



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2018-0025128
(43) 공개일자 2018년03월08일

- | | |
|--|---|
| <p>(51) 국제특허분류(Int. Cl.)
G06F 3/06 (2006.01) G06F 12/02 (2018.01)</p> <p>(52) CPC특허분류
G06F 3/0659 (2013.01)
G06F 12/0246 (2013.01)</p> <p>(21) 출원번호 10-2017-0027650</p> <p>(22) 출원일자 2017년03월03일
심사청구일자 없음</p> <p>(30) 우선권주장
201641029370 2016년08월29일 인도(IN)</p> | <p>(71) 출원인
삼성전자주식회사
경기도 수원시 영통구 삼성로 129 (매탄동)</p> <p>(72) 발명자
시바난드 스리칸스 톰쿠르
인도, 카르나타카, 톰쿠르-572103, 비드야
나가르, 4번가
코다세 샤라스 쿠마르
인도, 카르나타카, 벵갈루루-560062, 코나나쿤테,
뉴 벵크 콜로니, 2번가, 45
베르마 라지브
인도, 카르나타카, 벵갈루루-560037, 아우터 링
로드, 도나파쿤디, 포스트 편스 시티,
마라타할리, 바그만 콘스텔레이션 비즈니스 파크
서클, 피닉스 빌딩</p> <p>(74) 대리인
특허법인가산</p> |
|--|---|

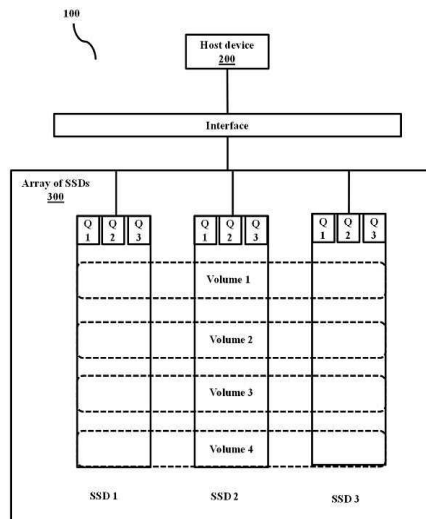
전체 청구항 수 : 총 20 항

(54) 발명의 명칭 SSD 어레이 관리를 위한 스트림 식별자 기반 스토리지 시스템

(57) 요약

SSD 어레이 관리를 위한 스트림 식별자 기반 스토리지 시스템이 제공된다. 스토리지 시스템은, 데이터를 저장하는 솔리드 스테이트 드라이브(SSD)의 어레이, 및 복수의 논리적 볼륨으로써 상기 SSD의 어레이를 관리하는 호스트 컨트롤러로, 각각의 상기 논리적 볼륨은 적어도 하나의 스트림 식별자(Identifier; ID)와 관련되고, 적어도 하나의 SSD 및 상기 호스트 컨트롤러는 스트림 ID 명령을 실행하여 상기 논리적 볼륨을 액세스하는 시스템 코드를 포함하는 호스트 컨트롤러를 포함한다.

대표도 - 도1



(52) CPC특허분류

G06F 12/0253 (2013.01)

G06F 3/0679 (2013.01)

G06F 3/0689 (2013.01)

명세서

청구범위

청구항 1

데이터를 저장하는 솔리드 스테이트 드라이브(SSD)의 어레이; 및

복수의 논리적 볼륨으로써 상기 SSD의 어레이를 관리하는 호스트 컨트롤러로, 각각의 상기 논리적 볼륨은 적어도 하나의 스트림 식별자(Identifier; ID)와 관련되고, 적어도 하나의 SSD 및 상기 호스트 컨트롤러는 스트림 ID 명령을 실행하여 상기 논리적 볼륨을 액세스하는 시스템 코드를 포함하는 호스트 컨트롤러를 포함하는 스토리지 시스템.

청구항 2

제 1항에 있어서,

상기 각각의 논리적 볼륨은 상기 SSD 각각의 메모리 영역에 전용(dedicate)되고, 상기 SSD 각각의 메모리 영역에 대응하는 상기 각각의 논리적 볼륨은 상기 적어도 하나의 스트림 ID를 이용하는 논리적 볼륨 그룹을 형성하도록 조합되는 스토리지 시스템.

청구항 3

제 2항에 있어서,

상기 호스트 컨트롤러는 상기 스트림 ID 명령을 이용하여 상기 적어도 하나의 스트림 ID를 이용하여 고유하게 (uniquely) 식별되는 논리적 볼륨 그룹과 통신하는 스토리지 시스템.

