

□ 기술개설 □

비휘발성 메모리를 이용한 데이터베이스 회복 구조 및 기법

한국전자통신연구소 백영식*
충남대학교 진성일**

● 목	차 ●
1. 서 론	3.1 이중면 비휘발성 메모리 구조
2. MMDB 관련 연구 현황	3.2 페이지 기반 이중면 비휘발성 메모리 구조
2.1 로그 기반 회복 구조	4. 비휘발성 메모리를 이용한 회복 구조 및 처리 알고리즘
2.2 그림자 페이지 기반 회복 구조	5. 결론 및 추후 연구 사항
2.3 비휘발성 메모리 응용	
2.4 비휘발성 메모리 관련 연구 현황	
3. 비휘발성 메모리 구조	

1. 서 론

현대 사회는 정보화 사회로 전환되면서 정보를 저장하는 데이터베이스 시스템이 더욱 중요하게 인식되고 있다. 데이터베이스 시스템은 대용량을 필요로 하는 분야, 실시간을 필요로 하는 분야 등 여러 가지 적용 분야가 있기 때문에 이에 적합한 데이터베이스 시스템 구현 기술을 개발하는 것이 필요하다.

주기억 장치 데이터베이스 시스템(MMDB : Main Memory Database)은 데이터들을 주기억장치에 저장하는 것만으로도 간단히 성능을 높일 수 있는 기술로 인식되고 있어, 이에 대한 많은 연구가 수행되었다. 지금까지 반도체 메모리 칩의 집적도는 매년 두 배 이상으로 증가하고 있으며, 비트 당 가격은 계속 떨어지는 추세이다. 현재 1G비트 용량의 DRAM(Dynamic Random Access Memory)이 개발된 상태[1]이기 때문에 대용량의 MMDB도 실현 가능하게 되었다.

디스크 기반 데이터베이스 시스템(DRDB :

Disk Resident Database)은 디스크를 기본적인 저장 매체로 사용하기 때문에 디스크의 특성에 최적화되어 있다. 그러나 MMDB는 디스크가 시리얼 접근을 하는 것과는 달리 무작위 접근을 하는 주기억장치를 기반으로 하기 때문에 DRDB와는 다른 기법이 필요하게 된다. MMDB는 데이터베이스를 무작위 접근을 하는 주기억장치에 저장하기 때문에 DRDB에서와 같이 복잡한 디스크 버퍼 관리 등이 필요하지 않다. 그러나 MMDB는 데이터베이스를 주기억장치에 모두 저장하기 때문에 회복을 위한 백업이 가장 큰 문제점이 된다. 이러한 문제점은 MMDB가 대용량화 될 수록 심각해 진다.

비휘발성 메모리는 이러한 문제점을 해결할 수 있는 하나의 기술로 인식되고 있다[2]. 그러나 실제적인 응용 방법 등에 대해서는 연구가 미흡한 상태이다. [3]에서는 배터리 백업을 이용한 방법에 대해서 기술하고 있지만, 배터리는 능동적인 방법으로 수동적인 방법인 비휘발성 메모리보다 신뢰성이 떨어진다. 비휘발성 메모리는 강 유전체를 사용한 메모리로 전원이 공급되지 않는 경우에도 10년간 정보를 잃어버리지 않는 특성을 가지고 있다. 현재 사용 가

*정회원
**종신회원

능한 비휘발성 메모리의 용량은 소용량이지만 대용량, 고속 및 저 전압의 잠재성으로 인해 연구가 진행 중에 있다[4].

본 논문에서는 MMDB 관련 연구 현황에 대해 기술하고, 비휘발성 메모리를 이용한 고속 회복 구조를 기술한다. 특히, 페이지 기반 이중면 비휘발성 메모리를 기반으로 설계된 회복 구조에서는 기존의 로그 기반 회복 구조 보다 그림자 페이지 기반 회복 구조가 더 우수한 것을 기술함으로써 발달된 반도체 기술을 응용하는 것이 필요한 것임을 기술한다. 이러한 응용 분야의 발굴은 반도체 기술 발전에도 박차를 가할 수 있을 것으로 보이며, 단순한 메모리 기능만이 아니라 약간의 기능을 첨가함으로써 기존에는 성능이 떨어지던 회복 기법이 우수한 특성을 발휘함을 보인다. 앞으로는 하드웨어 기술을 고려한 알고리즘 개발이 필요할 것으로 사료된다. 그림자 페이지 방식은 통신망과 같이 빠른 회복을 필요로 하는 곳에 적합한 방식이다.

2장에서는 기존의 MMDB 관련 연구 및 비휘발성 메모리 관련 연구에 대해 기술하고, 3장에서는 고속 회복 구조를 위해 필요한 페이지 기반 이중면 비휘발성 메모리에 대해서 기술한다. 4장에서는 일반적인 비휘발성 메모리를 사용한 회복 구조 및 3장에서 제안된 페이지 기반 이중면 비휘발성 메모리를 사용한 회복 구조 및 처리 방법에 대해서 기술한다. 5장에서는 결론 및 추후 연구 사항에 대해 기술한다.

