형태의 정보의 양을 초과할 것이라고 주장한다.(<그림 9> 참조) 게다가 저장된 정보는 전략적으로 중요하기 때문에 백업 혹은 재난 시 복구 기능은 필수적인 사항이나, 이러한 기능들을 복잡한 IT환경에서 구현하기는 어려운 상황이다. 기업들은 그들은 가치있는 데이터들의 보안에 대하여 관심을 갖고 있으며, 현재의 저장 아키텍처들은 보안 문제들을 야기할 가능성이 있다는 것을 인식하고 있다.

2002년 ESG 연구결과는 단지 37%의 기업 IT 조직 저장장치들이 RI를 위하여 사용되었다는 것을 나타낸다. 그러나, ESG는 2005년까지 그 비율은 54%로 증가할 것이라고 예측했다.



〈그림 9〉 RI 소스의 변천

ESG는 전통적인 디스크 저장 방법은 RI만의 요구사항을 만족시키기 위하여 설계되지 않았다고 주장한다. 최적화된 솔루션은 정보가 저장되어 있는 매체로부터 정보를 추상화하여 쉽게 검색이 가능하고 매우 안전하여야 할 것이다. 즉, 이상적인 미래 저장 기술은 데이터 공유와 동시에 보안의 문제를 해결해야 할 것이다.

이러한 문제 해결을 위해서, 저장 객체와 OSD (Object-based Storage Devices)의 사용은 기업의 급격하게 증가하는 저장 요구를 처리하는 하나의 방법으로 아주 유망한 것처럼 보인다.

OSD 기술은 확장가능하고, 고성능, 플랫폼에 의존하지 않고 안전한 데이터 공유 아키텍처를 가능하게 한다. 이러한 아키텍처에 기반을 둔 제품들은 고속도와 직접 접근 가능한 SAN의 장점과 데이터 공유와 보안성을 특징으로 하는 NAS의 장점들을 결합할 것이다.

연구 프로젝트는 고성능이면서, 플랫폼 독립적이고, 새로운 저장 집중 컴퓨팅 가능성을 달성하기 위한 방법들을 조사하고 있다. 이러한 연구는 여러 해 동안 이루어지고 있는 상당한 양의 계속되는 저장 시스템 연구에 기반을 두고 있다. 카네기멜론 대학의 Parallel Data Lab은 객체 기반 저장 모델 연구에 관한 선도 제창자들 중에 하나

이다.

이미 객체기술을 이용하는 오픈 소스 고성능 파일시스템을 개발하는 Lustre 프로젝트는 진행 중에 있다. 이 프로젝트는 Linux, XML(Extensible Markup Language), LDAP(Lightweight Directory Access Protocol), 오픈 소스 라이브러리, 현존하는 파일 시스템 등과 같은 오픈 표준을 최대한 이용하고 있다. 이 파일 시스템은 GNU GPL(General Public License) 아래에서 오픈 소스 소프트웨어로 라이선스되어 이용 가능하다. 개념의 증명이상으로, Lustre는 객체들과 스토리지가 서로를 위해 만들어진 것처럼 보인다는 사실을 강조하는 하나의 기능 시스템이다.

진행되고 있는 발전에 근간하여 새로운 응용과 제품이 가까운 미래에 시장에 출현할지도 모른다. 장기적인 안목에서 OSD 기술은 지능적인 처리와 새로운 형태의 저장 응용을 선도할 것으로 보인다.

전문가들은 객체는 파일과 블록 두 가지 기술의 수렴으로 간주될 있다고 기술하고 있다. 파일은 OS 플랫폼에 독립적으로 안전한 데이터 공유를 가능하게 해주는 상위 계층의 저장 장치 추상화를 응용프로그램에게 제공한다. 그러나, 때때로

파일 서버를 사용하기 위한 경쟁으로 제한된 성능을 나타낸다. 블록은 공유된 데이터에 신속하고 확장가능성이 있는 접근을 제공한다. 그러나, I/O를 인증하고, 메타데이터를 유지하는 파일 서버 없이는 블록의 직접적인 접근은 제한된 보안과 데이터 공유라는 문제를 낳게 된다.

객체는 파일과 블록 모두에게서 가장 좋은 특징만을 추려낸 결과이다. 블록과 같이 객체는 프리미티브이다. 즉, 서버가 필요하지 않고, 스토리지 디바이스 위에서 직접 접근될 수 있는 스토리지의 논리적인 단위이다. 특징, 보안 키, 허기를 포함하는 데이터에 관련된 정보와 함께 객체를 직접 접근하는데 필요한 정보를 제공하는 것이 메타데이터이다. 직접적인 접근은 블록에서와 같이 성능 면에서 좋은 점을 제공하고 파일과 같은 추상화는 데이터가 플랫폼에 관계없이 공유될 수 있는 것이 가능하도록 한다. 이런 것들이 객체 기반 스토리지의 장점들이다.

