(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2018년03월02일
(11) 등록번호 10-1834082
(24) 등록일자 2018년02월23일

(51) 국제특허분류(Int. Cl.)
G06F 3/06 (2006.01)
(52) CPC특허분류
G06F 3/0688 (2013.01)
G06F 3/064 (2013.01)
(21) 출원번호 10-2016-0111585
(22) 출원일자 2016년08월31일
심사청구일자 2016년08월31일
(56) 선행기술조사문헌
KR1020130026517 A*
KR1020160075229 A*
*는 심사관에 의하여 인용된 문헌

(73) 특허권자
울산과학기술원
울산광역시 울주군 언양읍 유니스트길 50
(72) 발명자
노삼혁
울산광역시 울주군 언양읍 유니스트길 50
김병석
울산광역시 울주군 언양읍 유니스트길 50
김재호
울산광역시 울주군 언양읍 유니스트길 50
(74) 대리인
특허법인 무한

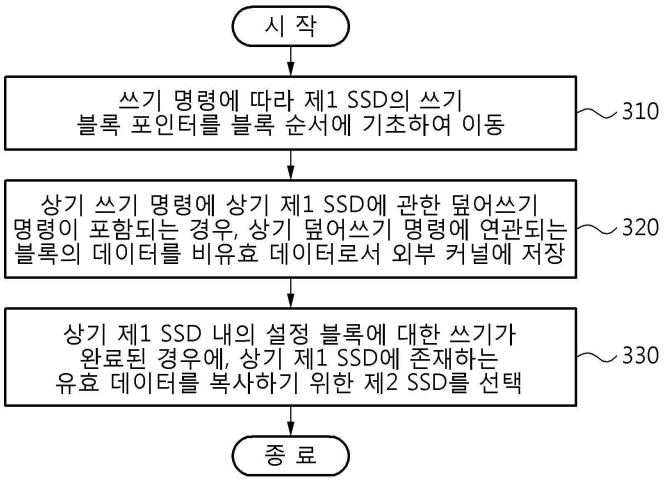
전체 청구항 수 : 총 7 항 심사관 : 이철수

(54) 발명의 명칭 **다중 SSD 시스템 관리 장치 및 방법**

(57) 요 약

SSD 시스템의 관리 장치는 쓰기 명령에 따라 제1 SSD의 쓰기 블록 포인터(writing block pointer)를 블록 순서에 기초하여 이동시키는 매핑부 및 상기 제1 SSD 내의 설정 블록에 대한 쓰기가 완료된 경우에, 상기 제1 SSD에 존재하는 유효 데이터를 복사하기 위한 제2 SSD를 선택하는 관리부를 포함할 수 있다.

대 표 도 - 도3



(52) CPC특허분류
G06F 3/0652 (2013.01)
G06F 3/0656 (2013.01)

이 발명을 지원한 국가연구개발사업
과제고유번호 2015R1A2A2A05027651
부처명 교육과학기술부
연구관리전문기관 한국연구재단
연구사업명 이공분야기초연구사업
연구과제명 플래시 기반 임베디드 시스템에서 PCM의 실용적 적용을 위한 운영체제 개발
기 여 율 1/1
주관기관 울산과학기술원
연구기간 2015.09.01 ~ 2018.08.31
공지예외적용 : 있음

명 세 서

청구범위

청구항 1

컴퓨터로 구현되는:

쓰기 명령에 따라 제1 SSD의 쓰기 블록 포인터(writing block pointer)를 블록 순서에 기초하여 이동시키는 매핑부;

상기 제1 SSD 내의 설정 블록에 대한 쓰기가 완료된 경우에, 상기 제1 SSD에 존재하는 유효 데이터를 복사하기 위한 제2 SSD를 선택하는 관리부; 및

상기 쓰기 명령에 상기 제1 SSD에 관한 덮어쓰기 명령이 포함되는 경우, 상기 덮어쓰기 명령에 연관되는 블록의 데이터를 비유효 데이터(invalid data)로 판단하는 메타 데이터를 저장하고, 상기 관리부에 의해 상기 제2 SSD가 선택된 경우에 상기 제1 SSD에 존재하는 유효 데이터를 상기 제2 SSD로 복사하고, 상기 제1 SSD 내의 모든 유효 데이터가 복사된 경우에, 상기 제1 SSD로 전체 데이터에 대한 TRIM 명령을 전송하여 상기 제1 SSD 내의 모든 블록에 대응하는 모든 데이터를 제거하는 커널

을 포함하고,

상기 커널은 상기 제1 SSD 및 상기 제2 SSD의 외부 커널 영역에 구현되고, 상기 TRIM 명령은 파일 시스템 내에서 비유효 처리된 데이터를 각각의 SSD에 알려주는 명령인 SSD(Solid State Disk) 시스템의 관리 장치.

청구항 2

제1항에 있어서,

상기 매핑부는 상기 제1 SSD에 포함되는 복수 개의 블록의 순서에 기초하여 순차적으로 상기 쓰기 블록 포인터를 이동시키는 SSD 시스템의 관리 장치.

청구항 3

제1항에 있어서,

상기 제1 SSD의 마지막 블록에 대한 쓰기가 완료된 경우에, 상기 관리부는 쓰기 명령을 수행하기 위한 제3 SSD를 선택하는 SSD 시스템의 관리 장치.

청구항 4

제3항에 있어서,

상기 매핑부는 쓰기 명령에 따라 상기 선택된 제3 SSD의 쓰기 블록 포인터(writing block pointer)를 블록 순서에 기초하여 이동시키는 SSD 시스템의 관리 장치.

청구항 5

삭제

청구항 6

삭제

청구항 7

삭제

청구항 8

컴퓨터가 수행하는 SSD 시스템 관리 방법에 있어서,

쓰기 명령에 따라 제1 SSD의 쓰기 블록 포인터를 블록 순서에 기초하여 이동시키는 단계;

상기 쓰기 명령에 상기 제1 SSD에 관한 덮어쓰기 명령이 포함되는 경우, 상기 덮어쓰기 명령에 연관되는 블록의 데이터를 비유효 데이터로 판단하는 메타 데이터를 외부 커널에 저장하는 단계;

상기 제1 SSD 내의 설정 블록에 대한 쓰기가 완료된 경우에, 상기 제1 SSD에 존재하는 유효 데이터를 복사하기 위한 제2 SSD를 선택하는 단계;

상기 제2 SSD가 선택된 경우에 상기 제1 SSD에 존재하는 유효 데이터를 상기 제2 SSD로 복사하는 단계; 및

상기 제1 SSD 내의 모든 유효 데이터가 복사된 경우에, 상기 제1 SSD로 전체 데이터에 대한 TRIM 명령을 전송하여 상기 제1 SSD 내의 모든 블록에 대응하는 모든 데이터를 상기 외부 커널에 의해 제거하는 단계

를 포함하고,

상기 외부 커널은 상기 제1 SSD 및 상기 제2 SSD의 외부 커널 영역에 구현되고, 상기 TRIM 명령은 파일 시스템 내에서 비유효 처리된 데이터를 각각의 SSD에 알려주는 명령인 SSD 시스템의 관리 방법.