2. MMDB 관련 연구 현황

본 장에서는 MMDB 관련 기존의 연구 및 비휘발성 메모리 연구 현황에 대해 기술한다. MMDB의 회복에 관한 대부분의 연구는 로그를 기반으로 하고 있지만, 하드웨어 기술의 발달로 인해 그림자 페이지 기반에 대해서도 조금씩 연구가 시작되었다.

2.1 로그 기반 회복 구조

데이터베이스 시스템의 회복을 위한 대표적인 방법으로 로그를 사용하는 방법 및 그림자

페이지를 사용하는 방법이 있다. 그림자 페이지 방법은 회복시 철회가 필요 없는 장점이 있으나, 디스크를 기반으로 하는 회복 구조에서는 일반적으로 로그를 사용하는 방법에 비해 성능이 떨어지고, 동시성을 위한 작은 단위 잠금이 어렵고, 페이지 테이블 관리 등의 부담으로 인해 현재 구현된 대부분의 데이터베이스 시스템들은 로그를 기반으로 하고 있다.

로그를 기반으로 하는 MMDB 시스템에서는 회복을 위해 안정된 로그 버퍼를 사용하며, 트랜잭션 수행시 로그 버퍼에 수정 내용을 먼저 기록하는 것이 필요하기 때문에 이를 처리하기 위한 시간이 필요하다. 이에 필요한 처리 시간을 줄이기 위해 미국의 프린스턴 대학에서 개발한 MMM(A Massive Memory Machine)에 구현된 데이터베이스 시스템에서는 하드웨어 로깅 기법인 HALO(Hardware Logging)를 개발하여 사용하고 있으며, 무 중단 백업 기법[5]에서도 하드웨어 로깅을 사용하고 있다.

로그를 기반으로 하는 MMDB 시스템에서는 주기적으로 데이터베이스의 내용을 비휘발성 매체에 저장하는 것이 필요하다. MMDB는 회복을 위해서만 비휘발성 저장 매체인 디스크를 사용하기 때문에, 트랜잭션 수행에 영향을 최소화하면서 백업을 하는 것이 중요하다.

로그를 기반으로 하는 검사점[6] 방법은 검사점 수행시 로그의 내용을 백업 메모리에 기록한 후 디스크에 적재하는 방법이다. 메모리 양은 실제 사용되는 데이터베이스의 2배가 필요하다. 만약 데이터베이스 백업용 메모리와 실제 데이터베이스용 메모리가 독립된 버스를 사용하게 되면 검사점 수행시 트랜잭션에 최소한의 영향을 주면서 디스크에 백업이 가능하다. 로그 버퍼는 안정 메모리를 사용하여야 되며, 데이터베이스 및 백업을 위한 메모리는 휘발성 메모리이다.

퍼지 검사점 방법[7]은 데이터베이스의 페이지를 순환적으로 검사하면서 디스크에 백업하는 방법이다. 이를 위해 비휘발성의 그림자 메모리를 사용하게 된다. 이 방법의 장점은 그림자 메모리의 양이 데이터베이스를 위한 메모리의 양보다 훨씬 적어도 되는 장점이 있다. 80-20 규칙을 적용할 경우 20% 이하의 그림

자 메모리만으로도 동작할 수 있다는 장점이 있다. 그러나 다른 두 방법에 비해 메모리의 사용은 적지만, 검사점 수행 시 메모리 이동이 일어나게 된다. 특히, 트랜잭션들이 페이지를 골고루 사용할 경우 더욱 많은 양의 그림자 메모리가 요구된다. 이 방법은 소수의 페이지를 집중적으로 사용하는 경우에 효율적이다.

비중단 검사점 방법[5]은 로그 기반[6]과 마찬가지로 실제 필요한 데이터베이스를 위한 메모리 양의 2배를 사용한다. 이 방법에서는 실제적인 데이터베이스에서는 읽기와 쓰기를 수행하고, 백업을 위한 메모리에는 쓰기만 수행한다. 검사점 수행 시에는 백업을 위한 메모리에 쓰기를 중단시키고 백업을 위한 메모리의 내용을 디스크에 백업한다. 디스크에 백업이 완료된 후 다시 쓰기 수행을 시작한다. 그러나 두개의 메모리 내용을 일치시키기 위해서는 실제적인 데이터베이스의 내용을 백업을 위한 데이터베이스에 복사를 해야 한다. 이러한 복사 및 디스크의 백업 수행은 전용 프로세서가 수행하게 된다. 그러므로 복사 수행 시 데이터베이스 전용 프로세서와 충돌이 발생하여 성능을 감소시키는 요인이 된다. 이를 최소화하기 위해 백업 또는 복사는 변화된 페이지에 대해서만 수행하게 된다. [5]는 이러한 2배의 메모리를 사용하는 방법 외에도 이중 포트 메모리 및 이중 버퍼를 사용하는 방법을 제시하고 있다.

