

메타데이터를 비휘발성 램에 유지하는 플래시 파일시스템에서 가비지 컬렉션 수행에 대한 효율성 평가

(Assessment of the Efficiency of Garbage Collection for the MiNV File System)

도인환[†] 최종무^{**}
(In Hwan Doh) (Jongmoo Choi)

이동희^{***} 노삼혁^{****}
(Donghee Lee) (Sam H. Noh)

요약 비휘발성 속성을 가지면서 램의 특성을 동시에 제공하는 비휘발성 램의 효과적인 활용을 목적으로, 모든 메타데이터를 비휘발성 램에서 관리하고 파일데이터만 낸드 플래시에서 관리하는 MiNV(Metadata in NVram) 파일시스템이 이미 제안된 바 있다. 본 연구에서는 MiNV 파일시스템에서 가비지 컬렉션 수행에 대한 효율성을 실험적으로 분석하고, 가비지 컬렉션의 효율성이 전체 파일시스템 성능에 미치는 영향에 대해서 살펴본다. MiNV 파일시스템은 동일한 가비지 컬렉션 기법을 적용하더라도 YAFFS보다

더 효율적으로 가비지 컬렉션을 수행한다. 성능 평가 결과에서 MiNV 파일시스템은 전체 낸드플래시에서 가용한 블록의 개수가 부족할 때 발생하는 공격적인 가비지 컬렉션의 발생 빈도를 줄임과 동시에 공격적인 가비지 컬렉션 수행 시점을 늦춤을 확인하였다. 이와 더불어, 실험 결과는 이러한 가비지 컬렉션에 대한 효율성이 파일시스템의 성능 향상에 기여함을 보여준다.

키워드 : 비휘발성 램, 플래시 메모리, 파일시스템, 메타데이터, 가비지 컬렉션

Abstract Non-volatile RAM (NVRAM) has both characteristics of nonvolatility and byte addressability. In order to efficiently exploit this NVRAM in the file system layer, we proposed the MiNV (Metadata in NVram) file system in our previous research. MiNV file system maintains all the metadata in NVRAM while storing file data in NAND Flash memory. In this paper, we experimentally analyze the efficiency for the execution of garbage collection in the MiNV file system. Also, we quantify the file system performance gains obtained from efficient garbage collection. Experimental results show that garbage collection on the MiNV file system executes more efficiently than on YAFFS even though these file systems adopt exactly the same garbage collection policy. Specifically, the MiNV file system invokes the aggressive garbage collection mechanism less frequently than YAFFS. Additionally, the MiNV file system postpones the first execution of the aggressive garbage collection mechanism in our experiments. From the experiments, we verify that the efficiency of garbage collection leads to performance improvements of the MiNV file system.

Key words : Non-volatile RAM, Flash memory, File system, Metadata, Garbage collection

· 이 논문은 2007년도 정부(과학기술부)의 재원으로 한국과학재단의 국가지정연구사업으로 수행된 연구임 (No. R0A-2007-000-20071-0)
· 이 논문은 제34회 추계학술대회에서 '메타데이터를 비휘발성 램에 유지하는 플래시 파일시스템에서 가비지 컬렉션 수행에 대한 효율성 분석'의 제목으로 발표된 논문을 확장한 것임

† 학생회원 : 홍익대학교 컴퓨터공학과
inhwando@hotmail.com

** 종신회원 : 단국대학교 컴퓨터학과 교수
choijm@dankook.ac.kr

*** 정회원 : 서울시립대학교 컴퓨터과학부 교수
dhi_express@uos.ac.kr

**** 종신회원 : 홍익대학교 정보컴퓨터공학부 교수
samhnoh@hongik.ac.kr

논문접수 : 2007년 12월 6일

심사완료 : 2008년 1월 24일

Copyright © 2008 한국정보과학회 : 개인 목적이나 교육 목적인 경우, 이 저작물의 전체 또는 일부에 대한 복사본 혹은 디지털 사본의 제작을 허가합니다. 이 때, 사본은 상업적 수단으로 사용할 수 없으며 첫 페이지에 본 문구와 출처를 반드시 명시해야 합니다. 이 외의 목적으로 복제, 배포, 출판, 전송 등 모든 유형의 사용행위를 하는 경우에 대하여는 사전에 허가를 얻고 비용을 지불해야 합니다.

