

스토리지 클래스 메모리 기반 파일시스템의 일관성 유지 기법 비교 및 분석

이현구*, 김정훈, 엄영익
성균관대학교 정보통신대학
e-mail:{hyungoo*, myhuni20, yieom}@skku.edu

Analysis and Comparison of Consistency Control Mechanism for Storage Class Memory based File System

HyunKu Lee, Junghoon Kim, YoungIk Eom
College of Information and Communication Engineering,
Sungkyunkwan University

요 약

최근 기존 스토리지의 물리적 한계를 극복하기 위해 차세대 스토리지로 불리는 스토리지 클래스 메모리(SCM)의 연구 및 개발이 활발히 진행되고 있다. 이러한 SCM의 장점을 활용하기 위한 SCM 전용 파일 시스템 또한 많은 연구가 되고 있다. 하지만, 현재까지 연구되어온 SCM 전용 파일 시스템은 빠른 속도에만 중점을 두어 부분적인 일관성 지원 또는 특정한 환경에서의 시스템 성능 저하 등의 문제 요인이 있다. 본 논문에서는 SCM과 SCM 전용 파일 시스템을 알아보고 일관성 유지 기법의 비교 및 분석을 통하여 현재까지 연구된 일관성 유지 기법의 문제점을 파악하고 SCM에 최적화된 새로운 일관성 유지 기법의 방향을 모색한다.

1. 소개

스토리지 클래스 메모리(SCM)는 메모리의 속도를 지니고 스토리지로 사용할 수 있는 특성을 가진 메모리를 총칭하는 용어이다. 많은 회사에서 연구 및 개발이 되고 있으며 PCM, FeRAM, STT-MRAM 등이 상용화 단계에 돌입하였다 [1]. 향후 몇 년 안에 대용량의 SCM이 개발될 예정이고 현재 기존의 스토리지를 대체할 차세대 메모리로 주목받고 있다 [2]. 하지만 SCM의 장점을 활용하기 위해서는 기존의 스토리지 관리 방식에서 변경된 몇 가지 요인이 필요하다. 첫째로 DRAM과 비슷한 속도를 지닌 SCM을 스토리지로 사용하면 스토리지와 메모리 사이의 속도 차이를 줄이기 위해 존재하던 페이지 캐시가 오히려 시스템 성능 저하의 요인이 된다. 페이지 캐시와 스토리지 두 곳에 동일한 데이터를 관리해야 하기 때문에 쓰기량이 증가되기 때문이다. 두 번째로 SCM의 특징을 이용한 파일 시스템을 사용해야 한다. 기존의 파일 시스템은 디스크 또는 플래쉬 기반의 스토리지를 이용하기 위해 개발되어 SCM의 바이트 단위 접근, 제로 탐색 시간 등의 특징을 이용할 수 없다. 그러므로 SCM의 장점을 최대한 활용하기 위해선 SCM의 특징을 이용하는 파일 시스템을 사용해야 한다. 마지막으로 SCM의 장점 중 하나인 바이트 단위 데이터 접근을 이용하기 위해 SCM을 메모리 버스

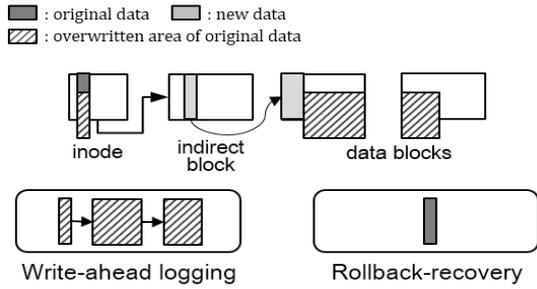
에 위치시켜야 한다. 이럴 경우 캐시 일관성을 위해 CPU 캐시와 SCM 간의 동기화가 필요하다. 현재 앞에서 고려한 SCM의 특징을 적용시킨 파일 시스템에 대한 연구가 진행되고 있으며 SCMFS, BPFS, Shortcut-JFS 등의 파일시스템이 제안되고 있다. 하지만 현재까지 개발된 파일 시스템은 SCM의 빠른 속도를 지원하기 위해 데이터 일관성을 무시하는 경향이 있고 페이지 캐시가 없는 환경에서는 시스템 성능이 저하되는 요인을 가지고 있다.

본 논문에서는 2장에서 일관성 유지 기법에 대해 알아보고 3장에서 SCM 전용 파일시스템의 일관성 유지 기법의 비교 및 분석을 통해 문제점을 분석하고 4장에서 새로운 일관성 유지 기법을 제안한다. 마지막으로 5장에서 본 논문의 결론과 함께 향후 연구 방안을 제시하도록 하겠다.

2. 일관성 유지 기법

파일 시스템에서 일관성 유지란 파일에 대한 쓰기 요청 진행 중 시스템 에러가 발생하였을 때 해당 데이터의 쓰기 명령이 재실행되거나 데이터가 쓰기 이전의 상태로 복구되어 데이터의 무결성이 항상 보장되는 것을 의미한다. 파일 시스템의 일관성 유지 기법은 저장하는 로그 데이터의 성격에 따라 Write-Ahead Logging(WAL)과 Rollback-Recovery(ROLLBACK)으로 나눌 수 있다. WAL은 스토리지의 데이터를 업데이트하기 전에 새로운 데이터를 로그 데이터로 저장한다. 데이터 업데이트 중

이 논문은 2013년도 정부(미래창조과학부)의 재원으로 한국연구재단-차세대 정보통신기술개발사업의 지원을 받아 수행된 연구임(No.2010-0020730)