청구항 9

삭제

청구항 10

삭제

청구항 11

제8항에 있어서,

상기 제2 SSD를 선택하는 단계는,

상기 제1 SSD의 마지막 블록에 대한 쓰기가 완료된 경우에, 상기 쓰기 명령을 수행하기 위한 제3 SSD를 선택하는 단계

를 포함하는 SSD 시스템의 관리 방법.

청구항 12

다중 SSD 시스템에 포함되는 SSD에 대해 가비지 컬렉션을 수행하는 프로그램을 수록한 컴퓨터 판독 가능한 기록 매체에 있어서, 상기 프로그램은:

쓰기 명령에 따라 제1 SSD의 쓰기 블록 포인터를 블록 순서에 기초하여 이동시키는 명령어 세트;

상기 쓰기 명령에 상기 제1 SSD에 관한 덮어쓰기 명령이 포함되는 경우, 상기 덮어쓰기 명령에 연관되는 블록의 데이터를 비유효 데이터로 판단하는 메타 데이터를 외부 커널에 저장하는 명령어 세트;

상기 제1 SSD 내의 설정 블록에 대한 쓰기가 완료된 경우에, 상기 제1 SSD에 존재하는 유효 데이터를 복사하기 위한 제2 SSD를 선택하는 명령어 세트;

상기 제2 SSD가 선택된 경우에 상기 제1 SSD에 존재하는 유효 데이터를 상기 제2 SSD로 복사하는 명령어 세트; 및

상기 제1 SSD 내의 모든 유효 데이터가 복사된 경우에, 상기 제1 SSD로 전체 데이터에 대한 TRIM 명령을 전송하여 상기 제1 SSD 내의 모든 블록에 대응하는 모든 데이터를 상기 외부 커널에 의해 제거하는 명령어 세트

를 포함하고,

상기 외부 커널은 상기 제1 SSD 및 상기 제2 SSD의 외부 커널 영역에 구현되고, 상기 TRIM 명령은 파일 시스템 내에서 비유효 처리된 데이터를 각각의 SSD에 알려주는 명령인 컴퓨터 판독 가능한 기록 매체.

발명의 설명

기 술 분 야

[0001] SSD(Solid State Disk) 시스템 관리 방법에 연관되며, 보다 구체적으로는 복수 개의 플래시 기반 SSD를 포함하는 다중 SSD 시스템을 위한 관리 방법에 연관된다.

배 경 기 술

[0002] 플래시 메모리 기반의 SSD는 반도체 메모리 소자를 이용하여 구현되기 때문에 하드디스크 보다 매우 빠른 속도로 데이터에 접근하는 것이 가능하다. 제조사의 SSD 관련 기술이 발전함에 따라 SSD는 내구성과 안정성 또한 확보하게 되어 오늘날 클라우딩 컴퓨팅 시스템이나 파일 시스템 등과 같은 다양한 영역에 이용되고 있다.

[0003] 다만, SSD의 기반이 되는 플래시 메모리는 소재의 특성 상 덮어쓰기가 불가능하여 입력된 데이터를 지운 뒤에 쓰기를 수행하여야 한다. 이는, SSD의 성능이 일정하게 유지되지 못하는 원인 중 하나로 작용되고 있다.

[0004] 종래에는 SSD가 관리하는 내부 자원을 각각의 가상 머신(VM: Virtual Machine) 내에 맞추어 고립시키는 방식이 이용되었다. 다만, 종래의 방식의 경우, SSD 내부의 블록에는 각기 다른 가상 머신의 페이지들이 포함되고, 특정 가상 머신에 의해 가비지 컬렉션(GC: Garbage Collection)이 발생하는 경우에도, 다른 가상 머신의 페이지들을 정리하는데 시간이 소요될 수 있다.

선행기술문헌

비특허문헌

[0005] (비특허문헌 0001) Kim, Jaeho. Lee, Donghee. Noh, Sam H. "Towards SLO Complying SSDs Through OPS Isolation" Proceedings of the 13th USENIX Conference on File and Storage Technologies, 2015, pp. 183-189

발명의 내용

과제의 해결 수단

[0006] 다중 SSD 시스템에서 안정적인 성능을 제공하도록 하는 관리 방법 및 이를 위한 관리 장치의 다양한 측면들 및 실시예들이 제시된다. 예시적으로, 그러나 한정되지 않은 몇 개의 측면들은 아래에서 서술된다.

[0007] 일측에 따르면, 적어도 하나의 프로세서를 포함하는 컴퓨터에 의해 구현되는 다중 SSD 시스템의 관리 장치가 제공된다. 상기 다중 SSD 시스템의 관리 장치는 쓰기 명령에 따라 제1 SSD의 쓰기 블록 포인터(writing block pointer)를 블록 순서에 기초하여 이동시키는 매핑부 및 상기 제1 SSD 내의 설정 블록에 대한 쓰기가 완료된 경우에, 상기 제1 SSD에 존재하는 유효 데이터를 복사하기 위한 제2 SSD를 선택하는 관리부를 포함할 수 있다.

[0008] 일실시예에 따르면, 상기 매핑부는 상기 제1 SSD에 포함되는 복수 개의 블록의 순서에 기초하여 순차적으로 상기 쓰기 블록 포인터를 이동시킬 수 있다.

[0009] 다른 일실시예에 따르면, 상기 제1 SSD의 마지막 블록에 대한 쓰기가 완료된 경우에, 상기 관리부는 쓰기 명령을 수행하기 위한 제3 SSD를 선택할 수 있다. 보다 구체적으로, 상기 매핑부는 쓰기 명령에 따라 상기 선택된 제3 SSD의 쓰기 블록 포인터를 블록 순서에 기초하여 이동시킬 수 있다.

[0010] 또 다른 일실시예에 따르면, 상기 다중 SSD 시스템의 관리 장치는 상기 쓰기 명령에 상기 제1 SSD에 관한 덮어쓰기 명령이 포함되는 경우, 상기 덮어쓰기 명령에 연관되는 블록의 데이터를 비유효 데이터(invalid data)로서

판단하는 메타 데이터를 저장하는 커널을 더 포함할 수 있다. 보다 구체적으로, 상기 커널은 식별된 상기 제1 SSD 내의 유효 데이터를 상기 제2 SSD로 복사할 수 있다. 또한, 상기 커널은 상기 제1 SSD 내의 모든 유효 데이터가 복사된 경우에, 상기 제1 SSD로 TRIM 명령을 전송하여 모든 데이터를 제거할 수 있다.

[0011] 다른 일측에 따르면, 덮어쓰기 명령에 연관되는 블록의 데이터를 비유효 데이터로서 판단하는 메타 데이터를 커널에 저장하여 가비지 컬렉션을 수행하기 위한 환경을 제공하는 다중 SSD 시스템의 관리 방법이 제공된다. 상기 다중 SSD 시스템의 관리 방법은 쓰기 명령에 따라 제1 SSD의 쓰기 블록 포인터를 블록 순서에 기초하여 이동시키는 단계, 상기 제1 SSD 내의 설정 블록에 대한 쓰기가 완료된 경우에, 상기 제1 SSD에 존재하는 유효 데이터를 복사하기 위한 제2 SSD를 선택하는 단계 및 상기 쓰기 명령에 상기 제1 SSD에 관한 덮어쓰기 명령이 포함되는 경우, 상기 덮어쓰기 명령에 연관되는 블록의 데이터를 비유효 데이터로서 판단하는 메타 데이터를 외부 커널에 저장하는 단계를 포함할 수 있다.