정보과학회논문지: 컴퓨팅의 실제 및 레터 제14권 제2호(2008.4)

1. 서론

비휘발성 램은 비휘발성 속성을 가지면서 동시에 램의 특성을 제공하는 신개념의 메모리이다. 비휘발성 램은 다수의 주요 반도체 회사들에 의해서 앞다투어 개발되고 있으며, FRAM과 MRAM의 형태로 이미 시판되고 있거나 PRAM의 형태로 상용화를 눈앞에 두고 있다 [1-3]. 또한 PRAM 개발을 주도하고 있는 삼성이 2008년부터 이동 전화기기의 노어플래시를 PRAM으로 대체하겠다고 언론에 밝혀 비휘발성 램에 대한 관심이 높아지고 있다[4].

램과 비휘발성 매체의 특성을 동시에 보유한 비휘발성 램이 시스템 소프트웨어에 의해서 효과적으로 활용된다면 전체 시스템의 성능 향상을 기대할 수 있다. 실제로 본 연구에 선행하는 연구에서는 파일시스템의 모든 메타데이터를 비휘발성 램에서 관리하고 파일 데이

타만 낸드플래시에 저장하는 파일시스템인 MiNV(Meta-data in NVram) 파일시스템을 제안한 바 있다[5].

MiNV 파일시스템은 메타데이터를 낸드플래시에 저장하지 않기 때문에 낸드플래시를 효과적으로 활용하게 되어, 가비지 컬렉션을 수행함에 있어서도 기존의 대표적인 플래시 파일시스템인 YAFFS보다 더 높은 성능을 제공할 것으로 판단된다[6]. 이에 본 연구에서는 MiNV 파일시스템에서 가비지 컬렉션 수행에 대한 효율성을 실험적으로 분석하고 정량화하고자 한다. 또한 성능 평가를 통해서 가비지 컬렉션의 효율성이 전체 파일시스템 성능에 미치는 영향에 대해서도 살펴본다.

본 연구의 성능 평가에서는 MiNV 파일시스템과 YAFFS가 다수의 가비지 컬렉션을 수행할 만큼 충분히 큰 워크로드에 대해서 각각의 파일시스템 성능을 관찰하였다. 성능 평가 결과는 MiNV 파일시스템에서 전체 낸드플래시에서 가용한 블록의 개수가 부족할 때 발생하는 공격적인 가비지 컬렉션 횟수가 적어짐과 동시에 공격적인 가비지 컬렉션의 시작 시점이 지연됨을 보여준다. 또한 파일시스템의 성능 평가 과정에서 발생하는 전체 가비지 컬렉션 수행 과정에서 유효한 페이지의 총 복사 횟수도 줄어든다는 것을 실험적으로 알 수 있다.

본 논문의 구성은 다음과 같다. 먼저 2절에서는 본 논문에서 고려하고 있는 MiNV 파일시스템과 YAFFS에 대해서 간략하게 살펴본다. 3절에서는 MiNV 파일시스템의 가비지 컬렉션 기법과 그의 효율성에 대해서 살펴보고, 4절에서는 성능 평가 결과를 보면서 가비지 컬렉션이 전체 파일시스템의 수행성능에 미치는 영향에 대해서 논의한다. 마지막으로 5절에서 결론을 내린다.

2. 배경 지식

본 절에서는 논문 전반에 걸쳐서 언급되는 YAFFS와 MiNV 파일시스템에 대해서 간략하게 설명한다.

