

논문 2017-12-20

NVDIMM의 동작 특성 분석 및 개선 방안 연구

(Characterization and Improvement of Non-Volatile Dual In-Line Memory Module)

박재현, 이형규*

(Jaehyun Park, Hyung Gyu Lee)

Abstract : High performance non-volatile memory system can mitigate the gap between main memory and storage. However, no single memory devices fulfill the requirements. Non-volatile Dual In-line Memory Module (NVDIMM) consisted of DRAMs and NAND Flashes has been proposed to achieve the performance and non-volatility simultaneously. When power outage occurs, data in DRAM is backed up into NAND Flash using a small-size external energy storage such as a supercapacitor. Backup and restore operations of NVDIMM do not cooperate with the operating system in the NVDIMM standard, thus there is room to optimize its operation. This paper analysis the operation of NVDIMM and proposes a method to reduce backup and restore time. Particularly, data compression is introduced to reduce the amount of data that to be backed up and restored. The simulation results show that the proposed method reduces up to 72.6% of backup and restore time.

Keywords : NVDIMM, Data compression, Memory system

1. 서론

컴퓨터 시스템이 작업을 수행하기 위해서는 연산결과를 저장하기 위한 메모리 시스템이 필수적이다. 메모리 시스템에 요구되는 특성은 빠른 접근 속도, 저렴한 가격, 긴 수명, 비휘발성 등이 있으나 현재 존재하는 어떠한 메모리 소자도 이러한 특성을 모두 갖추고 있지 못하다. 이를 극복하기 위해 여러 종류의 메모리 소자를 함께 사용하여 장점만을 살릴 수 있는 계층적 메모리 시스템이 사용되어 왔다.

계층적 메모리 시스템은 크게 캐쉬(Cache), 주기억장치(Main memory), 저장장치로 이루어져 있다. 이 중 하드디스크 또는 솔리드 스테이트 드라이브(SSD, Solid-state drive)가 사용되는 저장장치만 비휘발성 특성을 가진다. 이러한 저장장치는 캐쉬나

주기억장치에 비해 동작 속도가 느리며 데이터 접근이 파일 시스템을 통해 이루어져 추가적인 시간 부하가 발생한다.

이러한 단점을 극복하기 위하여 최근 주기억장치에 사용되는 휘발성 메모리 소자인 DRAM을 대신하여 비휘발성 메모리 소자인 Spin-Transfer Torque Random-Access Memory (STT-RAM), Phase-Change Memory (PCM), Resistive Random-Access Memory (RRAM)등의 소자를 사용하여 비휘발성 주기억장치를 구성하는 연구들이 이루어졌다. 하지만 아직까지 이러한 소자들은 짧은 수명, 긴 쓰기 시간, 미성숙 된 기술 수준 등의 문제로 인하여 실제 시스템에 사용되지 못하고 있다.

최근 빅데이터 및 데이터 마이닝과 같이 많은 양의 데이터를 빠르게 처리하고자 하는 요구가 더욱 늘어남에 따라 서버 시스템에서는 자주 사용되는 데이터를 주기억장치로 이루어진 램디스크에 저장하여 처리속도를 높이는 기법들이 활용되고 있다. 휘발성인 주기억장치에 데이터를 저장하고 있으므로 비상상황이 발생하여 시스템 재기동이 이루어지면 데이터를 복구하기 위해 많은 시간이 소모된다. 이러한 단점을 극복하기 위해 빠른 접근 속도를 가

*Corresponding Author(hglee@daegu.ac.kr)

Received: 12 May 2017, Revised: 15 May 2017.

Accepted: 23 May 2017.

J. Park: Arizona State University

H.G. Lee: Daegu University

※ 이 논문은 2012학년도 대구대학교 학술연구비 지원에 의한 논문임.

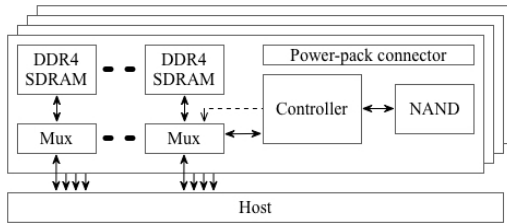


그림 1. NVDIMM의 구조도

Fig. 1 Block diagram of NVDIMM

지는 DRAM과 비휘발성인 NAND 플래시를 이용하여 DRAM과 동일한 읽기/쓰기 성능을 가지면서 비휘발성인 Non-Volatile Dual In-line Memory Module (NVDIMM)이 제안되었다.