[0012] 일실시예에 따르면, 상기 다중 SSD 시스템의 관리 방법은 상기 외부 커널에 저장된 상기 제1 SSD 내의 유효 데이터 정보를 이용하여 상기 제1 SSD 내의 유효 데이터를 상기 제2 SSD로 복사하는 단계를 더 포함할 수 있다.

[0013] 다른 일실시예에 따르면, 상기 다중 SSD 시스템의 관리 방법은 상기 제1 SSD 내의 모든 유효 데이터가 복사된 경우, 상기 제1 SSD로 TRIM 명령을 전송하는 단계를 더 포함할 수 있다.

[0014] 또 다른 일실시예에 따르면, 상기 제2 SSD를 선택하는 단계는 상기 제1 SSD의 마지막 블록에 대한 쓰기가 완료된 경우에, 상기 쓰기 명령을 수행하기 위한 제3 SSD를 선택하는 단계를 포함할 수 있다.

[0015] 또 다른 일측에 따르면, 다중 SSD 시스템에 포함되는 SSD에 대해 가비지 컬렉션을 수행하는 프로그램을 수록한 컴퓨터 판독 가능한 기록 매체가 제공된다. 상기 프로그램은 제1 SSD 내의 설정 블록에 대한 쓰기가 완료된 경우에, 상기 제1 SSD에 존재하는 유효 데이터를 복사하기 위한 제2 SSD를 선택하는 명령어 세트 및 상기 쓰기 명령에 상기 제1 SSD에 관한 덮어쓰기 명령이 포함되는 경우, 상기 덮어쓰기 명령에 연관되는 블록의 데이터를 비유효 데이터로서 판단하는 메타 데이터를 외부 커널에 저장하는 명령어 세트를 포함할 수 있다.

도면의 간단한 설명

[0016] 도 1은 일실시예에 따른 SSD 어레이 관리자의 구조를 나타내는 예시도이다.

도 2a는 일실시예에 따른 복수 개의 SSD를 이용하여 외부 가비지 컬렉션이 수행되는 초기 과정을 설명하는 예시도이다.

도 2b는 일실시예에 따른 제1 SSD에 대해 외부 가비지 컬렉션이 완료되는 과정을 설명하는 예시도이다.

도 3을 일실시예에 따른 다중 SSD 시스템을 관리하는 방법의 흐름도이다.

도 4는 일실시예에 다른 외부 가비지 컬렉션이 수행되는 경우에 호스트 버퍼가 이용되는 과정을 설명하는 예시도이다.

도 5는 일실시예에 따른 다중 SSD 시스템의 관리 장치를 설명하는 블록도이다.

발명을 실시하기 위한 구체적인 내용

[0017] 실시예들에 대한 특정한 구조적 또는 기능적 설명들은 단지 예시를 위한 목적으로 개시된 것으로서, 다양한 형태로 변경되어 실시될 수 있다. 따라서, 실시예들은 특정한 개시형태로 한정되는 것이 아니며, 본 명세서의 범위는 기술적 사상에 포함되는 변경, 균등물, 또는 대체물을 포함한다.

[0018] 제1 또는 제2 등의 용어를 다양한 구성요소들을 설명하는데 사용될 수 있지만, 이런 용어들은 하나의 구성요소를 다른 구성요소로부터 구별하는 목적으로만 해석되어야 한다. 예를 들어, 제1 구성요소는 제2 구성요소로 명명될 수 있고, 유사하게 제2 구성요소는 제1 구성요소로도 명명될 수 있다.

[0019] 어떤 구성요소가 다른 구성요소에 "연결되어" 있다고 언급된 때에는, 그 다른 구성요소에 직접적으로 연결되어 있거나 또는 접속되어 있을 수도 있지만, 중간에 다른 구성요소가 존재할 수도 있다고 이해되어야 할 것이다.

[0020] 단수의 표현은 문맥상 명백하게 다르게 뜻하지 않는 한, 복수의 표현을 포함한다. 본 명세서에서, "포함하다" 또는 "가지다" 등의 용어는 설시된 특징, 숫자, 단계, 동작, 구성요소, 부분품 또는 이들을 조합한 것이 존재함으로 지정하려는 것이지, 하나 또는 그 이상의 다른 특징들이나 숫자, 단계, 동작, 구성요소, 부분품 또는 이들을 조합한 것들의 존재 또는 부가 가능성을 미리 배제하지 않는 것으로 이해되어야 한다.

[0021] 다르게 정의되지 않는 한, 기술적이거나 과학적인 용어를 포함해서 여기서 사용되는 모든 용어들은 해당 기술 분야에서 통상의 지식을 가진 자에 의해 일반적으로 이해되는 것과 동일한 의미를 가진다. 일반적으로 사용되는 사전에 정의되어 있는 것과 같은 용어들은 관련 기술의 문맥상 가지는 의미와 일치하는 의미를 갖는 것으로 해석되어야 하며, 본 명세서에서 명백하게 정의하지 않는 한, 이상적이거나 과도하게 형식적인 의미로 해석되지 않는다.

[0022] 이하, 실시예들을 첨부된 도면들을 참조하여 상세하게 설명한다. 첨부 도면을 참조하여 설명함에 있어, 도면 부호에 관계없이 동일한 구성 요소는 동일한 참조 부호를 부여하고, 이에 대한 중복되는 설명은 생략하기로 한다.

[0023] 아래의 설명에서 가비지 컬렉션(GC: Garbage Collection)은 특정 SSD에 저장된 데이터 중 다른 주소에 Out of Place Update 방식으로 덮어쓰기된 기존의 데이터를 비유효 데이터로서 삭제하는 과정을 나타낼 수 있다. 본 실시예에서는, 복수 개의 SSD에 연관되는 커널을 통해 특정 SSD의 유효 데이터 및 비유효 데이터에 관한 매핑 데이터가 저장될 수 있다. 또한, 특정 SSD의 유효 데이터의 복사가 완료된 경우에 상기 커널은 상기 특정 SSD로 TRIM 명령을 전송하여 외부 가비지 컬렉션이 수행될 수 있도록 지원한다. 본 실시예에 따른 외부 가비지 컬렉션에 관한 자세한 설명은 아래에서 추가될 도면과 함께 보다 자세하게 설명될 것이다.

[0024] 도 1은 일실시예에 따른 SSD 어레이 관리자의 구조를 나타내는 예시도이다. 도 1을 참조하면, 커널 영역(kernel space)에 포함되는 파일 시스템(file system) 및 블록 레이어(block layer)가 도시된다. 또한, 사용자 영역(user space)에는 사용자에 의해 이용되는 복수 개의 가상 머신(VM: Virtual Machine)과 가상 머신에 의해 실행되는 어플리케이션(application)이 도시된다.