빠른 재시동을 위해서는 점진적 방법 등이 사용되고 있다[5]. 점진적 방법은 재시동 시 필요한 작은 수의 페이지만을 적재한 후 바로 트랜잭션을 수행하는 방법이다. 페이지를 우선 적재하는 방법 및 요구 시 적재되는 방법 등이 있지만, 회복이 완료되는 시간이 길어지고 트랜잭션이 자주 중단되어 오히려 처리량이 감소할 수 있다[8].

2.2 그림자 페이지 기반 회복 구조

DRDB(Disk Resident Database)를 기반으로 하는 기존의 그림자 페이지 기법의 한가지 문제점은 갱신된 페이지를 위해 항상 새로운 저장 장소를 사용하기 때문에 연관된 페이지가 흩어지게 된다. 그러므로 디스크 접근 시 페이지 탐색 시간이 증가하게 되는 단점이 있다.

그림자 페이지 기법은 물리적 페이지를 논리적 페이지로 관리하기 위해 페이지 테이블을 사용하는 부담이 있다. 이것은 페이지 테이블이 안정 저장 장치인 디스크에 저장되기 때문이다. 하나의 예로 페이지의 크기가 8KB이고 페이지 테이블 엔트리가 4바이트 크기 일 때, 100MB 크기의 데이터베이스를 위해서는 50KB의 페이지 테이블이 필요하고, 100GB를 위해서는 50MB 크기의 페이지 테이블이 필요하다.

[11]에서는 트리 형태의 페이지 테이블을 모두 주기억장치에 저장함으로써 페이지 테이블 관리의 부담을 경감시킬 수 있음을 보여 주었다.

[10]에서는 페이지 테이블을 통한 액세스 부담을 피할 수 있는 방법으로 두 가지 겹쳐 쓰는 기법 및 버전 선택 기법을 제안하였다. [10]에서 제시한 두 가지 겹쳐 쓰는 방법 모두 겹쳐 쓰기를 보장하기 위해 공통의 저장 영역을 사용한다.

겹쳐 쓰는 기법 중 재수행을 하지 않는 구조는 데이터 페이지를 갱신하기 전에, 원래 페이지(그림자 페이지)를 공통의 저장 공간에 저장한다. 갱신된 모든 페이지가 디스크에 기록된 후에 트랜잭션이 커밋 된다. 시스템 장애 후 회복을 위해 커밋 되지 않은 트랜잭션의 리스트가 필요하다. 회복은 공통의 저장 공간에 저장된 페이지를 사용하여 수행된다.

겹쳐 쓰는 기법 중 철회를 하지 않는 구조는 재수행을 하지 않는 구조와는 달리, 트랜잭션이 수행 중 그림자 페이지 페이지는 원래의 장소에 보존된다. 모든 갱신은 먼저 공통의 저장 공간에 기록한 후 트랜잭션은 커밋된 것으로 간주된다. 그러나 로크는 갱신된 페이지를 그림자 페이지 페이지로 교체한 후 풀어준다. 시스템 장애 후 회복을 위해 커밋된 트랜잭션의 리스트가 필요하다.

버전 선택은 페이지 테이블을 통한 오버헤드를 피할 수 있는 다른 한가지 방법으로 디스크 상에 물리적으로 이웃한 두개의 블록을 사용한다. 하나는 원래의 그림자 페이지 버전이고 다른 하나는 갱신된 현재 버전을 위한 블록이다. 버전 선택 알고리즘을 사용하여 이중 어

느 것이 현재 버전 인가를 구분한다. 현재 버전은 저장된 타임 스탬프를 이용하여 구분할 수 있다. 이 방법의 문제점은 겹쳐 쓰는 방법과는 달리 기존의 그림자 페이지 방법과 마찬가지로 두 배의 디스크 공간을 사용한다는 것이다.

[11]에서는 디스크와 주기억장치의 발달과 여러 가지 기술(로그-구성 화일 시스템, 스트리핑, 미러링, 그림자 갱신 및 그룹-커밋)들을 사용함으로써 그림자 페이지 기법이 현재 최고의 성능을 발휘하는 로그 기반 시스템과 비교할 수 있음을 보였다.

그림자 페이지의 다른 한가지 단점은 작은 단위 잠금이 어렵다는 것이다. 그 이유는 만약, 두개의 트랜잭션이 하나의 페이지중 서로 다른 영역을 갱신할 때, 한 트랜잭션은 성공하고 다른 트랜잭션은 실패하는 경우 로그가 없기 때문에 실패한 트랜잭션을 철회할 수 없게 되기 때문이다. [11]에서는 반도체 기술의 발달로 가용하게 된 주기억 장치 메모리에 여러 복사본을 사용함으로써 이러한 문제를 해결할 수 있음을 보였다.