NVDIMM은 평상시 DRAM을 이용하여 호스트의 데이터 접근을 처리한다. 동작 중 전원이 상실되면 DRAM의 모든 데이터를 NAND 플래시에 저장하고 전원이 복구되면 NAND 플래시에 저장된 데이터를 DRAM으로 복원한다. 이러한 과정을 통해 DRAM의 성능을 가지는 비휘발성 메모리로 동작한다.

메모리 표준을 제정하는 Joint Electron Device Engineering Council (JEDEC)에서 NVDIMM의 표준화 규약[1]이 제정되었고 Micron 사 등에서 상용 제품[2]이 출시되었지만 NVDIMM의 최적화 연구는 많이 이루어지지 않았다.

NVDIMM의 동작은 DRAM에 저장된 데이터에 대한 고려 없이 데이터의 백업과 복원이 이루어진다. 이에 따라 Micron사에서 출시된 NVDIMM의 경우 전원 상실 후 데이터 저장과 전원 공급 재개 후 데이터 복원 모두 수십 초의 시간이 소요된다.

본 연구에서는 NVDIMM의 동작에 대해 분석하고 NVDIMM에 존재하는 데이터를 고려한 최적화의 기법을 제시한다. 구체적으로 DRAM에서 NAND 플래시로 데이터 백업 시 압축 기법을 사용하여 플래시에 저장되는 데이터양을 줄이고자 한다. 이러한 데이터양의 절감은 저장에 소요되는 시간을 감소시킬 뿐만 아니라 복원에 소요되는 시간도 함께 감소시킨다. 실제 시스템의 메모리 데이터를 이용하여 시뮬레이션을 실시한 결과 제안된 기법을 통해 최대 72.6%의 백업 시간과 복원 시간이 감소되었다.

본 논문의 구성은 II장에서 NVDIMM의 구조와 동작 방식에 대해 설명한 후 백업과 복원 과정에서 병목 현상을 발생시키는 요인에 대해 분석한다. 이후 실험을 통해 제안한 기법의 효과를 제시하고 이에 대한 분석을 진행한다. 마지막으로 III장에서 결론을 맺는다.

II. 본 론

1. NVDIMM 구조 및 동작 방식

주기억장치용 메모리 슬롯에 연결되는 비휘발성 메모리인 NVDIMM은 Storage Networking Industry Association (SNIA)에서 주도적으로 제안되었다. NVDIMM은 목표로 하는 특성에 따라 N, F, P 방식이 제안되었다. DRAM의 용량과 동작 속도를 가지며 플래시를 이용하여 비휘발성 특성을 가지는 N 방식, NAND 플래시의 용량을 가지지만 속도가 느린 F 방식, 두 방식의 장점을 갖춘 P 방식이 제안되었다. N 방식의 경우 JEDEC에서 표준화되어 제품이 출시 중인 반면 F 방식은 제조업체의 기준에 따라 동작하고 P 방식은 JEDEC을 통해 표준화가 진행 중에 있다. 본 논문에서는 현재 활용되고 있는 N 방식에 대한 기법을 제시하고자 한다.

NVDIMM-N 방식은 기존의 battery-backed SRAM과 유사하게 동작한다. 그림 1은 NVDIMM의 구조를 보여주고 있다. 일반적인 상황에서는 DDR4 SDRAM DIMM과 같이 동작한다. 시스템의 전원이 상실되면 내부에 존재하는 제어기가 MUX의 제어 신호를 변경하여 DRAM과 호스트 간의 데이터 연결을 끊는다. 마지막으로 DRAM에 존재하는 모든 데이터를 비휘발성 장치인 NAND 플래시에 저장한다. 이 과정 동안 필요한 에너지는 별도의 전원 케이블로 연결된 슈퍼 캐패시터 모듈 등에 의해 공급된다. 시스템의 전원 공급이 재개되면 NAND 플래시에 저장된 모든 데이터를 DRAM에 복원한 뒤 MUX 제어 신호를 변경하여 호스트와 데이터 연결을 복원한다.

NVDIMM-N 규약에서 호스트 시스템과의 정보 교환은 SAVE_n 핀만으로 이루어진다. 이것이 의미하는 바는 NVDIMM의 백업과 복원 동작이 호스트 시스템과의 정보교환 없이 이루어진다는 것이다. 따라서 NVDIMM 내부에 존재하는 제어기는 DRAM의 주소공간 중 어느 부분이 유효한 데이터인지 판단할 수 없다. 그러므로 NVDIMM의 백업은 DRAM의 모든 주소 공간에 대해 이루어질 수밖에 없다.

2. NVDIMM의 동작 분석

NVDIMM의 동작을 분석하기 위해 먼저 DDR4 SDRAM과 NAND 플래시의 동작을 분석하였다.