[0025] 보다 구체적으로, 블록 레이어는 디바이스 맵퍼(device mapper)와 입력/출력 스케쥴러를 포함할 수 있다. 예시적으로, 디바이스 맵퍼는 커널에서 제공하는 디스크 볼륨 관리를 위한 개발 플랫폼을 나타낼 수 있다. 또한, 디바이스 맵퍼는 여러 개의 디바이스들을 가상 저장장치 또는 가상 블록 디바이스로서 생성할 수 있다. 디바이스 맵퍼는 상기 생성된 가상 저장장치 또는 상기 가상 블록 디바이스와 실제 디바이스 사이의 매핑 방식을 정의할 수 있다. 본 실시예에 따른 다중 SSD 시스템을 관리하는 관리 장치는 디바이스 맵퍼를 통해 구현될 수 있다.

[0026] 일실시예로서, 디바이스 맵퍼는 LVM(Linux Volume Manager)로서 이용될 수 있다. 보다 구체적으로, LVM은 물리 장치를 나타내는 PV(Physical Volume), 상기 PV들의 집합으로 구현되는 VG(Volume Group), 상기 VG 내에서 기설정된 용량에 따라 관리 방법이 적용되는 LV(Logical Volume)을 포함할 수 있다. 본 실시예에서 제안되는 다중 SSD 시스템을 관리하는 관리 장치는 디바이스 맵퍼 계층에서 LVM의 관리 모듈로써 구현될 수 있다.

[0027] 도 2a는 일실시예에 따른 복수 개의 SSD를 이용하여 외부 가비지 컬렉션이 수행되는 초기 과정을 설명하는 예시도이다. 도 2a를 참조하면, 다중 SSD 시스템 내에 포함되는 복수 개의 SSD(210, 220, 230)가 도시된다. 본 실시예에서는 다중 SSD 시스템 내에 4 개의 SSD가 포함되는 구성이 도시되지만, 이는 발명의 이해를 돕기 위한 예시적 기재일 뿐 다른 실시예의 범위를 제한하거나 한정하는 것을 해석되어서는 안될 것이다. 이를테면, 본 실시예의 사상은 40 개의 SSD를 포함하는 다중 SSD 시스템 또는 400 개의 SSD를 포함하는 다중 SSD 시스템과 같은 다양한 변형 실시예에도 적용될 수 있을 것이다.

[0028] 본 실시예에서 설명되는 다중 SSD 시스템은 각각의 SSD가 물리적으로 분리된 인터페이스로 연결되는 시스템을 나타낼 수 있다. 그에 따라, 다중 SSD 시스템은 하나의 SSD에서 수행되는 작업이 다른 SSD의 대역폭에 영향을 미치지 않는 효과를 기대할 수 있다.

[0029] 예시적으로, 제1 SSD(210)에 대해 쓰기(writing)가 수행되는 경우가 존재할 수 있다. 이 경우에, 다중 SSD 시스템의 관리 장치는 기설정된 규칙에 따라 제1 SSD(210)에 포함되는 블록에 대해 쓰기를 수행할 수 있다. 본 실시예에서 설명되는 다중 SSD 시스템의 관리 장치는 커널 계층 중 파일 시스템과 디바이스 드라이버 사이에서 모듈 형태로 구현될 수 있다.

[0030] 다중 SSD 시스템의 관리 장치는 Log-Structured 방식으로 제1 SSD(210)에 대해 쓰기를 수행될 수 있다. 그에 따라, 제1 SSD(210)는 요청된 쓰기 블록과는 독립적으로 제1 SSD(210) 내에 포함되는 첫 번째 블록에서부터 마지막 블록까지의 순차적인 순서로써 쓰기가 수행될 수 있다. 다중 SSD 시스템의 관리 장치는 제1 SSD(210)의 마지막 블록에 대한 쓰기가 완료된 경우에, 제2 SSD(220)로 쓰기 블록 포인터(writing block pointer)(250)를 이동시킬 수 있다.

[0031] 예시적으로, 그러나 한정되지 않게, 제1 SSD(210)의 제1 블록에 기입력된 제1 데이터에 대해 덮어쓰기 요청이 전달되는 경우가 존재할 수 있다. 이 경우에, 다중 SSD 시스템의 관리 장치는 앞서 설명된 Out of Place Update 방식에 따라 제1 SSD(210) 내의 제2 블록에 상기 제1 데이터에 대응하는 새로운 제2 데이터의 쓰기를 수행할 수 있다. 이 경우에, 덮어쓰기가 수행된 과거의 제1 데이터는 비유효(invalid) 데이터가 될 수 있다. 또한, 새롭게 덮어쓰기가 수행된 제2 데이터나 처음 입력된 이후에 다른 덮어쓰기가 수행되지 않은 제3 데이터는 유효 데이터(valid data)(211)로서 관리 될 수 있다.

[0032] 또한, 제1 SSD(210) 내의 설정 블록에 대한 쓰기가 완료된 경우에, 다중 SSD 시스템의 관리 장치는 제1 SSD(210)에 존재하는 유효 데이터(211)를 복사하기 위한 제3 SSD(230)를 선택할 수 있다. 선택된 제3 SSD(230)를 통해 유효 데이터가 복사되고, 제1 SSD(210)에 가비지 컬렉션이 수행되는 구체적 과정에 관한 설명은 이하 도 2b와 함께 설명될 것이다.

[0033] 도 2b는 일실시예에 따른 제1 SSD에 대해 외부 가비지 컬렉션이 완료되는 과정을 설명하는 예시도이다. 도 2b를 참조하면, 제3 SSD(230)로 복사된 제1 SSD(210)의 유효 데이터(212)가 도시된다. 본 실시예의 다중 SSD 시스템의 관리 장치는 제1 SSD(210)에 포함되는 블록에 저장된 유효 데이터(211)들을 제3 SSD(230)에 복사할 수 있다. 보다 구체적으로, 다중 SSD 시스템은 덮어쓰기에 대응하는 제1 SSD(210)의 비유효 데이터를 표시하고, 제1 SSD(210)에 포함되는 블록에 관한 유효 데이터와 비유효 데이터를 식별할 수 있다. 그에 따라, 제1 SSD(210) 자체는 내부 블록 전체에 포함되는 데이터들을 유효 데이터로서 인식하고, 별도의 가비지 컬렉션을 수행하지 않을 수 있다.

[0034] 일실시예로서, 제1 SSD(210)의 유효 데이터(211)가 제3 SSD(230)의 유효 데이터(212)로서 복사되는 과정은, 쓰기 블록 포인터(250)가 제1 SSD(210)에서 제2 SSD(220)로 이동된 경우에 수행될 수 있다. 또한, 제3 SSD(230)는 쓰기 블록 포인터(250)가 새롭게 옮겨진 제2 SSD(220)를 제외하고, 다중 SSD 시스템에 포함되는 나머지 SSD들 중 임의적으로 선택될 수 있다.

[0035] 다중 SSD 시스템의 관리 장치는 제1 SSD(210)의 유효 데이터(211)의 전체가 제3 SSD(230)로 복사된 경우에, 제1 SSD(210)에 남겨진 원본 데이터를 비유효 데이터로서 판단하는 메타 데이터를 저장할 수 있다. 또한, 다중 SSD 시스템의 관리 장치는 제3 SSD(230)에 새롭게 저장된 유효 데이터(212)를 파일 시스템 내의 메타 데이터와 매핑시킬 수 있다.

[0036] 또한, 다중 SSD 시스템의 관리 장치는 제1 SSD(210)로 TRIM 명령(260)을 전달할 수 있다. 제1 SSD(210)는 전달된 TRIM 명령(260)에 기초하여 모든 블록 내의 데이터들을 제거하고 안정적인 상태로 되돌아 갈 수 있다.