2.3 비휘발성 메모리 응용

비휘발성 메모리를 데이터베이스에 적용하는 연구가 진행 중이다[14]. 기존의 MMDB에 관한 연구들에서는 휘발성의 주기억 장치를 사용하기 때문에 회복을 위한 백업이 문제가 되었다. 그러나 비휘발성 메모리를 사용하면 이러한 백업 문제가 사라지게 된다. [7]에서는 적은 용량의 안정 메모리를 그림자 메모리로 사용한 퍼지 검사점 방법을 제안하였다. [5,9]에서는 프래쉬 메모리를 사용한 회복 구조 및 저장 장치를 제안하였으며, [3]에서는 배터리 백업을 사용한 저장 장치가 점차 사용 가능성을 기술하고 있다. [5]에서는 프래쉬 메모리를 디스크 대치용으로 사용하는 것을 기술하고 있지만, [9]에서는 배터리 백업 SRAM(Static Random Access Memory)을 프래쉬 메모리와 같이 사용하여 무작위 액세스가 가능한 저장 장치를 기술하고 있다. 이와 같이 시리얼 특성을 가진 프래쉬 메모리는 현재 대용량 제품이 나오고 있지만, 무작위 접근이 가능한 비휘발성 메모리는 현재 소용량이다. 그러나 2000년

대 초에는 대용량이 출현할 것으로 보이기 때문에 이를 데이터베이스에 적용하는 연구가 필요하다.

[2,12]에서는 비휘발성 메모리가 데이터베이스에 응용될 경우 많은 문제점이 사라질 것으로 기술하고 있지만, [14]에서는 대용량의 비휘발성 메모리가 출현할 경우 단순히 주기억장치를 비휘발성 메모리로 대체하는 것 보다 특수한 구조의 메모리를 사용할 경우 로그 기반 회복 알고리즘보다 그림자 페이지 기반 회복 알고리즘이 더 잘 적용되는 것을 기술하고 있다.

2.4 비휘발성 메모리 관련 연구 현황

현재 반도체 소자 기술을 대표하고 있는 DRAM은 1 G 비트 대 까지 고 집적화 되어 감에 따라 기술적으로, 경제적으로 한계에 이르고 있다. 이에 반해 낮은 전압이면서 고속이고 비휘발성인 메모리에 대한 요구가 증대하고 있다. 최근에 이르러 유전체 기술이 실용화됨에 따라 DRAM(Dynamic Random Access Memory)의 벽을 넘을 수 있는 길이 열리게 되고, 저 전압, 고속, 비휘발성 메모리의 실현에 밝은 전망을 보여주고 있다. 현재 디스크 대용으로 많이 사용되는 프래쉬 메모리가 시리얼 액세스 특성을 가지는 것과는 달리 무작위 액세스가 가능한 FRAM(Ferroelectric Random Access Memory)은 이러한 강 유전체를 이용한 메모리로, 기존의 반도체 메모리처럼 회로에서 분리하지 않은 상태에 쓰기/읽기가 가능하며, 한번 쓰면 전원 차단 시에도 10년간 정보가 유지된다. 그러나 기존의 반도체 메모리와는 달리 쓰기 횟수 제한이 있다.

비휘발성 메모리는 제조 공정 등의 문제 외에도, 상기한 바와 같이 쓰기 횟수 제한이 있어, 정보 처리 시스템 적용에 걸림돌로 작용하고 있다. 강 유전체는 분극 반전에 의한 막 피로 현상이 있기 때문에 고쳐 쓰기/읽어 내기 횟수가 제한을 받게 된다. 이를 해결하기 위해서 특성이 좋은 강 유전체의 개발에 대한 연구, 이중 셀과 같은 구조적인 해결 방법 및 소프트웨어적인 해결책인 칩 수명 연장[15]에 대한 연구가 있다.

10MHz(100 nsec 사이클)의 연속 액세스를 가정한다면 10년간의 동작을 보정하기 위해 $3 * 10^{13}$ 회가 필요하게 된다. 현재 발표된 결과에 따르면[7], 10^{10} 까지 횡수가 증가되었고, 현재 까지 개발된 특성이 좋은 강 유전체 Y1인 경우에는 10^{12} 회 이상의 고쳐 쓰기를 확인한 바 있다. 데이터 유지 시간도 실온에서 10년은 쉽게 넘어서리라는 예상이다. 그러나 데이터 유지 시간은 실제 적용에 있어 쓰기 횡수 제한 문제보다는 큰 문제가 되지 않는다.

3. 비휘발성 메모리 구조

3.1 이중면 비휘발성 메모리 구조

강 유전체를 사용한 비휘발성 메모리는 쓰기를 수행할수록 유전체가 열화 되어 기록을 할 수 없게 된다. 이러한 문제점을 해결하기 위해 많은 연구가 수행되고 있다. 유전체의 열화를 해결하기 위한 하나의 방법은 좋은 특성의 강 유전체를 개발하는 것이다. 다른 하나의 방법으로 특수한 메모리 구조를 사용하는 방법이 있다.