2.1 DDR4 SDRAM의 동작

DDR4 SDRAM에 존재하는 데이터에 접근하기

위해서는 먼저 접근하고자 하는 행(Row)를 선택하여 해당 행의 데이터를 모두 행 버퍼(Row buffer)로 가져오는 Activation 동작을 수행하여야 한다. 행 버퍼의 크기는 모듈의 용량과 데이터 핀 수에 따라 512~2,048 바이트 중 결정된다. 즉, 행 버퍼에 존재하는 모든 데이터에 대해 읽기 또는 쓰기 동작이 이루어지는 것이 아니라 일부 열을 선택하여 읽기 또는 쓰기 동작을 수행한다. 설정에 따라 일정 시간 후 두 사이클 또는 네 사이클 동안 총 4번 또는 8번의 데이터 입출력이 연속적으로 이루어진다. 다른 행에 존재하는 데이터를 접근하거나 일정 시간 후에는 행 버퍼의 내용을 해당 행에 다시 기록하는 Precharge 동작을 수행하여야 한다.

읽기/쓰기 동작 시작 후 실제 데이터 입출력이 이루어지기까지 지연시간이 발생하며 이전의 행 주소와 다른 행 주소의 데이터에 대한 읽기/쓰기 동작이 이루어질 경우 지연시간이 증가한다. 이 시간을 줄이고자 여러 개의 뱅크(Bank)를 나누어 서로 다른 동작을 동시에 수행할 수 있도록 하였다. DDR4 SDRAM의 경우 최대 16개의 뱅크를 가지고 있어 데이터 읽기/쓰기가 끊임없이 이루어질 수 있다.

2.2 NAND 플래시의 동작

NAND 플래시의 주소는 블록(Block), 페이지(Page), 열(Column)으로 구성된다. NAND 플래시는 읽기와 쓰기에 제한이 없는 DRAM과 달리 한번 기록(Program)된 페이지는 새로운 데이터를 쓰기 전에 지우기(Erase) 동작이 이루어져야 한다. 또한, 데이터 읽기와 쓰기는 페이지 단위로 이루어지는데 비해 지우기는 블록 단위로 이루어진다. 이와 같은 제약 사항과 동작 단위의 불일치, 소자의 수명 제한으로 인해 NAND 플래시를 효율적으로 관리하기 위한 플래시 관리 정책(Flash translation layer)이 제안되었다.

데이터를 읽기 위해서는 먼저 NAND 플래시의 어레이에서 한 페이지의 데이터를 NAND 플래시 내부에 존재하는 캐쉬 레지스터(Cache register)로 옮기는 과정이 필요하다. 여기에 소요되는 시간이 길기 때문에 한 페이지의 데이터를 캐쉬 레지스터에서 출력하는 동안 다음 페이지 데이터를 별도의 데이터 레지스터(Data register)로 옮기는 Read page cache sequential 명령이 지원된다. 해당 명령을 통해 NAND 플래시 어레이에서 데이터를 읽는 동작과 데이터 전송 동작을 병렬적으로 수행할 수 있다. 이와 비슷하게 쓰기 시에는 캐쉬 레지스터에 데이터를 옮긴 후 명령어를 통해 NAND 플래시 어

레이에 데이터를 기록한다. Read page cache sequential 명령과 마찬가지로 데이터 레지스터를 이용하여 데이터 입력과 기록 동작을 동시에 수행할 수 있는 Program page cache 명령이 존재한다. 여기에 더해 데이터 처리량(Throughput)을 높이고자 짝수와 홀수 플레인(Plane)이 동시에 NAND 플래시 어레이로부터 읽거나 쓸 수 있는 멀티 플레인 동작(Multi-plane operation)이 지원된다.

이러한 명령을 통해 NAND 플래시 어레이에 병렬 동작을 지원한 결과 데이터 전송의 병목현상을 발생시키는 동작이 기존의 어레이에 대한 접근 동작에서 데이터 입출력 동작으로 변경되었다. 이를 해결하기 위해 기존에 지원되던 비동기식 입출력에 더해 동기식 입출력이 지원되기 시작하였다. 최근 페이지의 크기가 증가함에 따라 동기식 입출력만으로는 원활한 데이터 입출력을 수행할 수 없어 한 사이클에 2번의 데이터 입출력이 이루어지는 더블 데이터 레이트(DDR, Double Data Rate) 방식을 지원한다.