(그림 1) 일관성 유지 기법의 로그 데이터 비교

시스템 에러가 발생할 경우 저장해 놓은 로그 데이터를 이용해 중단된 쓰기를 다시 수행하며 대표적인 예로 저널링과 웨도우 페이지가 있다. 이와 반대로 ROLLBACK은 스토리지의 데이터를 업데이트하기 전에 변경되는 원본 데이터를 로그 데이터로 저장 하며 데이터 업데이트 중 시스템 에러가 발생할 경우 저장해 놓은 로그 데이터를 이용하여 변경 이전의 데이터로 복구한다. 이 방식은 데이터베이스 시스템에서 많이 사용된다. 두 일관성 유지 기법은 로그 데이터의 속성이 다르기 때문에 로그 사이즈도 차이가 난다. 기본적으로 그림 1에서와 같이 변경되는 데이터만 저장하는 ROLLBACK 기법이 변경하는 모든 데이터를 저장하는 WAL의 기법보다 작다.

3. SCM 전용 파일 시스템의 일관성 유지 기법

SCMFS(Storage Class Memory File System) [3]는 메타데이터를 배열로 관리하며 가상 메모리에 데이터를 저장하는 SCM 전용 파일 시스템이다. 데이터 일관성 유지를 위해 MFENCE와 CLFLUSH를 이용하여 CPU 캐시의 연산순서를 보장 하고 메타데이터가 변경 될 때마다 CPU 캐시와의 동기화를 한다. 하지만 데이터 변경에 관한 일관성 유지 기법이 없어 스토리지의 데이터 업데이트 중 시스템 에러가 발생 되었을 경우 해당 파일의 데이터의 일관성을 보장할 수 없다. BPFs(Byte-addressable Persistent File System) [4]는 SCSP(Short-circuit Shadow Paging)을 사용해 기존의 Shadow Paging 기법 보다 데이터 쓰기 비용을 줄인 파일 시스템이다. epoch barrier 이름의 제안한 하드웨어를 이용하여 CPU 캐시와 스토리지와의 일관성을 유지한다. 하지만 BPFs를 현재의 시스템에서 사용하기 위해선 많은 CPU 캐시가 유발된다. 또한 SCSP를 이용한 데이터 쓰기 시 64bit 원자적 업데이트가 가능 할때까지 데이터 복사가 전과되므로 다른 일관성 유지 기법보다 데이터 쓰기 량이 많다. Shortcut-JFS [5]는 SCM의 특성을 이용하도록 수정한 저널링 기법을 사용한다. 저널링 영역에 저장한 한 블록의 데이터 크기가 블록 크기 반보다 작으면 기존의 저널링 기법과 마찬가지로 Checkpoint시 데이터를 기존의 블록에 덮어쓰기를 한다. 만약 반 보다 크면 덮어쓰지 않고 간접 포인터 변경을 통하여 저널링 영역의 데이터를 원본 데이터로 여긴다. 이런 방식을 통하면 데이터 쓰기 량을 줄일 수 있는 장점이

<표 1> SCM 전용 일관성 유지 기법 비교

구 분	SCMFS	BPFs	Shortcut-JFS
데이터 일관성 유지 기법	없 음	SCSP	저널링
CPU 캐시와 SCM 동기화 기법	MFENCE CLFLUSH (주기적)	epoch barrier	없 음
비 고	데이터 일관성을 보장하지 못함	하드웨어의 지원이 없으면 일관성 유지 비용이 높음	페이지 캐시가 필요함

있지만 페이지 캐시가 없는 환경에서는 CPU 캐시와의 동기화가 많이 발생하여 일관성 유지비용이 늘어나게 된다.

표 1 에서와 같이 현재까지 연구된 SCM 전용 파일 시스템은 일관성 유지 기법이 시스템 성능을 감소시키거나 완벽한 데이터 일관성을 유지하지 못하는 단점이 있다. 그러므로 SCM을 효율적으로 활용하기 위한 새로운 데이터 일관성 기법의 연구가 필요하다.

4. 결 론

현재까지 연구된 SCM 전용 파일시스템의 일관성 유지 기법들은 현재의 시스템에서 성능 저하와 데이터 일관성을 보장 못하는 단점이 있다. 특히 가장 많이 사용하고 있는 WAL 방식의 일관성 유지 기법은 새로운 데이터를 적용하기 전에 해당 내용을 따로 저장해야 하므로 페이지 캐시가 없는 환경에서 많은 양의 쓰기와 CPU 캐시 동기화를 유발한다. 또한 변경중인 내용을 사용자가 바로 인지할 수 없다. 만약 WAL방식 대신 ROLLBACK 방식을 사용한다면 로그 량을 줄일 수 있고 CPU 캐시와의 일관성 유지 비용을 감소시킬 수 있을 것이다. 향후에는 SCM을 스토리지로 사용하는 환경에서 ROLLBACK을 이용한 일관성 유지 기법의 구체화 하고 유지 비용이 적은 일관성 유지 기법을 연구할 예정이다.

참고문헌

[1] <http://www.mram-info.com/everspins-ceo-explains-companys-technology-and-business>
 [2] <http://www.mram-info.com/everspin-officially-announces-worlds-first-st-mram-chip-will-be-available-2013>.
 [3] Wu, X. and Reddy A "Scmfs: a file system for storage class memory" Proceedings of 2011 International Conference for High Performance Computing, Networking, Storage and Analysis
 [4] Condit, J. "Better I/O through byte -addressable, persistent memory?" Proceedings of the ACM SIGOPS 22nd symposium on Operating systems principles
 [5] Lee, E. Yoo, S. Jang, J.E. and Bahn, H. "Shortcut-jfs: A write efficient journaling file system for phase change memory." Mass Storage Systems and Technologies 2012