[0037] 본 실시예에 따르는 외부 가비지 컬렉션 과정은 쓰기 블록 포인터(250)가 새롭게 이동된 제2 SSD(220)의 쓰기 과정에 그대로 적용될 수 있다. 제2 SSD(220)에 포함되는 모든 블록에 대해 첫 번째 블록에서부터 마지막 블록까지 순차적으로 쓰기가 진행되고, 쓰기 블록 포인터(250) 역시도 순차적으로 이동될 수 있다. 제2 SSD(220)의 마지막 블록에 쓰기가 수행된 경우, 다중 SSD 시스템의 관리 장치는 제2 SSD(220)에 존재하는 유효 데이터를 복사하기 위한 새로운 SSD를 다중 SSD 시스템 내에서 임의적으로 선택할 수 있다.

[0038] 다중 SSD 시스템의 관리 장치는 커널 영역에 존재하는 어레이 관리자를 이용하여 특정 SSD에 연관되는 비유효 데이터 정보를 관리할 수 있다. 다만, 다중 SSD 시스템의 관리 장치는 상기 특정 SSD로는 의도적으로 비유효 데이터에 관한 매핑 정보를 전달하지 않을 수 있다. 그에 따라, 특정 SSD는 가비지 컬렉션을 수행하지 않고 그대로 쓰기 작업을 지속할 수 있고, 가비지 컬렉션에 의해 대역폭 변동과 성능 하락을 방지하는 효과를 기대할 수 있다.

[0039] 도 3을 일실시예에 따른 다중 SSD 시스템을 관리하는 방법의 흐름도이다. 도 3을 참조하면, 다중 SSD 시스템을 관리하는 방법이 도시된다. 다중 SSD 시스템을 관리하는 방법은 쓰기 명령에 따라 제1 SSD의 쓰기 블록 포인터를 블록 순서에 기초하여 이동하는 단계(310), 상기 쓰기 명령에 상기 제1 SSD에 관한 덮어쓰기 명령이 포함되는 경우, 상기 덮어쓰기 명령에 연관되는 블록의 데이터를 비유효 데이터로서 외부 커널에 저장하는 단계(320) 및 상기 제1 SSD 내의 설정 블록에 대한 쓰기가 완료된 경우에, 상기 제1 SSD에 존재하는 유효 데이터를 복사하기 위한 제2 SSD를 선택하는 단계(330)를 포함할 수 있다. 다중 SSD 시스템을 관리하는 방법은 적어도 하나의 프로세서를 포함하고, 상기 적어도 하나의 프로세서에 의해 적어도 일시적으로 구현되는 다중 SSD 시스템의 관리 장치에 의해 수행될 수 있다.

[0040] 단계(310)에서 다중 SSD 시스템의 관리 장치는 제1 SSD의 쓰기 블록 포인터를 블록 순서에 기초하여 이동시킬 수 있다. 제1 SSD는 다중 SSD 시스템에 포함되는 복수 개의 SSD 중에서 현재 쓰기 명령이 수행되고 있는 SSD를

나타낼 수 있다. 다중 SSD 시스템의 관리 장치는 블록 단위의 Out of Place Update 방식에 따라 쓰기 및 덮어쓰기를 제1 SSD 내에 포함되는 새로운 블록에 지속적으로 할당 할 수 있다.

[0041] 일실시예로서, 다중 SSD 시스템의 관리 장치는 제1 SSD 내에 존재하는 첫 번째 블록에서부터 마지막 블록까지 순차적으로 쓰기 작업을 수행할 수 있다. 본 실시예에서, 첫 번째 블록은 제1 SSD 내에 존재하는 0번 주소에 대응하는 블록을 나타낼 수 있으며, 제1 SSD 내의 첫 번째 블록에 관한 물리적 위치에는 다양한 변경이 가능할 것이다. 다중 SSD 시스템의 관리 장치는 수행되는 쓰기 명령에 따라 쓰기 블록 포인터를 이동시킬 수 있다.

[0042] 단계(320)에서 다중 SSD 시스템의 관리 장치는 상기 쓰기 명령에 상기 제1 SSD에 관한 덮어쓰기 명령이 포함되는 경우, 상기 덮어쓰기 명령에 연관되는 블록의 데이터를 비유효 데이터로서 판단하는 메타 데이터를 외부 커널에 저장할 수 있다. 본 실시예에서 다중 SSD 시스템의 관리 장치는 데이터에 관한 변경이 발생할 경우에 상기 외부 커널에만 메타 데이터 형태로서 비유효 데이터를 저장하고, 상기 제1 SSD로는 통지하지 않는다. 따라서, 상기 덮어쓰기 명령에 의한 데이터의 업데이트가 지속적으로 발생하는 경우에도, 제1 SSD 내부에서는 모든 데이터들을 유효 데이터로서 판단하여 계속적으로 쓰기 작업이 수행될 수 있다. 본 실시예에 따를 때, 다중 SSD 시스템의 관리 장치가 제1 SSD의 내부에 의도적으로 비유효 페이지를 만들지 않는 이상, 제1 SSD는 가비지 컬렉션을 수행하지 않아 상기 가비지 컬렉션에 의한 성능의 변동이 방지되는 효과를 기대할 수 있다.

[0043] 단계(330)에서 다중 SSD 시스템의 관리 장치는 상기 제1 SSD 내의 설정 블록에 대한 쓰기가 완료된 경우에, 상기 제1 SSD를 가비지 컬렉션 SSD로서 인식할 수 있다. 예시적으로, 설정 블록은 제1 SSD 내에 존재하는 마지막 주소에 대응하는 블록을 나타낼 수 있다. 아래의 설명에서 가비지 컬렉션 SSD는 전체 블록에 대한 쓰기가 완료되어, 외부 가비지 컬렉션이 수행되고 데이터가 정리될 필요가 있는 SSD를 나타낼 수 있다. 다중 SSD 시스템의 관리 장치는 SSD 내부에서 발생하게 될 가비지 컬렉션을 예방하기 위해 상기 관리 장치 측면에서 다른 SSD를 이용하여 가비지 컬렉션을 수행하는 과정을 나타낼 수 있다.

[0044] 또한, 단계(330)에서 다중 SSD 시스템의 관리 장치는 상기 제1 SSD에 존재하는 유효 데이터를 복사하기 위한 제2 SSD를 선택할 수 있다. 다중 SSD 시스템의 관리 장치는 제2 SSD를 타겟(target) SSD로서 인식할 수 있다. 아래의 설명에서 타겟 SSD는 제1 SSD가 모두 채워진 경우에, 상기 제1 SSD에 대한 외부 가비지 컬렉션이 수행될 수 있도록 상기 제1 SSD의 유효 데이터가 복사되는 SSD를 나타낼 수 있다. 제2 SSD에 유효 데이터를 복사하고, 상기 제1 SSD에 대해 TRIM 명령이 수행되는 과정에 대해서는 이하의 도 5에서 추가적으로 자세히 설명될 것이다.