2.1 MiNV 파일시스템

MiNV 파일시스템은 비휘발성 램을 활용하는 낸드플래시 기반의 파일시스템이다[5]. MiNV 파일시스템은 모든 메타데이터를 비휘발성 램에 유지 및 관리하며, 파일데이터만 낸드플래시에 저장한다. MiNV 파일시스템은 파일데이터와 비교해서 상대적으로 작은 크기이면서 빈번하게 업데이트되는 메타데이터를 비휘발성 램에서 관리함으로써 기존의 플래시 파일시스템에 비해서 수행 성능과 마운트 속도 측면에서 높은 성능 향상을 기대할 수 있다.

비휘발성 램은 바이트 단위의 무작위적 접근이 가능할 뿐만 아니라 낸드플래시보다 빠른 접근 속도를 제공하므로 메타데이터에 대한 빈번한 업데이트는 MiNV 파일시스템의 성능 향상에 크게 기여한다. 또한 비휘발

성 램에 저장되어 있는 메타데이터들은 시스템의 재부팅 이후에도 여전히 비휘발성 램에 남아 있으므로 MiNV 파일시스템은 메타데이터를 낸드플래시와 같은 저장매체에 중복해서 관리할 필요가 없다. 따라서 메타데이터의 동시성 제공을 위해서 소요되는 비용을 줄일 수 있다. 더불어, MiNV 파일시스템의 메타데이터는 이미 구조화된 상태로 항상 비휘발성 램에 존재하기 때문에 즉각적인 파일시스템 마운트가 가능하다.

2.2 YAFFS

YAFFS는 임베디드 시스템 개발자들에 의해 널리 알려져 있는 대표적인 낸드플래시 기반 파일시스템이다[6]. YAFFS는 파일데이터와 메타데이터를 모두 낸드플래시에 저장한다. 메타데이터는 낸드플래시의 데이터 영역에 저장되는 파일의 헤더와 스페어 영역에 저장되는 파일의 오프셋 값으로 구성된다.

YAFFS는 마운트시에 낸드플래시에 흩어져 있는 메타데이터를 수집하고 램에서 수집된 메타데이터를 구조화한다. 이로 인해서 초기 마운트시간이 길어진다. 또한 메타데이터가 램과 낸드플래시에 중복해서 존재하므로 동시성을 유지하는데 많은 비용이 든다.

3. MiNV 파일시스템의 가비지 컬렉션

본 연구에서는 MiNV 파일시스템이 메타데이터를 낸드플래시에 저장하지 않고 비휘발성 램에서만 관리함으로써, 가비지 컬렉션 수행 측면에서 얻을 수 있는 효율성에 대해 살펴본다.

3.1 가비지 컬렉션 기법

MiNV 파일시스템의 가비지 컬렉션은 YAFFS의 가비지 컬렉션 기법과 정확하게 일치한다. YAFFS의 가비지 컬렉션 기법이 우수해서 그대로 채택한 것이 아니라, 메타데이터를 비휘발성 램에 저장함으로써 얻을 수 있는 가비지 컬렉션상의 효과를 보다 정확하게 파악하기 위함이다.

MiNV 파일시스템과 YAFFS의 가비지 컬렉션 수행 절차는 다음과 같다. 먼저, 낸드플래시에 쓰기 연산을 수행할 때마다 가비지 컬렉션을 수행할지 여부를 판단한다. 만약 7개 이상의 비어있는 가용한 블록이 존재한다면 소극적으로, 그렇지 않으면 공격적으로 가비지 컬렉션을 수행한다. 소극적인 가비지 컬렉션은 유효한 페이지가 3개 이하인 블록을 소거할 대상으로 선정하지만 공격적인 가비지 컬렉션은 모든 블록 중에서 가장 유효한 페이지를 적게 포함하는 블록을 소거할 대상으로 선정한다. 가비지 컬렉션을 위한 소거 대상 블록이 선정되면 해당 블록에 존재하는 유효한 페이지들을 다른 곳으로 복사한 다음에 해당 블록을 소거한다.