하나의 객체는 가변 길이의 특성이 있기 때문에 객체는 어떤 종류의 응용에 관한 데이터라도 저장할 수 있다. 예를 들면, 멀티미디어, 의료 영상, 데이터 베이스 기록, 3D 그래픽스 등의 것들이 포함된다. 사실 어떤 하나의 객체는 어떤 파일 시스템 전체 혹은 데이터 베이스를 포함할 수 있다. 객체에 무엇을 저장할 것인가에 관해서는 실

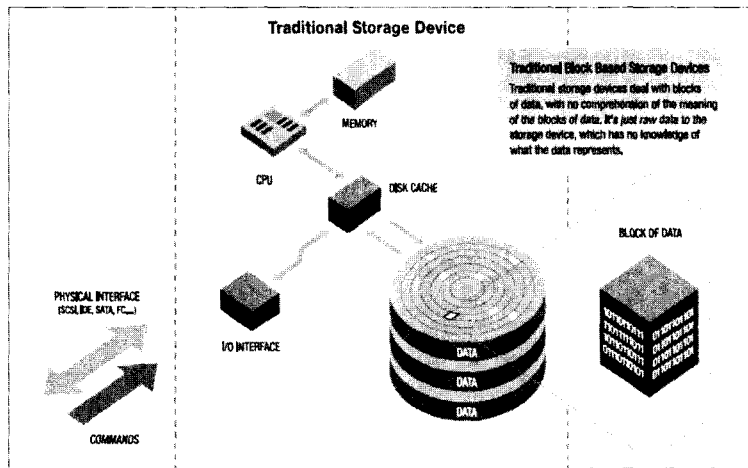
제적으로 응용프로그램에 달려있다. 객체는 역동적으로 크기가 증가하거나 감소할 수 있기 때문에 스토리지 디바이스 즉 OSD는 모든 공간 관리에 책임이 있다.

<그림 10>의 예에서 보면 각 객체는 어떤 데이터베이스에서 하나의 레코드를 나타낸다. OSD에 보내진 조회가 balance amount, due date, balance due에 따라서 검색되도록 OSD는 레코드를 정리하여 저장할 수 있다. 예를 들어, 오늘까지 기한인 모든 계좌를 얻기 위한 명령어가 응용 프로그램으로부터 왔다면, 이러한 조회요구는 OSD가 자체적으로 처리할 수 있다. 전통적인 스토리지 시스템에서는 응용 프로그램이 모든 레코드를 요구해서 필요한 정보를 검색해야만 할 것이다.

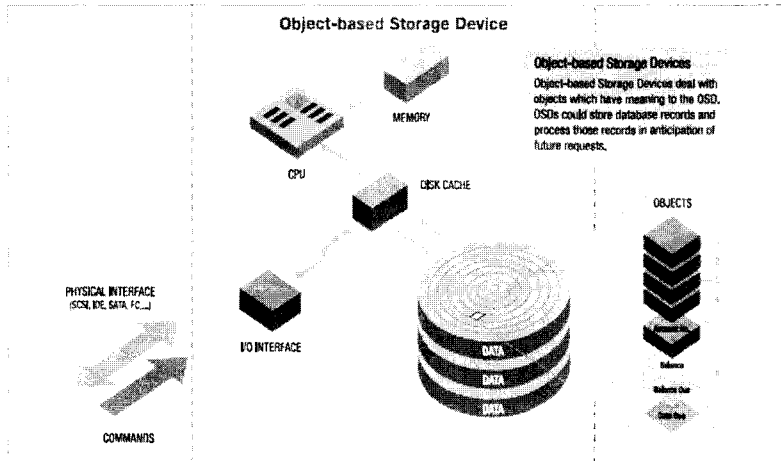
객체는 NAS와 SAN 아키텍처의 장점만을 취하고 두 기술의 사용에 있어서 단점이 되는 trade-off들을 제거한다.

객체는 NAS의 상위 계층 파일 추상화를 가능하게 하고, 이것은 정책 기반의 보안과 플랫폼 독립적인 데이터 공유를 가능하게 한다. 객체는 SAN의 고속도 직접 접근과 스위치 fabric에 의해서 제공되는 확장성을 최대한 이용할 수 있도록 한다.