[0045] 도 4는 일실시예에 다른 외부 가비지 컬렉션이 수행되는 경우에 호스트 버퍼가 이용되는 과정을 설명하는 예시도이다. 도 4를 참조하면, 커널 소프트웨어 내에서 블록 레이어에 포함되는 버퍼와 볼륨 매니저의 구성이 도시된다. 본 실시예에 따른 다중 SSD 시스템의 관리 장치는 디스크 볼륨을 관리할 수 있는 디바이스 맵퍼를 통해 구현될 수 있다. 다중 SSD 시스템의 관리 장치는 유틸리티 형태로 구현되어, 새로운 블록 디바이스들을 생성하고 그에 대한 관리 방식을 인자의 이름을 통해 호출할 수 있다. 다중 SSD 시스템의 관리 장치는 상기 관리 방식을 커널 모듈 형태로 포함할 수 있고, 호출되었을 경우에 적재시킬 수 있다.

[0046] 상기 매핑 알고리즘에 기초하여, 다중 SSD 시스템의 관리 장치는 특정 SSD 내에 쓰기 블록 포인터(430)를 위치시킬 수 있다. 또한, 다중 SSD 시스템의 관리 장치는 쓰기 블록 포인터(430)의 위치를 이동시키면서 쓰기 명령을 수행할 수 있다. 본 실시예에 따른, 다중 SSD 시스템의 관리 장치는 특정 SSD 내의 마지막 섹터까지 순서대로 쓰기가 수행된 경우, 특정 SSD의 상태를 외부 가비지 컬렉션이 필요한 상태로 기록하고, 다음 SSD에 관한 쓰기를 계속적으로 수행할 수 있다.

[0047] 다중 SSD 시스템의 관리 장치는 외부 가비지 컬렉션이 수행되는 경우에, 과정(410)을 통해 유효 데이터가 커널 내에 존재하는 버퍼로 읽어지도록 할 수 있다. 또한, 다중 SSD 시스템의 관리 장치는 과정(420)을 통해 커널 내의 버퍼에서 제2 SSD로 쓰기가 수행되도록 하여 외부 가비지 컬렉션이 수행되도록 할 수 있다.

[0048] 도 5는 일실시예에 따른 다중 SSD 시스템의 관리 장치를 설명하는 블록도이다. 도 5를 참조하면, 다중 SSD 시스템의 관리 장치(500)는 매핑부(510), 관리부(520) 및 커널(530)을 포함할 수 있다. 다중 SSD 시스템의 관리 장치(500)는 적어도 하나의 프로세서를 포함하며, 상기 적어도 하나의 프로세서에 의해 적어도 일시적으로 구현될 수 있다.

[0049] 매핑부(510)는 쓰기 명령에 따라 제1 SSD에 존재하는 쓰기 블록 포인터를 블록 순서에 기초하여 이동시킬 수 있다. 보다 구체적으로, 매핑부(510)는 상기 블록 순서에 따라 제1 SSD 내의 첫 번째 블록에서부터 마지막 블록까지 순차적으로 상기 쓰기 블록 포인터를 이동시킬 수 있다.

[0050] 관리부(520)는 상기 제1 SSD 내의 설정 블록에 대한 쓰기가 완료된 경우, 상기 제1 SSD에 존재하는 유효 데이터를 복사하기 위한 제2 SSD를 선택할 수 있다. 또한, 관리부(520)는 상기 제1 SSD의 마지막 블록에 대한 쓰기가 완료된 경우에, 쓰기 명령을 계속적으로 수행하기 위한 제3 SSD를 선택할 수 있다. 상기 제2 SSD는 새로운 쓰기 작업이 수행될 제3 SSD와는 다른 SSD로서, 상기 제1 SSD의 유효 데이터가 옮겨질 타겟 SSD로서 지칭될 수 있다. 매핑부(510)는 쓰기 명령에 따라 새롭게 선택된 제3 SSD의 쓰기 블록 포인터를 블록 순서에 기초하여 이동시키면서, 쓰기 명령이 지속적으로 수행되게 지원할 수 있다.

[0051] 커널(530)은 쓰기 명령에 상기 제1 SSD에 관한 덮어쓰기 명령이 포함되는 경우, 상기 덮어쓰기 명령에 연관되는 블록의 데이터를 비유효 데이터로서 표시하고, 해당 데이터를 매핑 데이터로서 저장할 수 있다. 커널(530)은 기저장된 매핑 데이터를 이용하여 상기 제2 SSD가 선택된 경우에, 상기 제1 SSD 내의 유효 데이터를 상기 제2 SSD로 복사할 수 있다. 커널(530)은 상기 제1 SSD 내의 모든 유효 데이터가 복사된 경우에, 상기 제1 SSD로 TRIM 명령을 전송하여 제1 SSD 내의 모든 블록에 대응하는 모든 데이터를 제거할 수 있다. 본 명세서에 기재된 TRIM 명령은 파일 시스템 내에서 비유효 처리된 데이터를 각각의 SSD에 알려주는 명령을 나타낼 수 있다. 본 실시예에서는 SSD 내부의 가비지 컬렉션을 방지하여 전체 시스템의 성능을 유지하는 취지에서, SSD 외부의 커널(530)에 의해 유효 데이터가 사전에 복사되고 전체 데이터에 대한 TRIM 명령이 전달될 수 있다.

[0052] 이상에서 설명된 실시예들은 하드웨어 구성요소, 소프트웨어 구성요소, 및/또는 하드웨어 구성요소 및 소프트웨어 구성요소의 조합으로 구현될 수 있다. 예를 들어, 실시예들에서 설명된 장치, 방법 및 구성요소는, 예를 들어, 프로세서, 콘트롤러, ALU(arithmetic logic unit), 디지털 신호 프로세서(digital signal processor), 마이크로컴퓨터, FPGA(field programmable gate array), PLU(programmable logic unit), 마이크로프로세서, 또는 명령(instruction)을 실행하고 응답할 수 있는 다른 어떠한 장치와 같이, 하나 이상의 범용 컴퓨터 또는 특수 목적 컴퓨터를 이용하여 구현될 수 있다. 처리 장치는 운영 체제(OS) 및 상기 운영 체제 상에서 수행되는 하나 이상의 소프트웨어 애플리케이션을 수행할 수 있다. 또한, 처리 장치는 소프트웨어의 실행에 응답하여, 데이터를 접근, 저장, 조작, 처리 및 생성할 수도 있다. 이해의 편의를 위하여, 처리 장치는 하나가 사용되는 것으로 설명된 경우도 있지만, 해당 기술분야에서 통상의 지식을 가진 자는, 처리 장치가 복수 개의 처리 요소(processing element) 및/또는 복수 유형의 처리 요소를 포함할 수 있음을 알 수 있다. 예를 들어, 처리 장치는 복수 개의 프로세서 또는 하나의 프로세서 및 하나의 콘트롤러를 포함할 수 있다. 또한, 병렬 프로세서(parallel processor)와 같은, 다른 처리 구성(processing configuration)도 가능하다.