2.3 NVDIMM의 동작

본 논문에서는 상용으로 출시된 Micron 사의 8GB NVDIMM 모듈인 MTA18ASF1G72PF1Z[2]을 참고하여 DRAM과 NAND 플래시가 각각 DDR4-2400와 NV-DDR2 Mode 6로 동작하는 NVDIMM의 동작을 분석하였다. NAND 플래시의 페이지 크기는 16KB이며 두개의 플레인으로 구성된 한 개의 논리 단위(Logical unit, LUN)를 갖추고 있다. JEDEC 표준에 정의된 NVDIMM의 물리적 크기는 일반적인 DDR4 SDRAM DIMM과 동일한 것에 비해 MUX와 제어기, NAND 플래시 등이 추가되었다. 이로 인해 한 개의 NAND 플래시만 사용되는 구조가 일반적이다.

다양한 주소에 대한 쓰기 동작이 빈번하게 발생하는 저장장치로 다르게 NVDIMM의 백업 및 복원 시 각각 NAND 플래시 쓰기 동작과 읽기 동작만 발생하며 순차적인 데이터 접근이 이루어진다. 또한 데이터 복원 후 백업까지의 시간을 이용하여 지우기 동작을 수행할 수 있으므로 NAND 플래시를 저장장치로 사용할 때와 달리 복잡한 플래시 관리 정책이 필요하지 않다.

백업 과정은 제어기가 DRAM의 모든 주소공간에 존재하는 데이터를 읽어 NAND 플래시에 저장하는 과정이다. 백업이 이루어지는 동안 한정된 용량의 에너지를 통해 전원 공급이 이루어지기에 성능을 최대한 높여야 한다. 따라서 이미 열려있는

행 버퍼에서 데이터를 가져오는 동안 다른 뱅크에 대한 Activation 동작과 Precharge 동작을 수행하여 Activation과 Precharge에 소요되는 시간을 감추는 방식으로 동작한다. 이를 통해 데이터 핀 당 2,400 Mb/s, NVDIMM 모듈 당 21.6 GB/s의 데이터 전송 속도 달성이 가능하다[3].

DRAM으로부터 가져온 데이터는 NVDIMM 제어기 내부에 존재하는 버퍼에 임시 저장된 후 NAND 플래시로 옮겨져 기록된다. 표1은 NVDIMM 동작 분석에 사용된 NAND 플래시의 성능을 보여주고 있다. 고속 데이터 전송 규약인 NV-DDR2 Mode 6을 지원하는 NAND 플래시는 최대 333 MB/s의 데이터 전송이 가능한 반면 NAND 플래시 어레이에 대한 기록속도는 멀티 플레인 동작과 Program page cache 명령을 모두 사용하더라도 약 34 MB/s이 최대이다. 따라서 페이지 기록과정이 병목현상을 발생시키며 DRAM 읽기와 NAND 플래시의 캐쉬 레지스터로 데이터를 옮기는데 소요되는 시간은 모두 가려진다. NVDIMM 제어기 내의 버퍼공간은 제한되어 있기 때문에 DRAM에서의 읽기를 멈추는 방식으로 속도조절이 이루어진다.

이에 비해 복원 과정의 병목현상은 다른 곳에서 발생한다. Read page cache sequential 명령을 사용할 경우 표1에 제시된 바와 같이 NAND 플래시 어레이에서 한 페이지의 데이터를 읽기위해 소요되는 시간은 평균 29 μ s으로 약 560 MB/s의 데이터 읽기가 가능하다. 반면 캐쉬 레지스터에서 NVDIMM 제어기로의 데이터 전송속도는 최대 333 MB/s이다. DRAM의 쓰기 시간은 읽기 시간과 동일하므로 NAND 플래시에서 NVDIMM 제어기로의 데이터 이동이 병목현상을 발생시킨다.

3. 데이터 압축을 통한 NVDIMM 최적화

위에서 살펴본 바와 같이 현재 NVDIMM은 구조적으로 NAND 플래시가 병목현상을 발생시키며 백업과 복원에 소요되는 시간은 데이터양에 비례한다. 따라서 NAND 플래시에 읽고 쓰는 데이터양을 줄일 수 있다면 백업과 복원 시간도 줄일 수 있다. 본 논문에서는 압축을 이용하여 데이터양을 줄이는 기법을 제안하고자 한다. 먼저 본 논문에 적용하고자 하는 압축 알고리즘에 대해 살펴본 후 적용 방법에 대해 설명한다.