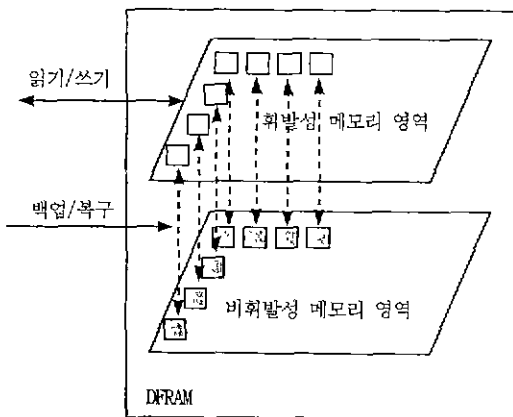


그림 1 이중면 FRAM(DFRAM) 구조

그림 1에서는 이중 셀을 사용하여 유전체의 열화 문제를 해결한 이중면 비휘발성 메모리의 내부 구조를 보여준다. 이중면 비휘발성 메모리 각각의 비트는 휘발성 셀과 비휘발성 셀의 쌍으로 구성된다. 이러한 이중 셀의 목적은 유전체의 열화 문제를 보정하기 위한 것이다. 즉,

평상시의 읽기와 쓰기는 쓰기 횡수 제한이 없는 휘발성의 셀에 수행하고 전원 차단 시에만 휘발성 영역의 내용을 비휘발성 영역으로 백업하여 비휘발성 영역에 쓰기 횡수를 대폭적으로 줄임으로써 쓰기 횡수 증가에 따른 유전체 열화 문제를 해결한다.

이중면 비휘발성 메모리는 전원 차단 시 복구 및 재생을 위해 제공되는 제어 선을 이용하여 휘발성 셀의 데이터를 비휘발성 셀에 쓰기 때문에 데이터를 잃어 버리지 않는다. 그러므로 전원 차단 시 이를 감지하여 복구 및 재생을 위해 제공되는 제어 선을 구동하여 백업 명령을 내리는 하드웨어가 필요하다.

3.2 페이지 기반 이중면 비휘발성 메모리 구조

이중면 비휘발성 메모리의 원래 목적은 유전체의 열화를 방지하기 위한 것이지만, 본 논고에서는 이러한 이중면 비휘발성 메모리를 데이터베이스 회복 구조에 적용할 경우 일반적인 단일면 구조의 비휘발성 메모리를 사용한 회복 구조보다 우수한 것을 보여 준다. 특히 디스크 기반에서는 로그 방식에 밀려 빛을 보지 못하였던 그림자 페이지 기법이 이중면 비휘발성 메모리를 사용한 구조에 가장 적합하다는 것을 보여 준다.

그림자 페이지 기법은 현재 버전과 그림자 페이지 버전을 사용하는 기법이다. 그러므로 이중면 비휘발성 메모리의 두개면 중 하나는

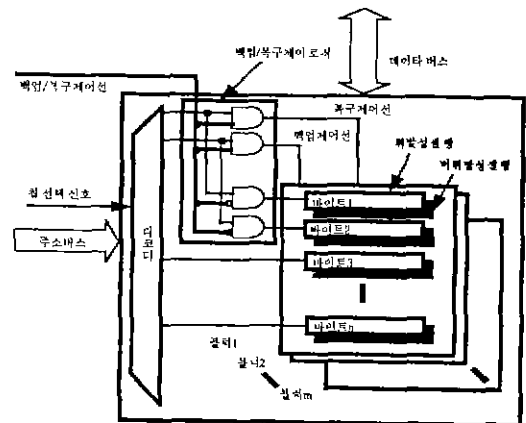


그림 2 페이지 단위로 백업 및 복구가 가능한 이중면 FRAM

현재 버전으로 사용하고, 다른 한 면은 그림자 페이지 버전으로 사용함으로써 그림자 페이지 기법을 구현할 수 있다. 그러나 현재 사용 가능한 이중면 비휘발성 메모리는 정전 시 백업을 위하여 전체 면을 하나의 제어 선으로 제어하기 때문에 하나의 현재 버전과 하나의 그림자 페이지 버전만이 존재한다. 그러므로 여러 트랜잭션이 동시에 수행될 경우 페이지 단위로 제어할 수 없게 되며, 동시성 제어가 어려워진다.

이러한 문제점은 이중면 비휘발성 메모리의 백업 및 복구를 분할된 블럭(또는 페이지) 단위로 수행함으로써 해결된다. 그림 2에서는 이러한 페이지 단위의 백업 및 복구를 가능하게 하는 이중면 비휘발성 메모리 구조를 보여준다. 새로운 메모리 칩은 외부와 연결되는 물리적인 핀의 수가 기존의 메모리와 호환성을 가지는 것이 요구된다. 이를 위해 본 논고에서 제안된 페이지 단위로 백업 및 복구 제어가 가능한 이중면 비휘발성 메모리는 기존의 이중면 비휘발성 메모리와 같은 수의 핀을 사용하지만, 내부의 디코더는 기존의 이중면 비휘발성 메모리와 차이가 난다.