OSD는 하나의 디스크드라이브로부터 드라이



<그림 10> 전통적인 블록 기반 스토리지 디바이스



〈그림 11〉 객체 기반 스토리지 디바이스

브 배열까지 다양한 형태를 취하면서 다양한 기능을 제공한다. 랜덤 접근 혹은 균일한 쓰기 디바이스 이외에도 테이프 드라이브와 광 미디어도 객체를 저장하기 위하여 사용될 수 있다. OSD와 전통적인 블록 기반 디바이스와의 큰 차이점은 물리적인 미디어가 아니라 OSD가 이해하는 명령어들이다. OSD는 영역 관리를 디바이스 자체가 처리하기 때문에, 파일시스템 혹은 데이터베이스와 같은 저장 응용 프로그램으로부터 생성되는 메타데이터도 또한 자체적으로 관리한다. 데이터의 물리적인 위치를 기술하는 메타데이터는 OSD로 옮겨지고 디바이스 인터페이스는 블록에서 객체로 변경된다. 메타데이터를 스토리지 디바이스로 이동시키면서, 오늘날의 파일 시스템에 의하여 제공되는 NAS 시스템의 핵심 기능과 유사한 플랫폼 독립적인 인터페이스를 제공한다. 호스트들이 비록 데이터를 처리하기 위하여 다른 응용프로그램을 사용할지라도 객체는 호스트들이 데이터를 공유하는 것을 가능하게 한다. 예를 들면, 서로 다른 호스트에서 실행되고 있는 어떤 데이터 베이스 응용 프로그램과 어떤 파일 서버는 다른 백업 응용과 데이터를 공유할 수 있을 것이다. 또한, 응용 프로그램은 그들의 구조체들을 블록들의 집합체로 저장하기 보다 객체로 저장할 수 있기 때문에 보안 정책은 객체 수준에서 수립될 수 있다. 이러한 정책은 객체의 집합으로부터

하나의 객체, 심지어 하나의 객체에서 바이트 단위로 보안 정책 범위를 결정할 수 있다.

데이터 공유 및 보안에 대한 기능이 성립되면서, OSD 기술의 가장 즉각적인 두 가지 이점인 플랫폼 독립적인 데이터 공유와 응용 수준의 보안이 향상되었다. 객체의 향상된 데이터 공유는 저장되고 있는 데이터를 나타내는 속성뿐만 아니라 스토리지를 접근하는 높은 수준의 인터페이스의 결과이다. 객체는 파일시스템 의미론을 사용한다. OSD 매소드들 중에는 다음과 같은 것들이 있다.

- Create and remove
- Open and close
- Read, write and append
- Get and set attributes for both objects and groups

인터페이스는 객체 내에서 데이터를 조작하기 위한 응용프로그램 특정한 매소드들을 제공하기 위하여 쉽게 확장될 수 있다. 게다가, SNIA TWG(Technical Work Group)는 QoS(Quality of Service) 혹은 보안 보장 같은 응용 프로그램 특정 매개 변수들을 성립하기 위하여 디바이스 내에서 응용 프로그램들이 세션을 생성하는 것을 가능하도록 하려는 시각을 갖고 있다. 객체에 기반을 둔 스토리지 디바이스는 새

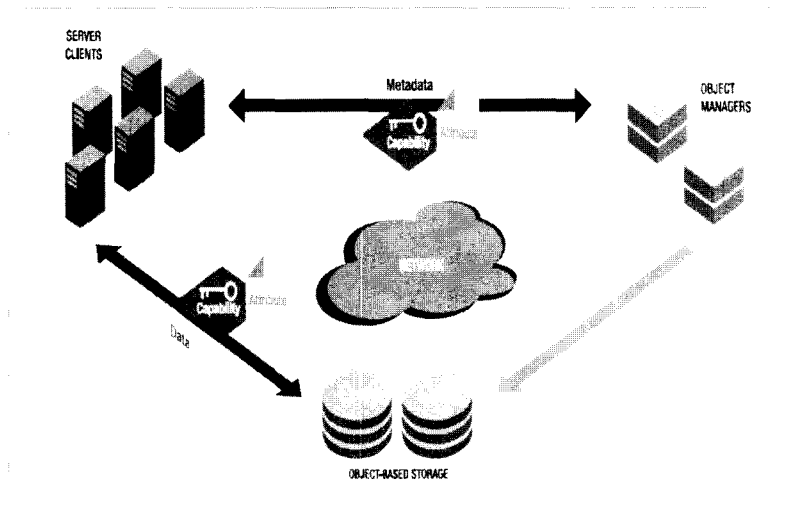
로운 수준의 프로그래밍 가능성을 제공하고, 각 OSD 클라이언트 단위로 OSD 클라이언트에 데이터를 서비스할 수 있다.