3.1 부하가 적은 데이터 압축 알고리즘

전원이 상실된 후의 백업 과정은 NVDIMM 내에 존재하는 제어기에 의해 진행된다. 따라서 데이터

표 1. NAND 플래시의 특성[4]

Table 1. NAND Flash feature

Interface	NV-DDR2 Mode6
Maximum throughput	333 MB/s
Page size	16,384 bytes
Maximum page read time	105 μ s
Typical page program time	1,260 μ s
Typical cache read time	29 μ s
Typical cache program time	1,030 μ s

압축에 사용할 수 있는 자원이 제한되며 압축에 많은 시간이 소요될 경우 압축 알고리즘이 병목현상을 발생시킬 수 있다. 그러므로 부하가 적은 압축 알고리즘이 적절하다.

데이터 전송량을 줄이거나 제한된 용량을 효율적으로 사용하기 위해 압축 알고리즘이 컴퓨터 시스템의 많은 부분에 적용되고 있다. 최하위 수준의 캐쉬(Last-level cache)[5,6]나 주기억장치[7]에서도 압축 알고리즘이 적용된 연구들이 발표되었다. 해당 장치들은 빠른 응답속도가 요구되고 많은 데이터 접근이 이루어지기 때문에 시간 및 에너지 부하가 적은 알고리즘이 필요하다. 이러한 요구조건은 NVDIMM에 사용하고자 하는 압축 알고리즘의 요구조건과 동일하다. 따라서 본 논문에서는 최하위 수준의 캐쉬나 주기억장치에 적용되는 압축 알고리즘을 사용하고자 한다.

최하위 수준의 캐쉬에 사용하는 대표적인 압축 알고리즘으로 Base-delta-immediate (BDI)[5]와 Frequent pattern compression (FPC)[6]이 있다. BDI는 하나의 압축 단위를 작은 블록을 나누어 대푯값과 대푯값으로부터의 차이를 이용하여 데이터를 표현하는 압축 알고리즘이다. 따라서 모든 데이터가 0 또는 1로 동일하거나 데이터의 변화가 적은 경우 좋은 압축률을 보인다. FPC는 모두 0이거나 모두 1과 같이 데이터에서 자주 나타나는 유형을 파악하여 해당 유형이 나타날 경우 압축을 진행하는 알고리즘이다. BDI와 마찬가지로 하나의 압축 단위를 작은 블록으로 나누어 대표유형과 비교한 후 값이 일치하면 적은 정보를 이용하여 표현하는 방식으로 압축한다. 두 압축 알고리즘은 압축 및 해제에 1~5 사이클이 소요되며 간단한 회로를 통해 구현이 가능해 NVDIMM 최적화를 위한 압축 알고리즘의 요구조건에 부합한다.

3.2 NVDIMM 최적화 기법의 동작

먼저 DRAM에서 NAND 플래시로 데이터 전송이 이루어지는 백업과정에 대한 최적화 기법의 동작을 살펴보고자 한다. 앞에서 논의한 바와 같이 DRAM 읽기는 모든 주소공간에 대해 연속적인 읽기(Sequential read)가 이루어진다. 최대 전송속도는 한 번의 읽기 명령에 대해 8개의 데이터 전송이 이루어질 때 달성된다. 여러 개의 DRAM에 대한 동시 접근이 이루어지는 NVDIMM의 데이터폭은 72 비트이므로 한 번의 읽기 명령을 통해 ECC를 포함, 72 바이트의 데이터가 DRAM으로부터 NVDIMM 제어기로 옮겨진다. 이 크기는 한 캐쉬라인의 크기와 비슷하므로 해당 단위로 압축을 진행한 후 버퍼에 저장한다.

NAND 플래시의 캐쉬 레지스터로 데이터를 이동시킨 후 내부 어레이에 기록을 시작하기까지 시간 간격은 최대한 작게 정되어있지 않으므로 적절한 양의 데이터가 버퍼에 모이면 데이터 이동을 시작한다. 압축된 데이터를 기록하더라도 NAND 플래시 어레이에 동시에 기록할 수 있는 데이터양은 최대 2페이지로 동일하므로 최대 기록속도는 이전과 같이 약 34 MB/s이다. 하지만 기록하여야 하는 데이터 총량은 압축으로 인해 줄었기 때문에 백업에 소요되는 시간은 감소한다.

복원과정도 데이터 압축을 적용하기 전과 비슷하다. NAND 플래시 어레이로부터 Read page cache sequential 명령을 통해 압축된 데이터를 가져 온 뒤 NVDIMM 제어기 내의 버퍼로 이동시킨다. 한번의 Read page cache sequential 명령을 통해 어레이로부터 읽어오는 동작과 NAND 플래시에서 NVDIMM의 제어기로 옮기는 동작은 데이터 압축 여부에 관계없이 이루어진다. 즉, 이전과 동일하게 NAND 플래시에서 NVDIMM 제어기까지의 데이터 이동이 병목현상을 발생시킨다. 제어기 내에 저장된 데이터는 압축 해제 과정을 거쳐 DRAM에 저장된다. 백업 과정과 마찬가지로 전체 메모리 주소공간을 복원하기 위해 NAND 플래시로부터 가져와야 하는 데이터 총량이 압축으로 인해 줄어들어 복원에 소요되는 시간이 감소한다.