3.2 가비지 컬렉션의 효율성

YAFFS는 자체적으로 페이지 읽기와 쓰기 캐시를 가진다. 이는 안정성을 저해한다는 이유로 MiNV 파일시스템에서는 사용하지 않는다. 같은 수준의 안정성에서 성능 평가를 진행하기 위해 본 실험에서 YAFFS는 자체 캐시를 사용하지 않도록 설정하였다. 또한 안정성의 이유와 함께 버퍼캐시의 성능이라는 변수를 고려하지 않기 위해서 MiNV 파일시스템과 YAFFS는 버퍼캐시를 사용하지 않도록 설정하였다. 마지막으로 정확한 성능 측정을 위해서 낸드플래시의 오류상황에서도 오류 메시지를 출력하지 않도록 설정하였다.

본 실험에서 사용하는 파일시스템 워크로드로는 Gal과 Toledo가 제안한 TFFS(Transaction Flash File System)를 평가하기 위해서 사용되었던 워크로드 시나리오를 사용하였다[7]. 본 연구에서는 TFFS 워크로드 시나리오를 그대로 재현하는 TFFS 벤치마크 툴을 제작하였다.

4.1.2 하드웨어 환경

성능 평가를 위해서 사용된 임베디드 개발보드는 FALINUX사의 EZ-M28 보드이다[8]. 성능 평가 과정에서 사용된 낸드플래시 메모리의 용량은 32MB이다. MiNV 파일시스템이 필요로 하는 비휘발성 램은 실험상의 편의를 위해서 메인 메모리 중에서 1MB를 비휘발성 램으로 에뮬레이션 하였다.

4.2. 실험 결과

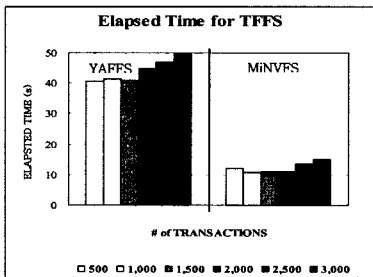
본 실험에서는 MiNV 파일시스템과 YAFFS가 다수의 가비지 컬렉션을 수행할 만큼 충분히 큰 워크로드에 대해서 MiNV 파일시스템과 YAFFS에 대한 성능평가를 실시한다. 본 실험에서는 TFFS 벤치마크가 제공하는 워크로드 시나리오 중에서 휴대전화 워크로드 시나리오를 사용해서 MiNV 파일시스템과 YAFFS의 성능을 평가한다.

TFFS 벤치마크의 휴대전화 워크로드는 한번의 트랜잭션 동안 300B로 고정된 크기의 3개 파일에 대해 각각 다수의 업데이트를 발생시키고, 2개의 파일에 대해서는

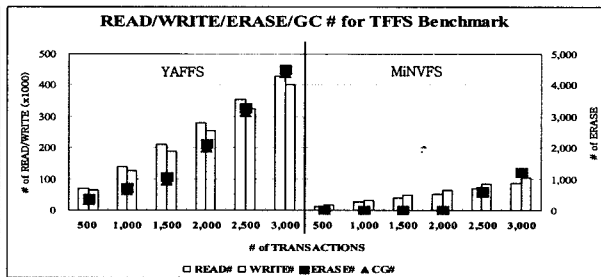
3회에 걸쳐 각각 파일 내용을 추가한다. 전자는 휴대전화의 통화 기록을 관리하는 모습을, 후자는 문자메시지가 계속해서 쌓여가는 모습을 흉내 낸다. 이 워크로드는 휴대전화와 더불어 작은 파일들을 저장하는 호출기나 문자 전송 기기와 같은 장치들의 동작을 대변할 수 있다.