객체 속성은 접근 시간, 크기, 사용자 정보와 같은 객체와 객체의 환경에 관한 정보를 포함한다. 또한, 백업, QoS 요구사항, 객체 접근 패턴(연속적 혹은 임의적), 다른 객체들과의 관계와 같은 정보들이 객체 속성으로 저장될 수 있다. 속성은 더욱 큰 상호 운용적인 객체 접근을 제공하고, 객체의 의미를 묘사하는 정보를 OSD에 제공한다. 이러한 정보는 디스크 상의 배치와 자원 준비와 같은 더욱 효과적인 데이터 관리를 용이하게 할 수 있다. 일단 OSD가 객체를 묘사하는 속성에 접근을 하면 OSD는 디바이스에서 향상을 생성하는데 속성을 사용할 수 있다. 예를 들면, OSD는 자주 접근되는 데이터는 더욱 빨리 이용 가능하거나, 이전 버전의 객체가 객체 자신이 변하고 있는 것을 감지할 때 자동으로 저장되는 것을 확실히 할 수도 있을 것이다. OSD가 자신이 포함하고 있는 데이터의 의미를 이해했을 때, OSD는 그러한 정보를 OSD 처리 능력을 이용하는 새로운 스토리지 응용 프로그램을 생성하는데 사용할 수 있다.

보안은 객체 기반 스토리지와 블록 기반 스토

리지를 구별해주는 또 다른 하나의 요인이다. 블록 기반 SAN은 디바이스와 fabric 수준에서 정교하지 못한 계층의 보안을 제공한다. 이는 어떤 호스트가 어떤 디바이스에 대한 접근 권한을 획득하던지 혹은 못하던 지의 수준이다. 보안은 사용자 혹은 그룹 수준에서 구현되지 않았다. 반면에, 객체 기반 스토리지는 SAN 보안 아키텍처보다 우수한 새로운 수준의 유연성과 정교함을 제공한다. 객체는 큰 디바이스 혹은 fabric을 작은 개개의 보안 영역으로 분할 할 수 있다. 이렇게 분할된 각각의 보안 영역에 대한 접근 정책은 스토리지 응용 프로그램에 의하여 결정된다. 모든 접근은 인가되는 과정을 통하지만, 중앙 집중 인가 과정을 통하기 위한 통신에 수반되는 지연은 발생하지 않는다.

객체 기반 스토리지에 대한 SNIA OSD TWG의 보안 아키텍처는 <그림 12>에 나타나 있다. 이 접근 방법은 클라이언트 권한의 tamper-proof 묘사를 포함함으로써 암호기법적으로 강한 기능에 기반을 두고 있다. 객체 관리자는 보안 정책을 관리하고 capability를 생성함으로써 클라이언트를 위한 스토리지 접근 권한을 허가하는 데에 책임이 있다. <그림 12>에서 알 수 있듯이, OSD와 객체 관리자는 키를 공유하지만, 클라이언트



<그림 12> 객체 기반 스토리지 보안

는 그 키를 갖고 있지 않다.

클라이언트가 암호 기법을 사용하여 객체 관리자에게 성공적인 스토리지 요구를 한다면, 객체 관리자는 OSD에 대한 키를 기반으로 한 capability를 생성한다. 이 capability는 인가된 스토리지 운용을 기술한다. 그리고, 클라이언트는 객체 기반 스토리지 디바이스에 요청을 보내기 위하여 받은 capability를 사용한다. Capability는 클라이언트의 보안 정책을 나타내고, 어떤 객체에 대한 클라이언트의 접근 권한을 결정하기 위해 OSD에 의하여 검사된다. OSD는 또한 보안 키를 갖고 있으므로, OSD는 인증을 위하여 클라이언트로부터 전달 받은 capability를 검사할 수 있고, 이렇게 하여 클라이언트의 요청을 수락하던지 아니면 거부한다. OSD는 자체적으로 보안 정책을 제공하지 않고, 오로지 보안을 강화시키는 메커니즘 즉, 확장 가능한 보안 아키텍처를 구성하는 핵심인 분리를 제공한다. 클라이언트 특정한 인증 정보를 디바이스 위에서 유지할 필요가 없다는 것은 스토리지 디바이스가 시스템에서 클라이언트의 숫자에 관계없이 독립적으로 확장될 수 있다는 것을 의미한다.