만약 압축 및 해제 시간이 길면 해당 과정이 병목현상을 일으킬 수 있다. 하지만 제안한 기법에서 사용한 BDI와 FPC 압축 알고리즘은 3.1절에서 분석한 바와 같이 64 바이트의 압축 및 해체에 최대 5 사이클이 소요된다. 이는 표1에 제시된 바와 같이 각각 29 μ s과 1,030 μ s 소요되는 페이지 캐쉬 읽기 및 쓰기에 비해 매우 적다. 또한, 백업 과정에

표 2. 실험 데이터의 메모리 사용률 및 압축률
Table 2. Memory usage and compression ratio

	Memory usage	compression ratio	
		BDI [5]	FPC [6]
Case 0	0.325	0.378	0.395
Case 1	0.438	0.274	0.315
Case 2	0.600	0.819	0.431
Case 3	0.863	0.680	0.634

서는 메모리의 전 영역에 대한 읽기를 수행하여야 하며 반대로 복원 과정에서는 쓰기를 수행하여야 한다. 따라서 한 페이지의 데이터를 플래시에 쓰거나 읽는 동안 다른 페이지의 데이터에 대해 압축 또는 해제 동작을 수행할 수 있다. 그러므로 제안한 기법에서 추가된 압축 및 해제 과정은 시스템의 처리량에는 영향을 미치지 않는다. 이와 같이 압축 및 해제에 소요되는 시간을 가릴 수 있으므로 압축 및 해제에 의한 시간 부하는 마지막 페이지에 대해서만 유효하며 백업 및 복원시간에 견주어 볼 때 무시할 수 있는 수준이다.

4. 실험

4.1 실험 환경

본 논문에서 제안한 데이터 압축을 이용한 NVDIMM 최적화 기법의 효과는 임의의 시점에서 주기억장치에 저장된 메모리의 내용을 추출한 뒤 해당 데이터를 백업 및 복원하는 시뮬레이션을 통해 검증하였다. 메모리 내용을 추출한 시스템은 Intel i7-3770 프로세서를 장착하고 있으며 주기억장치 용량은 8 GB이다. 윈도우7을 운영체제로 사용하고 있으며 각 실험에 사용된 응용 프로그램은 이후에 설명하고자 한다.

표2는 이렇게 추출된 데이터의 메모리 사용률과 압축률을 보여주고 있다. 여기서 압축률은 압축하지 않은 데이터양 대비 압축 후의 데이터양으로 낮을수록 더 많이 압축된 것을 의미한다. Case 0은 운영체제의 기동을 마치고 기본 응용 프로그램이 수행되기 전에 주기억장치 전체의 내용이며 Case 1은 기본 응용 프로그램들이 수행된 이후의 주기억장치 전체 내용이다. Case 2는 오피스와 웹 브라우저를 사용하였을 때의 주기억장치 전체 내용이며 Case 3은 Case 2에 사용된 응용 프로그램 이외에 동영상 재생과 설계자동화 툴을 사용하여 메모리 사용률이 증가한 뒤의 주기억장치 전체 내용이다.

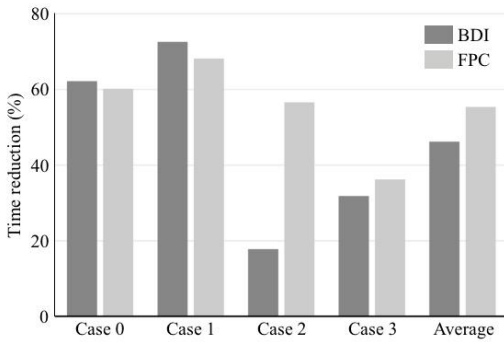


그림 2 제안한 최적화 기법의 시간 감소율
Fig. 2 Time reduction results

3.1절에서 설명한 바와 같이 BDI 압축 알고리즘의 경우 압축단위 내에서 데이터 변화가 작을 때 높은 압축률을 보이는 반면 데이터 변화가 크다면 압축을 수행할 수 없다. 따라서 압축단위 내의 데이터가 모두 0이거나 1인 비율이 높을 경우 높은 압축률을 보인다. 실험에 사용된 데이터를 분석해보면 압축단위 내의 데이터가 모두 0이거나 1인 비율은 case0부터 case3까지 각각 60.3%, 71.5%, 8.3%, 22.8%로 해당 비율이 높을수록 압축이 잘 되는 것을 확인할 수 있다. 반면 FPC의 경우 4바이트 단위로 압축을 진행하여 동작 시나리오에 따른 압축률이 차이가 적다.