그림 2는 TFFS 벤치마크의 트랜잭션 수를 증가시키면서 YAFFS와 MiNV 파일시스템이 각각 워크로드를 처리하기 위해서 나타내는 성능을 보여준다. 그림 2(a)는 YAFFS와 MiNV 파일시스템이 벤치마크를 3,000회까지 처리하는 데 소요된 시간을 500회 트랜잭션 단위로 보여준다. 그림 2(b)는 각각 YAFFS와 MiNV 파일시스템이 워크로드를 처리하면서 낸드플래시에 요청한 페이지 읽기, 쓰기, 그리고 블록 소거 연산의 횟수와 수행된 가비지 컬렉션 횟수를 나타낸다. 실험 결과는 트랜잭션의 수가 증가함에 따라 메타데이터에 대한 업데이트가 많이 발생할수록 MiNV 파일시스템의 성능이 상대적으로 향상됨을 보여준다.

본 실험 결과는 동일한 가비지 컬렉션 기법을 사용하더라도 MiNV 파일시스템이 YAFFS보다 더 효율적으로 가비지 컬렉션을 수행함을 보여준다. 그림 2(b)에서 YAFFS와 MiNV 파일시스템은 블록 소거 연산의 대부분이 가비지 컬렉션에 의해서 발생함을 알 수 있다. 이는 수행중인 워크로드의 특성상 블록의 페이지 전체가 모두 무효화되는 경우가 거의 없다는 것을 의미한다. 실제로 벤치마크는 2개의 파일에 대해서는 단 한번도 데이터에 대한 삭제를 수행하지 않는다. MiNV 파일시스템은 트랜잭션 2,000회까지 블록 소거 연산을 거의 수행하지 않다가 트랜잭션 2,500 회부터 공격적으로 가비지 컬렉션을 수행한다. 반면에 YAFFS는 트랜잭션 500부터 1,500까지는 소극적인 가비지 컬렉션을 수행하다가 트랜잭션 2,000회 이상부터는 공격적으로 가비지 컬렉션을 수행한다. 그림 2(a)에서 알 수 있듯이 공격적인 가비지 컬렉션을 시작하는 순간부터 500개의 트랜잭션을 처리하는데 점점 더 오랜 시간을 소모함을 알 수 있다.



(a) 500회 트랜잭션마다 소요된 수행시간



(b) 낸드플래시에 요청된 연산 횟수

그림 2 MiNVFS와 YAFFS의 TFFS 벤치마크 수행 결과

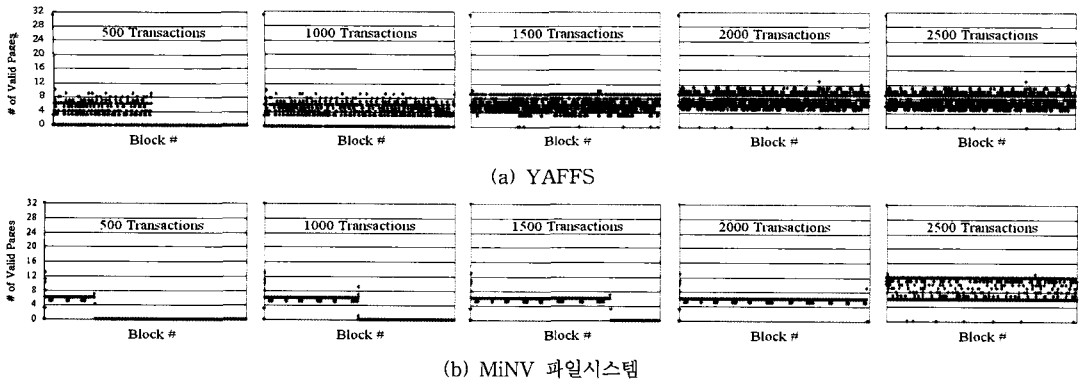


그림 3 각 파일시스템이 트랜잭션을 수행한 후 낸드플래시의 각 블록당 유효한 페이지 수

따라서 MiNV 파일시스템의 수행 성능 향상에 대한 하나의 원인으로써, MiNV 파일시스템에서 가비지 컬렉션 수행에 대한 효율성을 들 수 있다.