일반적인 메모리 칩은 주소 버스와 칩 선택 신호를 사용하여 비트 또는 바이트 단위의 메모리 영역을 선택하고, 데이터 버스를 통하여 정보를 읽거나 기록한다. 메모리 내부의 디코더는 주소 버스로 입력된 정보를 판독하여 데이터 영역을 선택하는 기능을 수행한다. 기존의 이중면 비휘발성 메모리에서는 이러한 주소 버스 및 데이터 버스와 더불어 백업 및 복구를 제어하기 위한 제어 선을 제공한다. 이 제어 선은 전체 메모리 셀을 제어하여 정전시 휘발성 영역의 내용을 비휘발성 영역으로 저장하는 기능 및 전원 공급 시 그 반대의 기능을 수행한다.

본 논고에서 제안하는 페이지 기반 이중면 비휘발성 메모리는 주소 버스, 데이터 버스 및 백업/복구 제어 선을 사용하지만, 디코더의 기능은 기존의 이중면 비휘발성 메모리와 상이하다. 기존의 이중면 비휘발성 메모리에서는 제어 선에 가해지는 신호의 형태에 의해 백업 또는 복구가 결정되었다. 즉, 하나의 제어 선으로 두 가지 기능을 제어한 방법이다. 그러나 페이

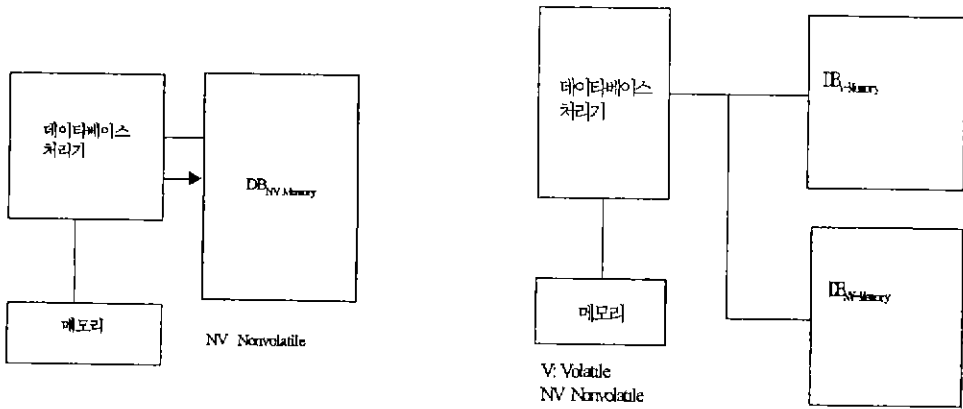
지 기반 이중면 비휘발성 메모리는 제어 선이 0일 경우에는 일반적인 메모리 액세스를 하게 하고, 제어 선이 1인 경우에는 메모리 내부 각각의 페이지에 연결된 백업 및 복구 제어선을 구동시키게 된다. 제어 선이 1인 경우 주소 버스가 각각의 블럭 중 0번 주소를 가리킬 경우에는 백업을, 1번 주소일 경우에는 복구를 수행하게 함으로써 백업 및 복구를 페이지별로 제어할 수 있게 된다. 백업/복구 제어 로직은 이러한 기능을 수행하며, 그림 3에서는 복잡성을 피하기 위해 블럭 1을 위한 백업/복구 제어 로직만을 기술하였다.

4. 비휘발성 메모리를 이용한 회복 구조 및 처리 알고리즘

본 장에서는 비휘발성 메모리를 이용하여 고속 회복 구조를 설계할 수 있음을 기술한다. 특히, 페이지 기반 이중면 메모리를 이용할 경우에 그림자 페이지 기법이 로그 기반 보다 우수한 특성을 발휘함을 보인다.

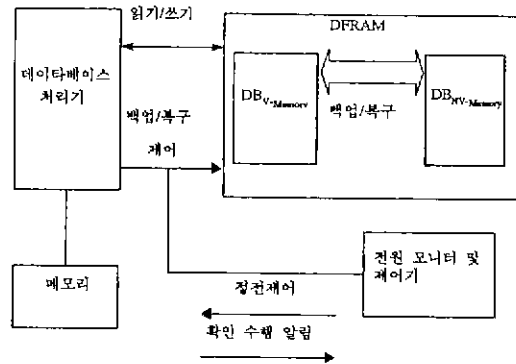
그림 3에서는 비휘발성 메모리를 이용하여 구성 가능한 회복 구조들을 보여준다. NVARCH-I(NonVolatile Architecture-I)은 주기억 장치를 모두 비휘발성 메모리로 구성한 경우를 보여주고, NVARCH-II(NonVolatile Architecture-II)는 휘발성 메모리와 비휘발성 메모리를 같이 사용한 경우를 보여준다. NVARCH-D(NonVolatile Architecture-Dual)은 3.2절에서 기술한 페이지 기반 이중면 비휘발성 메모리를 사용한 구조를 보여준다.