NVDIMM은 앞에서 동작 분석 시 제시한 바와 같이 DDR4-SDRAM은 DDR4-2400으로 동작하며 NAND 플래시의 특성은 표1에 제시된 바와 같다. NAND 플래시의 데이터 처리 속도를 최대화하기 위해 백업 시에는 Multi-plane program page cache 명령을, 복원 시에는 Read page cache sequential 명령을 사용한 동작에 대해 데이터 압축을 통한 최적화 기법을 적용하였다.

4.2 실험 결과 분석 및 논의

그림2는 데이터 압축을 사용하지 않는 경우 백업 및 복원에 소요되는 시간 대비 데이터 압축을 적용한 경우 소요되는 시간의 감소율을 보여주고 있다. 막대의 높이가 높을수록 시간 감소 효과가 큰 것을 나타낸다. 짙은 회색의 막대그래프는 압축 알고리즘으로 BDI를 사용한 경우를 나타내고 있으며 옅은 회색의 막대그래프는 FPC를 압축 알고리즘으로 사용한 경우를 보여주고 있다. 데이터 백업에 걸리는 시간이 복원에 걸리는 시간에 비해 길지만 압축을 사용하지 않은 경우를 기준으로 시간 감소율

을 계산할 경우 백업과 복원 과정의 시간 감소율은 동일하다. 따라서 제안한 기법의 효과를 하나의 그래프로 제시하였다.

앞에서 분석한 바와 같이 데이터 백업 및 복원 과정에서 병목현상을 일으키는 요인은 NAND 플래시의 데이터 처리 속도이다. 백업과 복원 동작으로 인해 데이터가 변경되어서는 안 된다. 따라서 백업 과정에서 압축에 의해 감소된 데이터양과 복원 과정에서 압축에 의해 감소된 데이터양은 동일하며 백업과 복원 과정의 시간 감소를 또한 동일하다.

메모리 데이터별 시간 감소율을 비교해보면 Case 1의 경우처럼 압축이 효과적으로 이루어질수록(압축률이 낮을수록) 더 높은 시간 감소율을 보여준다. 이러한 실험 결과는 앞에서 분석한 NVDIMM의 동작 특성과 일치한다.

BDI 압축 알고리즘은 데이터의 변화가 적은 경우 압축에 유리하다. 따라서 Case 0과 Case 1 같이 메모리 사용률이 낮고 운영체제가 기동한 직후의 데이터에서 시간 감소율이 큰 반면 메모리 사용률이 높아지는 Case2와 Case 3의 경우 압축이 잘 이루어지지 않아 시간 감소율이 줄어든다. 압축 알고리즘으로 BDI를 이용하는 경우 백업과 복원에 소요되는 시간이 평균 46.2%, 최대 72.6% 줄어든다. 반면 표1에서 확인할 수 있듯이 데이터에 따른 압축률의 변화가 적은 FPC를 압축 알고리즘으로 사용할 경우 백업과 복원에 소요되는 시간은 평균 55.6%, 최대 68.5% 줄어들며 메모리 사용률의 영향을 덜 받는다.

본 연구에서 제안한 압축을 이용한 최적화 기법은 DRAM의 읽기 및 쓰기의 최소 단위 데이터별로 적용되며 기존의 백업과 복원 과정이 데이터 내용에 기반 한 동작을 수행하지 않기 때문에 NVDIMM 제어기의 다른 동작에 영향을 미치지 않는다. 뿐만 아니라 압축이 적용되지 않은 NVDIMM과 같은 횡수의 백업 및 복원이 이루어질 경우 같은 크기의 NAND 플래시에 적은 양의 데이터를 기록하므로 쓰기 분산효과를 발생시켜 NVDIMM의 수명을 증가시키는 부가효과가 있다.

전원 상실 후 DRAM의 데이터를 NAND 플래시로 백업하는 시간의 절감은 에너지 절감으로 연결된다. 이 과정 동안의 에너지는 별도로 연결된 전원 공급 모듈을 통해 이루어진다. 따라서 본 연구에서 제안된 최적화 기법은 전원 공급 모듈의 용량도 최적화 할 수 있는 여지를 제공한다. 다만 전원 공급 모듈의 용량은 최악의 상황에 대비하여 설계된다. BDI와 FPC 압축 알고리즘은 압축이 이루어지지 않

는 경우도 있기 때문에 최악의 경우 압축이 전혀 이루어지지 않을 수 있다. 따라서 압축을 통해 전원 공급 모듈의 최적화가 직접적으로 이루어지지 않는다. 만약 일정 수준의 압축률을 보장하는 압축 알고리즘이 있다면 전원 공급 모듈도 압축률에 비례하여 최적화 할 수 있다.