[0053] 소프트웨어는 컴퓨터 프로그램(computer program), 코드(code), 명령(instruction), 또는 이들 중 하나 이상의 조합을 포함할 수 있으며, 원하는 대로 동작하도록 처리 장치를 구성하거나 독립적으로 또는 결합적으로(collectively) 처리 장치를 명령할 수 있다. 소프트웨어 및/또는 데이터는, 처리 장치에 의하여 해석되거나 처리 장치에 명령 또는 데이터를 제공하기 위하여, 어떤 유형의 기계, 구성요소(component), 물리적 장치, 가상 장치(virtual equipment), 컴퓨터 저장 매체 또는 장치, 또는 전송되는 신호 파(signal wave)에 영구적으로, 또는 일시적으로 구체화(embody)될 수 있다. 소프트웨어는 네트워크로 연결된 컴퓨터 시스템 상에 분산되어서, 분산된 방법으로 저장되거나 실행될 수도 있다. 소프트웨어 및 데이터는 하나 이상의 컴퓨터 판독 가능 기록 매체에 저장될 수 있다.

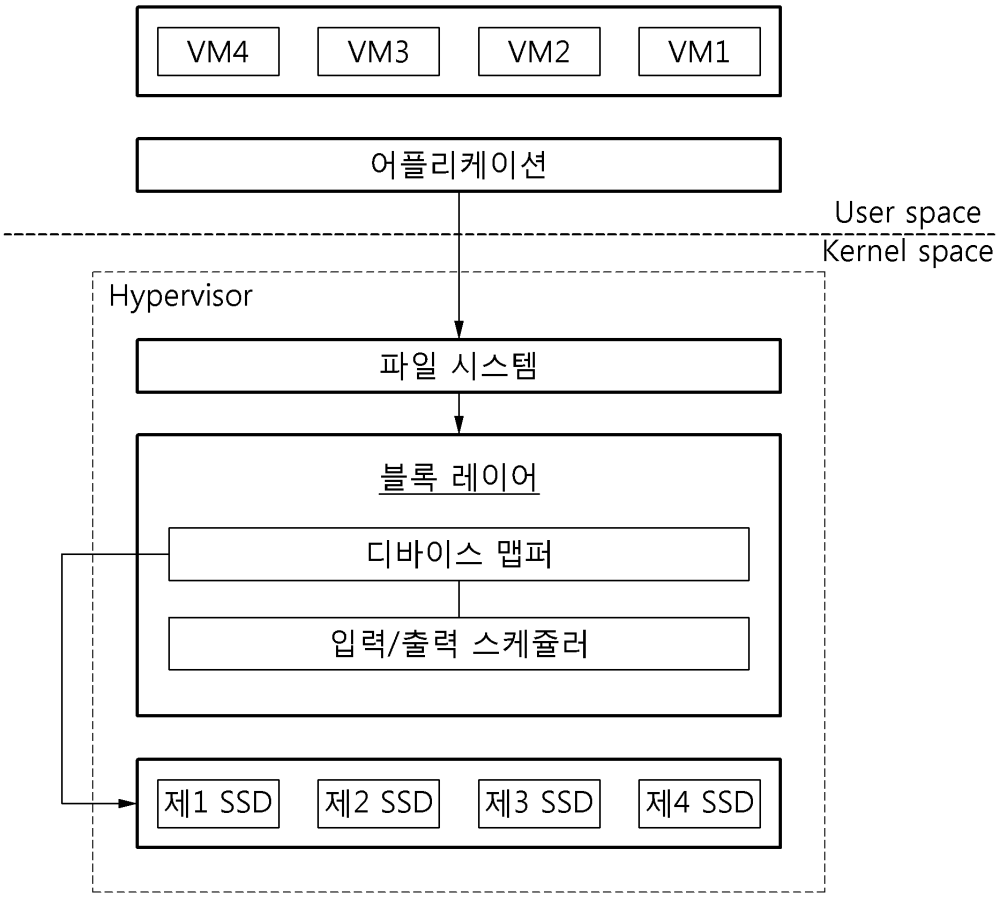
[0054] 실시예에 따른 방법은 다양한 컴퓨터 수단을 통하여 수행될 수 있는 프로그램 명령 형태로 구현되어 컴퓨터 판독 가능 매체에 기록될 수 있다. 컴퓨터 판독 가능 매체는 프로그램 명령, 데이터 파일, 데이터 구조 등을 단독으로 또는 조합하여 포함할 수 있다. 컴퓨터 판독 가능 매체에 기록되는 프로그램 명령은 실시예를 위하여 특별히 설계되고 구성된 것들이거나 컴퓨터 소프트웨어 당업자에게 공지되어 사용 가능한 것일 수도 있다. 컴퓨터 판독 가능 기록 매체의 예에는 하드 디스크, 플로피 디스크 및 자기 테이프와 같은 자기 매체(magnetic media), CD-ROM, DVD와 같은 광기록 매체(optical media), 플롭티컬 디스크(floptical disk)와 같은 자기-광 매체(magneto-optical media), 및 롬(ROM), 램(RAM), 플래시 메모리 등과 같은 프로그램 명령을 저장하고 수행하도록 특별히 구성된 하드웨어 장치가 포함된다. 프로그램 명령의 예에는 컴파일러에 의해 만들어지는 것과 같은 기계어 코드뿐만 아니라 인터프리터 등을 사용해서 컴퓨터에 의해서 실행될 수 있는 고급 언어 코드를 포함한다. 상기된 하드웨어 장치는 실시예의 동작을 수행하기 위해 하나 이상의 소프트웨어 모듈로서 작동하도록 구성될 수 있으며, 그 역도 마찬가지이다.

[0055] 이상과 같이 실시예들이 비록 한정된 도면에 의해 설명되었으나, 해당 기술분야에서 통상의 지식을 가진 자라면 상기를 기초로 다양한 기술적 수정 및 변형을 적용할 수 있다. 예를 들어, 설명된 기술들이 설명된 방법과 다른 순서로 수행되거나, 및/또는 설명된 시스템, 구조, 장치, 회로 등의 구성요소들이 설명된 방법과 다른 형태로 결합 또는 조합되거나, 다른 구성요소 또는 균등물에 의하여 대치되거나 치환되더라도 적절한 결과가 달성될

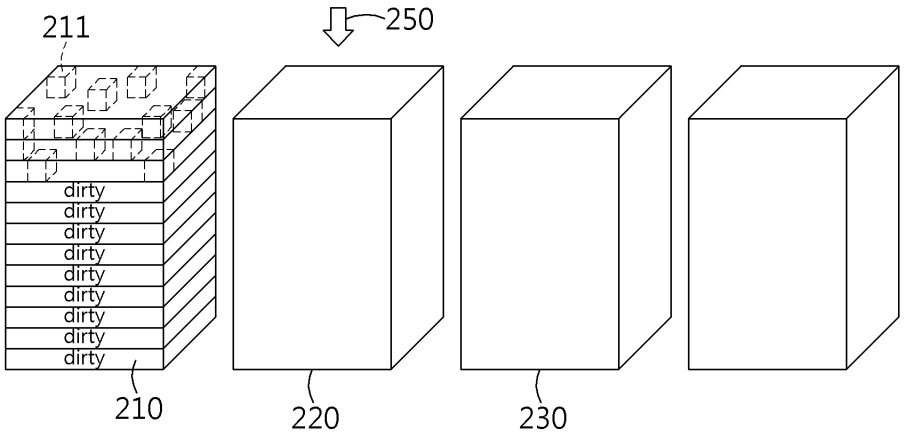
수 있다.