iSCSI는 OSD를 구현하기 위하여 요구되지는 않는다. 그러나, OSD는 새로운 SCSI 명령어 집합으로서 구현될 수 있기 때문에 새로운 스토리지 솔루션으로서 어떤 가능성을 갖고 있다. 이러한 솔루션은 NAS와 SAN 기술의 가장 좋은 점만을 결합한 것이고, 상호운용성과 상위 계층의 보안을 향상시킨다. 결국에는 TCO를 매우 효과적으로 감소시킨다. iSCSI와 OSD는 OSD가 상위에서 IP가 동작하는 어떠한 물리 네트워크에도 연결될 수 있도록 할 것이다.^[11]

V. 결 론

기업의 비즈니스 환경이 인터넷의 발달, 스토리지 도입 비용의 감소, 스토리지 관리 비용 증가, 정보의 기하급수적인 증가로 인해 급격히 변

하고 있으며, 이에 따라 새로운 개념의 저장장치가 요구되고 있다. 이러한 급변하는 IT 환경에 대처하기 위해서는 기존의 스토리지 시스템에 변화가 있어야 된다. 본 논문에서는 이러한 새로운 개념의 스토리지 시스템에서 사용되거나 사용될 것으로 예상되는 새로운 파일 접근 시스템인 DAFS, 고성능 대용량인 InfiniBand 연결망을 간단히 살펴 보았다. 또한, 파일 시스템에 객체 지향적인 스토리지 환경이 예고되고 있음을 OSD를 통해 보여 주었으며, 이러한 객체 지향적인 스토리지 시스템이 IP 스토리지에서 부분적으로 연결의 시도가 일어나고 있음을 보여 주었다.

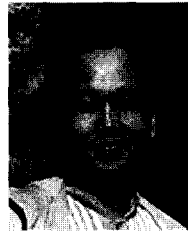
스토리지 환경은 기업의 비즈니스 환경의 변화에 따라 같이 변하게 된다. 비즈니스의 환경은 새로운 문화의 변화에서 기인한다고 보면, 새로운 변화에 쉽게 적응하기 위해서는 새로운 스토리지 기술, 새로운 스토리지 시스템의 출현에 관심을 기울여야 한다.

참 고 문 헌

- [1] Aberdeen Group, "iSCSI: New Business Benefits and Solutions An Executive White Paper", <http://www.snia.org/tech-activities/ip-storage/iscsi-new-business/>, 2002, 08
- [2] Next Generation I/O: A New Approach to Server I/O Architecture, Technical White Paper of Aberdeen Group, Feb. 1999.
- [3] Compaq, "PCI-X: An Evaluation of the PCI bus", TC990903TB, White Paper of Compaq Computer System Co., Sept. 1999
- [4] InfiniBand Architecture Overview, IBTA Presents at IDF Spring 2000, Feb. 15, 2000,
- [5] 박경, "InfiniBand의 개요", 주간기술동향, 통권 967호, pp.13~22, 2000. 10.

- [6] InfiniBand Trade Association, "InfiniBand Technology Overview", <http://www.infinibandta.org/ibta>, 2003
- [7] Virtual Interface Architecture Specification Version 1.0, Dec. 16, 1997.
- [8] 홍승표, 정현수, "인피니밴드 기술 및 시장 전망", IITA, 2002. 05.
- [9] 장성훈, "새로운 프로토콜 DAFS의 등장", 시사컴퓨터, 2002. 6
- [10] DAFS Implementers's Forum, <http://www.snia.org/tech-activities/dafs/>
- [11] Intel, "Object-based Storage : The Next Wave of Storage Technology and Devices", White paper, July 2003

저자 소개



김성운

1998년 충남대학교 전자공학과(석사), 2003년 충남대학교 대학원 전자공학과(박사수료), 1989년 2월~2003년 현재: 한국 전자통신연구원 입출력시스템 연구팀장, <주관심 분야: 스토리지 시스템, 컴퓨터 구조, SoC 설계>



김보관

1978년 KAIST 전기전자공학과(석사), 1989년 University of Wisconsin 전기 컴퓨터 공학과(박사), 1980년 3월~1991년 2월: 금호 공과대학 조교수, 1991년 3월~2003년 현재: 충남대학교 전자공학과 교수, <주관심 분야: VLSI CAD, VLSI 시스템 설계 및 모델링, HW/SW Co-Design, 통신시스템용 VLSI 설계>