데이터 압축을 이용한 최적화 기법 이외에도 DRAM에 저장된 데이터에 바탕을 둔 다양한 최적화 기법이 적용될 수 있다. 페이지와 같은 특정 단위별로 변경 플래그(Dirty bit)를 두어 사용되지 않은 페이지는 백업 및 복원에서 제외하는 최적화 기법, 주기적으로 유효한 데이터 페이지를 NAND 플래시에 선제적으로 기록하여 백업 시 저장하여야 할 데이터양을 줄이는 기법 등이 DRAM에 저장된 데이터에 바탕을 둔 최적화 기법의 예시라 할 수 있다. 하지만 이러한 기법들을 적용하기 위해서는 DRAM에 대한 호스트의 접근을 추적, 기록할 필요가 있다. 따라서 DRAM이 아니라 NVDIMM 제어가 호스트의 데이터 접근 요청을 처리하도록 NVDIMM의 구조가 변경되거나 호스트와 정보교환이 가능하도록 변경되어야 한다. 이 경우 NVDIMM 제어가 관리하여야 할 정보가 늘어나므로 제어기 구현의 복잡도가 상승하게 된다.

IV. 결 론

본 논문에서는 DRAM의 동작 속도에 비휘발성 특성을 지니는 NVDIMM의 동작을 살펴보았다. DRAM과 NAND 플래시로 이루어진 NVDIMM은 호스트 시스템과의 정보교환이 제한되어 있기 때문에 시스템 전원 상실 시 DRAM에 존재하는 모든 데이터를 NAND 플래시에 저장한다. 본 논문에서는

NVDIMM 동작의 분석을 통해 DRAM의 읽기/쓰기 능력과 NAND 플래시의 읽기/쓰기 능력의 차이로 인해 병목현상이 발생함을 보였다. 마지막으로 데이터 압축기법을 이용하여 백업과 복원에 소요되는 시간을 최대 72.6% 감소시킬 수 있음을 보였다.

References

- [1] JEDEC, "JESD248 DDR4 NVDIMM-N Design Standard," 2016.
- [2] Micron, "DDR4 SDRAM NVRDIMM," 2015.
- [3] Micron, "DDR4 SDRAM," 2017.
- [4] Micron, "NAND Flash Memory - MLC+," 2014.
- [5] G. Pekhimenko, V. Seshadri, O. Mutlu, P. B. Gibbons, M. A. Kozuch, T. C. Mowry, "Base-Delta-Immediate Compression: Practical Data Compression for On-Chip Caches," Proceedings of the ACM 21st international conference on Parallel architectures and compilation techniques, pp.377-388, 2012.
- [6] J. Park, H.G. Lee, "Location-Aware Hybrid SLC/MLC Management for Compressed Phase-Change Memory Systems," IEMEK J. Emded. Sys. Appl., Vol. 11, No. 2, pp.107-116, 2016 (in Korean).
- [7] A.R. Alameldeen, D.A. Wood, "Frequent Pattern Compression: A Significance-Based Compression Scheme for L2 Caches," Technical Report 1500, Computer S University of Wisconsin-Madison, 2004.

Jaehyun Park (박재현)

Jaehyun Park received the B.S. in the Department of Electrical Engineering and Ph.D degree in the school of Electrical Engineering & Computer Science from Seoul National University, Korea, in 2006 and 2015, respectively. He is currently an Exchange Scholar with School of Electrical, Computer & Energy Engineering, Arizona State University, USA.

Research interest: Embedded system, low-power system, memory system.

Email: jaehyun.park.2@asu.edu

Hyung Gyu Lee (이형규)

Hyung Gyu Lee received the M.S. and Ph.D. degrees from the School of Computer Science & Engineering, Seoul National University, Korea, in 2001 and 2007, respectively. He was a Senior Engineer with Samsung Electronics, Korea, from 2007 to 2010, and a Post-Doctoral Research Fellow with the Georgia Tech., USA, from 2010 to 2012. He is currently an Assistant Professor with the School of Computer & Communication Engineering, Daegu University, Korea.

Research interest: Embedded system design, low power systems, energy harvesting, non-volatile memory/storage architecture

Email: hglee@daegu.ac.kr