대표적인 회복 기법으로 로그 기반 및 그림자 페이지 기반을 들 수 있다. 디스크에 데이터베이스를 저장하는 기존의 데이터베이스 시스템에서는 로그 방식이 그림자 페이지 방식보다 여러 가지 측면에서 우수한 것으로 인식되고 있다. 그 이유는 디스크는 주기억 장치보다 속도가 훨씬 느리고, 직렬 접근을 하기 때문에 페이지 단위로 입출력을 수행하기 때문이다. 그러나 그림 3에 기술된 바와 같이 비휘발성 메모리를 사용하는 경우에는 백업이 필요 없기 때문에 이러한 제약이 사라지게 된다. 그러므로 디스크 특성에 맞게 최적화된 기존의 알고



(a) 일반적인 비휘발성 메모리를 이용한 회복 구조-I(NVARCH-I)

(b) 일반적인 비휘발성 메모리를 이용한 회복 구조-II(NVARCH-II)



(c) 페이지 기반 이중면 비휘발성 메모리를 이용한 회복 구조(NVARCH-D)

그림 3 비휘발성 메모리를 이용한 고속 회복 구조

리즘을 무작위 액세스가 가능한 비휘발성 주기억 장치에 적용할 경우 다른 효과가 나타난다.

기술된 바와 같이 비휘발성 메모리를 MMDB에 사용할 경우 백업을 할 필요가 없어지기 때문에 로그 기반인 경우에는 트랜잭션 수행 시 로그를 처리하는 부담에 의해 성능에 영향을 준다. NVARCH-I은 주기억 장치가 비휘발성이기 때문에 즉시 갱신 회복 알고리즘이 적용된다. 그러므로 트랜잭션 수행 시 이전 값 및 이후 값을 로그 버퍼에 저장 해야 한다. 반면에 NVARCH-II에서는 휘발성 영역을 작업 영역으로 사용하기 때문에 지연 갱신 회복 알고리즘이 적용된다. 그러므로 이후 값만을 로그 버퍼에 저장하면 된다. 그러나 트랜잭션 처리

시 비휘발성 영역으로부터 휘발성 영역으로 복사를 해야 하는 부담이 있다. 이 부담은 기존의 디스크 기반에서는 당연히 발생하는 것이었다. 그러나 NVARCH-D에서도 지연 갱신을 수행하지만, 페이지 또는 블록 단위로 동시에 복사하기 때문에 복사 시간 부담을 거의 무시할 수 있다. 그러므로 NVARCH-I 또는 NVARCH-II보다 좋은 성능을 발휘한다.

[16]의 그림자 페이지 방식을 NVARCH-I에 적용할 경우에는 한번의 레코드 또는 페이지 복사가 발생한다. 2장에서 기술한 버전 선택 방법을 적용하면 페이지 테이블이 필요 없어지기 때문에 이에 관련된 부담이 없어진다. [16]에 기술된 그림자 페이지 알고리즘을

NCARCH-II에 적용할 경우에는 기존의 디스크 상주 데이터베이스와 같이 두 번 복사를 해야 한다. NVARCH-D에서는 복사가 레코드 또는 페이지 단위로 동시에 일어나기 때문에 이 시간을 무시할 수 있다. 그러므로 트랜잭션을 위해 부가적으로 필요하였던 모든 부담이 사라지게 된다.

그림자 페이지 기법은 작은 단위 잠금이 어려워 동시성을 증가시키기가 어려운 것으로 인식되고 있다. 만약 경쟁이 적은 경우에 주기억 장치 상주 데이터에 대해서는 전체 데이터베이스에 대해 잠금을 하여 시리얼로 트랜잭션을 처리하는 것이 좋다는 결과도 있다[12]. 그러나 긴 트랜잭션이 있는 경우 또는 다중 프로세서를 사용하는 경우에는 동시성 문제를 해결해야 한다. 다중 프로세서를 사용하는 경우에는 스핀 잠금 또는 제한된 스핀 잠금 등을 사용하는 것이 좋은 성능을 발휘한다[13]. 이때 잠금의 단위를 작게 함으로써 동시성을 증가시킬 수 있다.

페이지는 기존의 구조에서 디스크의 입출력 단위이다. 즉, 디스크를 사용할 경우에는 디스크를 관리를 최적화 하기 위해 일정 크기의 페이지 단위로 관리한다. 이것은 디스크가 시리얼 접근 방식이기 때문이다. 이와 같이 디스크를 사용할 경우 그림자 페이지 기반에서는 페이지를 안정한 디스크에 기록하는 것을 페이지 단위로 하기 때문에 작은 단위 잠금이 어려워진다. 그러나 본 논고에서와 같이 무작위 접근이 가능한 비휘발성 메모리를 사용할 경우에는 디스크 입출력이 없기 때문에 페이지 단위로 관리하는 것이 의미를 잃게 되고, 레코드 단위로 관리가 가능하며 실제로는 레코드 단위로 처리할 경우 복사하는 부담이 줄어들어 페이지 단위로 처리할 때 보다 성능이 증가 되는 효과가 있다. 본 논고에서 기술된 페이지 기반 이중면 비휘발성 메모리의 경우 실제로는 블록의 크기를 레코드 단위로 작게 할 수 있기 때문에 작은 단위 잠금이 자연스럽게 실현된다.