청구항 4

제 1항에 있어서,

상기 스트림 ID 명령은 스트림 ID 활성화 명령, 스트림 ID 유희 명령, 스트림 ID 상태 획득(get) 명령 중 하나인 스토리지 시스템.

청구항 5

제 4항에 있어서,

상기 스트림 ID 유희 명령은 적어도 하나의 파워 모드(power mode) 동작을 정의하는 동작 유형(operation type), 가비지 컬렉터(garbage collector) 동작, 데이터 연산 동작 및 사용자 지정(custom) 동작과 관련되는 스토리지 시스템.

청구항 6

제 1항에 있어서,

상기 적어도 하나의 SSD 및 상기 호스트 컨트롤러는 상기 시스템 코드를 실행(implement)하고 운영 체제(operating system)에 의하여 로드된 RAID 드라이버와 공존하는 스토리지 시스템.

청구항 7

제 1항에 있어서,

상기 호스트 컨트롤러는,

상기 복수의 논리적 볼륨으로부터 요청을 수신하여 논리적 볼륨의 LBA(Logical Block Addressing)에 액세스하고,

상기 논리적 볼륨에 대응하는 스트림 ID를 결정하고,

비휘발성 랜덤 액세스 메모리(Non-Volatile Random Access Memory; NVRAM)로부터 상기 논리적 볼륨에 대응하는 스트림 ID의 상태를 검출하되, 상기 상태는 활성(active) 또는 유휴(passive) 중의 하나이고,

상기 스트림 ID의 상태가 활성으로 검출된 경우,

상기 논리적 볼륨의 상기 스트림 ID에 대응하는 상기 복수의 SSD 중 적어도 하나의 SSD의 상기 LBA에 대한 액세스를 허용하고,

상기 스트림 ID의 상태가 유휴로 검출된 경우,

상기 논리적 볼륨의 상기 스트림 ID에 대응하는 상기 복수의 SSD 중 적어도 하나의 SSD에 대하여 가비지 컬렉션 기능을 수행하는, 스토리지 시스템.

청구항 8

제 7항에 있어서,

조율(orchestration) 로직에 기초한 패리티 그룹(parity group)으로부터 상기 LBA에 대한 상기 적어도 하나의 SSD가 선택되고,

상기 패리티 그룹은 상기 LBA에 기초하여 결정되는, 스토리지 시스템.

청구항 9

제 7항에 있어서,

상기 복수의 SSD로부터의 상기 적어도 하나의 SSD에 대한 상기 가비지 컬렉션 기능은 조율 시퀀스에 기초하여 수행되는 스토리지 시스템.

청구항 10

제 7항에 있어서,

상기 요청은 읽기 요청, 쓰기 요청 및 LBA 연산 요청 중 하나인, 스토리지 시스템.

청구항 11

제 7항에 있어서,

상기 NVRAM은 상기 LBA에 액세스하기 위한 상기 요청을, 상기 스트림 ID에 대응하는 상기 논리적 볼륨으로 전송하기 전에 버퍼링하는, 스토리지 시스템.

청구항 12

제 11항에 있어서,

상기 호스트 컨트롤러는 상기 스트림 ID 명령에 기초하여 상기 NVRAM 내의 상기 각각의 스트림 ID의 상태를 설정하는, 스토리지 시스템.

청구항 13

프로세서; 및

상기 프로세서 및 SSD 어레이에 연결되는 컨트롤러를 포함하되,

상기 호스트 컨트롤러는,

상기 SSD 어레이를 복수의 논리적 볼륨으로 관리하되,

상기 각각의 논리적 볼륨은 적어도 하나의 스트림 ID와 관련되고, 상기 호스트 컨트롤러는 상기 각각의 논리적 볼륨에 액세스하기 위한 스트림 ID 명령을 실행하는 시스템 코드를 포함하는, 호스트 장치.

청구항 14

제 13항에 있어서,

상기 호스트 컨트롤러는,

상기 복수의 논리적 볼륨으로부터 요청을 수신하여 논리적 볼륨의 LBA에 액세스하고,

상기 논리적 볼륨에 대응하는 스트림 ID를 결정하고,

NVRAM으로부터 상기 논리적 볼륨에 대응하는 상기 스트림 ID의 상태를 검출하되, 상기 상태는 활성 및 유틸 상태이고,

상기 스트림 ID의 상태가 활성으로 검출된 경우, 상기 논리적 볼륨의 상기 스트림 ID에 대응하는 상기 복수의 SSD 중 적어도 하나의 SSD의 상기 LBA에 대한 액세스를 허용하고,

상기 스트림 ID의 상태가 유틸로 검출된 경우, 상기 논리적 볼륨의 상기 스트림 ID에 대응하는 상기 복수의 SSD 중 적어도 하나의 SSD에 대하여 가비지 컬렉션을 수행하는, 호스트 장치.