그림 3은 MiNV 파일시스템에서 가비지 컬렉션 수행에 대한 효율성을 YAFFS와 비교하여 직관적으로 보여준다. 그림 3은 YAFFS와 MiNV 파일시스템이 500부터 2,500회의 트랜잭션을 수행하고 난 후에, 낸드플래시 각각의 블록에 포함되어 있는 그 내용이 보존되어야 하는 페이지의 개수를 보여준다. 이러한 보존되어야 하는 데이터를 포함한 페이지들은 가비지 컬렉션 과정에서 하나의 블록에서 다른 블록으로 복사되어야 하기 때문에, 이들 페이지의 개수가 가비지 컬렉션이 수행되도록 선택된 블록 내에 적게 존재할수록 가비지 컬렉션 비용은 줄어들게 된다. MiNV 파일시스템은 낸드플래시에 파일데이터만 유지하므로 그림 3(b)처럼 트랜잭션 2,000회를 수행할 때까지 다수의 빈 블록들을 확보하고 있다. 따라서 2,000회의 트랜잭션까지는 가비지 컬렉션이 발생하지 않는다.

YAFFS는 그림 3(a)에서와 같이 트랜잭션 1,000회까지 블록들을 모두 한번씩 다 사용하였지만 여전히 중간 중간 다수의 빈 블록들을 확보한다. 이는 3개 이하의 유효한 페이지를 포함하는 블록들이 이미 소극적인 가비지 컬렉션에 의해서 가용한 상태로 전환되었기 때문이다. 하지만 트랜잭션 1,000회 이후에는 공격적인 가비지 컬렉션이 발생하게 된다. 이처럼 YAFFS는 같은 수의 트랜잭션을 수행하더라도 MiNV 파일시스템보다 더 많은 블록들을 소모하기 때문에 더 빨리 공격적인 가비지 컬렉션을 수행할 수 밖에 없다.

5. 결론 및 향후 과제

본 논문에서는 메타데이터를 비휘발성 램에 관리하는 MiNV 파일시스템에서 가비지 컬렉션의 효율성을 살펴

보았으며, 이러한 가비지 컬렉션의 효율성이 전체 파일 시스템 성능에 미치는 영향에 대해서 실험 결과를 살펴 보면서 분석하였다.

성능 평가 결과, MiNV 파일시스템은 YAFFS와 비교해서 전체 낸드플래시에서 가용한 블록의 개수가 부족할 때 수행되는 공격적인 가비지 컬렉션을 더 늦은 시점에서 발생시킬 뿐만 아니라 공격적인 가비지 컬렉션을 더 적게 수행함을 확인하였다. 이는 같은 가비지 컬렉션 기법을 적용하더라도 MiNV 파일시스템이 YAFFS보다 더 효율적으로 가비지 컬렉션을 수행함을 의미한다. 이러한 효율성이 전체 수행 성능에 대한 성능 향상을 가져옴을 실험 결과를 통해서 확인하였다.

참 고 문 헌

- [1] Freescale Semiconductor, <http://www.freescale.com>.
- [2] Ramtron International - Nonvolatile Memory, Integrated Memory and Microcontrollers, <http://www.ramtron.com>
- [3] Tech-On News, http://techon.nikkeibp.co.jp/english/NEWS_EN/20070226/128173
- [4] ETNEWS, Commercialization of Pram is at hand: what will happen to nor flash?, <http://english.etnews.co.kr/news/detail.html?id=200706070003>
- [5] In Hwan Doh, Jongmoo Choi, Donghee Lee, and Sam H. Noh, "Exploiting Non-Volatile RAM to Enhance Flash File System Performance," In *Proceedings of the 7th ACM & IEEE International Conference on Embedded Software (EMSOFT'07)*, pp. 164-173, Oct. 2007.
- [6] Aleph One Company, YAFFS (Yet Another Flash File System), <http://www.yaffs.net>
- [7] E. Gal and S. Toledo, A Transactional Flash File System for Microcontrollers, In *Proceedings of the USENIX Annual Technical Conference (USENIX 2005)*, pp. 89-104, Apr. 2005.
- [8] FALINUX, <http://falinux.com/zproducts/ez-m28.php>