도면

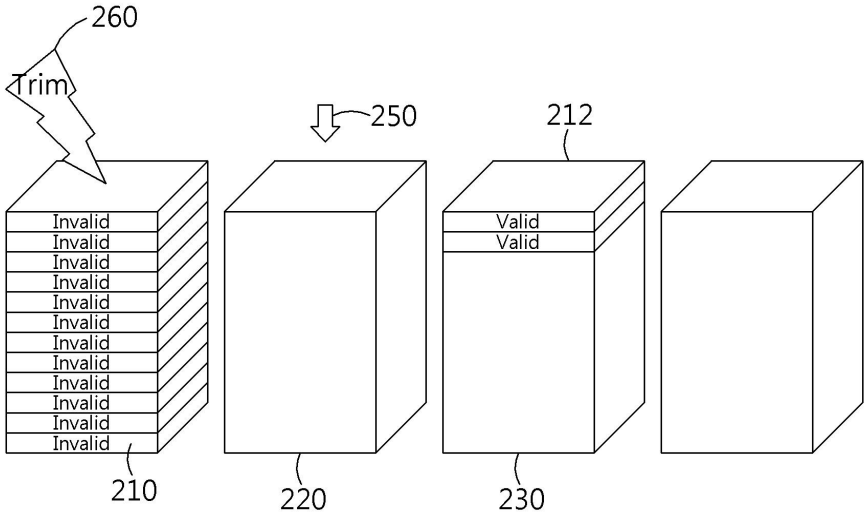
도면1



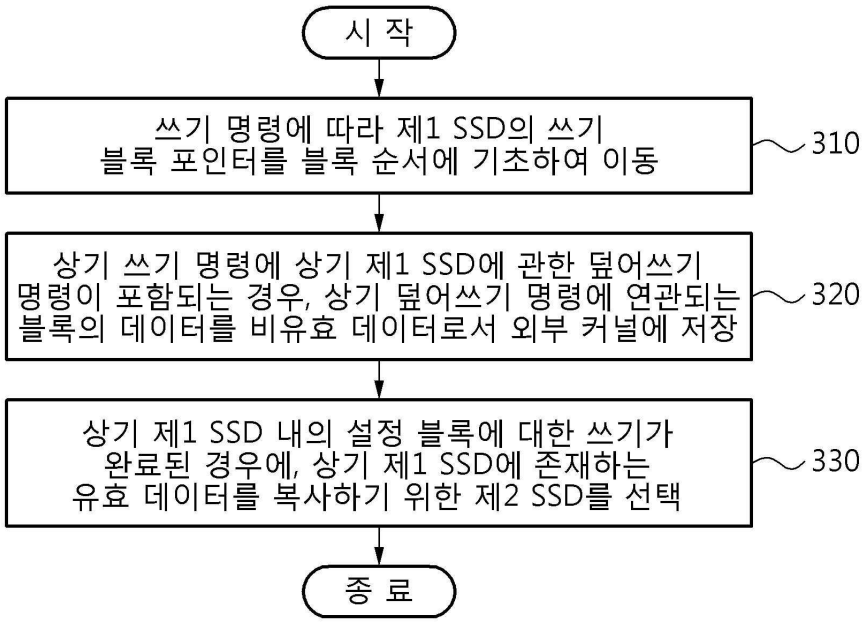
도면2a



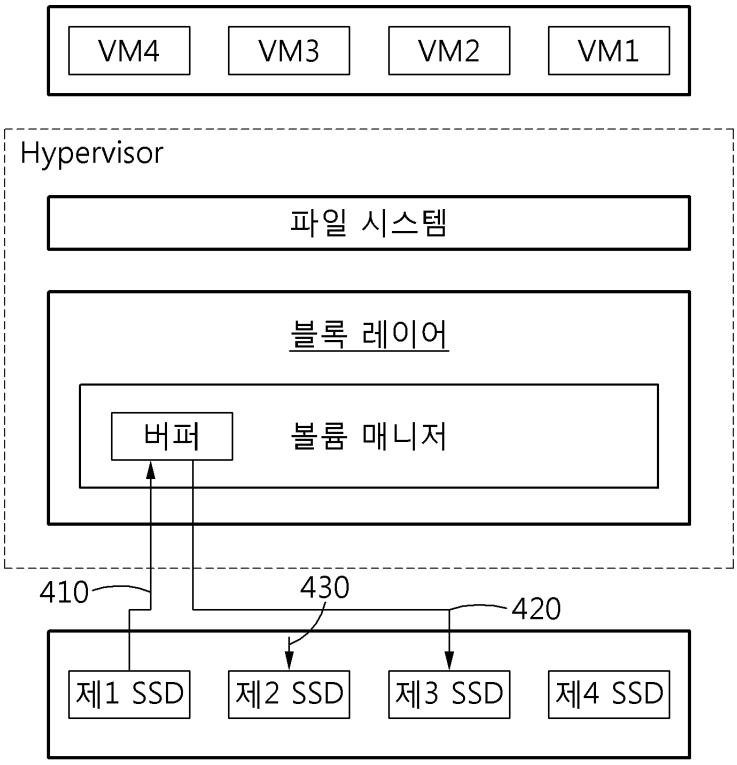
도면2b


260
Trim
250
212
Invalid
Invalid
Invalid
Invalid
Invalid
Invalid
Invalid
Invalid
Invalid
Invalid
Invalid
Valid
Valid
210
220
230

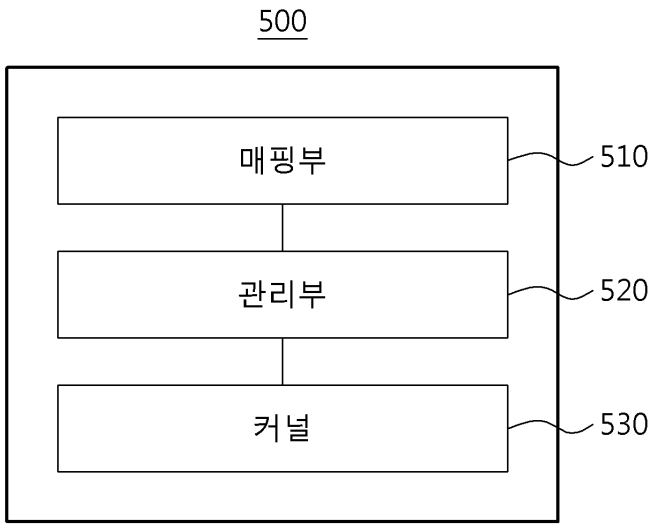
도면3


시 작
쓰기 명령에 따라 제1 SSD의 쓰기
블록 포인터를 블록 순서에 기초하여 이동
310
상기 쓰기 명령에 상기 제1 SSD에 관한 덮어쓰기
명령이 포함되는 경우, 상기 덮어쓰기 명령에 연관되는
블록의 데이터를 비유효 데이터로서 외부 커널에 저장
320
상기 제1 SSD 내의 설정 블록에 대한 쓰기가
완료된 경우에, 상기 제1 SSD에 존재하는
유효 데이터를 복사하기 위한 제2 SSD를 선택
330
종 료

도면4


VM4
VM3
VM2
VM1
Hypervisor
파일 시스템
블록 레이어
버퍼
볼륨 매니저
410
430
420
제1 SSD
제2 SSD
제3 SSD
제4 SSD

도면5


500
매핑부
510
관리부
520
커널
530