5. 결론 및 추후 연구 사항

본 논고에서는 기존의 MMDB에 관한 연구

중 회복 구조에 대하여 기술하였으며, 반도체 기술의 발달로 사용 가능하게 된 비휘발성 메모리를 적용하는 방안에 대해 기술하였다. 특히 이중면 구조를 가진 메모리 구조를 이용할 경우 기존에는 알고리즘 구동 시 큰 부담이었던 페이지 복사가 무시할 만큼 짧은 시간에 수행됨으로써 새로운 효과가 발생한다는 것을 기술함으로써 새로운 연구 방향을 제시하였다. 추후 연구 사항으로 비휘발성 주기억 장치인 경우에도 미디어 장애가 발생한 경우를 대비해 트랜잭션 수행에 최소한의 영향을 주면서 백업을 하는 방법이 필요하다.

참고문헌

- [1] Muroya, Sugibayashi, "화일 용도의 1 GDRAM," 월간 Semiconductor World, pp. 34-39, 1995.
- [2] D. Bondurant, F. Gnadinger, "Ferroelectrics for nonvolatile RAMs," IEEE Spectrum, pp.30-33, Jul., 1989.
- [3] G. Copeland, T. Keller, R. Krishnamurthy, M. Smith, "The Case for Safe RAM," Proc. of the 15th International Conference on Very Large Databases, pp.327-335, 1989.
- [4] R. Moazzami, P.D. Maniar, R.E. Jnoes, C. J. Mogab, "Integration of ferroelectric capacitor technology with CMOS," Proc. of Symposium on VLSI Technology Digest of technical papers, Jun., pp.55-56, 1994.
- [5] Hiroki Takakura, Yahiko Kambayashi. "A Design of a Transparent Backup System Using a Main Memory Database", Proc. of Intl. Conf. on Database Systems for Advanced Applications, pp.178-184, 1993.
- [6] Eliezer Levy, Abraham Silberschatz, "Log-Driven Backups : A Recovery Scheme for Large Memory Database Systems," Proc. of the Fifth JCIT, pp.22-25, 1990.
- [7] Xi Li, Margaret H. Eich, "Post-crash Log Processing for Fuzzy Checkpointing Main Memory Databases," Proc. of Intl. Conf. on Data Engineering, IEEE, pp.117-124,

1993.

[8] E. Levy, A. Silberschatz, "Incremental Recovery in Main Memory Database Systems," IEEE Transactions on Knowledge and Data Engineering, Vol. 4, No. 6, Dec., pp.529-540, 1992.

[9] Michael Wu, Willy Zwaenepoel, "eNVy : A Non-Volatile, Main Memory Storage System," Intl. Conference on Architectural Support for Programming Languages and Operating Systems, pp.86-97, Oct. 1994.

[10] Rakesh Agrawal, "Recovery Architectures for Multiprocessor Database Machines," ACM SIGMOD, pp131-145, 1985.

[11] T. Ylonen, "Shadowing Paging Is Feasible," Licentiate's thesis, Department of Computer Science, Helsinki University of Technology, 1994.

[12] Hector Garcia-Molina, Kenneth Salem, "Main Memory Database Systems : An Overview," IEEE Transactions on Knowledge and Data Engineering, Vol. 4, No. 6, Dec., pp.509-516, 1992.

[13] 진 성일, "공유기억장치 다중처리 시스템에서의 다중 서버 데이터베이스 관리 시스템의 성능평가," 박사학위논문, 한국과학기술원, DCS-78199, 1994.

[14] 백 영식, "주기억장치 데이터베이스 시스템에서 비휘발성 메모리를 이용한 회복 기법," 박사 학위 논문, 충남대학교, 1996.

[15] Y.S. Baek, S.I. Jin, "Wear Leveling and Restart Mechanism for FRAM and

EEPROM Applications," Intl. Symposium on Next Generation Database Systems and Their Applications," pp.288-292, 1993.

[16] R.A. Lorie, "Physical Integrity in a Segmented Database," ACM Transactions on Database Systems, Vol. 2, No. 1, , pp.91-104 , Mar. 1977.

백 영 식



1981 경북대학교 전자공학과 학사
 1984 경북대학교 전자공학과 석사
 1984~현재 한국전자통신 연구소 선임 연구원
 1994~1995 NTT 객원 연구원
 관심분야 : MMDB 회복 기법, ATM-LAN, TINA

진 성 일



1978 서울대학교 계산 통계학과 학사
 1980 한국 과학기술원 전산학과 석사
 1994 한국 과학기술원 전산학과 박사
 1983~1989 충남대학교 컴퓨터 과학과 조교수
 1987~1989 Northwestern 대학 전산학과 객원 교수

1990~1992 충남대학교 전자 계산 소장
 1989~1994 충남대학교 전산학과 부교수
 1994~ 충남대학교 컴퓨터과학과 교수
 관심분야 : 데이터베이스, 성능분석, 정보모델링, 멀티미디어