청구항 15

제 14항에 있어서,

조율 로직에 기초한 패리티 그룹으로부터 상기 LBA에 대한 상기 적어도 하나의 SSD가 선택되고, 상기 패리티 그룹은 상기 LBA에 기초하여 결정되는, 호스트 장치.

청구항 16

제 14항에 있어서,

상기 복수의 SSD 중 상기 적어도 하나의 SSD에 대하여 수행되는 상기 가비지 컬렉션은 조율 시퀀스에 기초하여 수행되는, 호스트 장치.

청구항 17

제 14항에 있어서,

상기 요청은 읽기 요청, 쓰기 요청 및 LBA 연산 요청 중 하나인, 호스트 장치.

청구항 18

제 14항에 있어서,

상기 NVRAM은 상기 LBA에 액세스하기 위한 상기 요청을, 상기 스트림 ID에 대응하는 상기 논리적 볼륨으로 전송하기 전에 버퍼링하는, 호스트 장치.

청구항 19

호스트 컨트롤러에 의해, SSD 어레이로부터 논리적 볼륨의 LBA에 액세스하는 요청을 수신하고,

상기 호스트 컨트롤러에 의해, 상기 논리적 볼륨에 대응하는 스트림 ID를 결정하고,

상기 호스트 컨트롤러에 의해, 상기 논리적 볼륨에 대응하는 상기 스트림 ID의 상태를 검출하되, 상기 상태는 활성 및 유틸 중 하나이고,

상기 호스트 컨트롤러에 의해, 상기 스트림 ID의 상태가 활성으로 검출된 경우 상기 논리적 볼륨의 상기 스트림 ID에 대응하는 상기 복수의 SSD 중 적어도 하나의 SSD의 상기 LBA에 대한 액세스를 허용하고,

상기 스트림 ID의 상태가 유틸로 검출된 경우, 상기 논리적 볼륨의 상기 스트림 ID에 대응하는 상기 복수의 SSD 중 적어도 하나의 SSD에 대하여 가비지 컬렉션 기능을 수행하는 것을 포함하는, SSD 어레이의 관리 방법.

청구항 20

제 19항에 있어서,

상기 호스트 컨트롤러는 NVRAM으로부터 상기 논리적 볼륨에 대응하는 상기 스트림 ID의 상기 상태를 검출하고, 상기 NVRAM은 LBA에 액세스하기 위한 상기 요청을 상기 스트림 ID에 대응하는 상기 논리적 볼륨으로 전송하기 전에 버퍼링하는, SSD 어레이의 관리 방법.

발명의 설명

기술 분야

[0001] 본 발명은 스토리지 시스템에 관한 것이며, 더욱 상세하게는 솔리드 스테이트 디스크(Solid State Disk; SSD)를 위한 스트림 식별자(identifier; ID)에 기초한 스토리지 시스템에 관련된 것이다.

배경 기술

[0002] SSD는 반도체 소자이고, 다른 자기 디스크와는 달리 기계적으로 움직이는 부품이 없다. 이는 디스크 또는 SSD 헤드 탐색 레이턴시를 줄이고 다양한 입출력(IO) 작업에 대한 성능을 향상시킬 수 있다. 또한 SSD는 자기 디스크에 비하여 기계적 장애에 탄력적(resilient)이다. 현재 SSD는 세 개의 기본적 IO 동작, 즉 쓰기, 읽기 및 삭제를 지원한다. SSD의 기본 데이터 저장 단위는 페이지(일반적으로 4KB 용량의 플래시 메모리 셀 그룹)이다. 페이지는 블록으로 더욱 그룹화된다. SSD에서 읽기 및 쓰기의 단위(granularity)는 페이지 레벨이지만, 삭제 작업의 단위는 블록 레벨이다.

[0003] SSD에서의 삭제 동작은 가비지 컬렉션(Garbage Collection; GC)으로 불리는 메커니즘을 사용함으로써 해결된다. 가비지 컬렉션은 무효화된 페이지를 회수하고 SSD에 사용 가능한 여유 공간을 만드는 프로세스를 정의하는 용어이다. 가비지 컬렉션 프로세스가 SSD의 특정 블록에서 진행될 때마다, 해당 특정 블록에 해당하는 모든 유입된(incoming) 요청이 진행 중인 가비지 컬렉션 프로세스가 완료될 때까지 연기된다. 그러므로, 유입된 요청이 지연되어 큐에 배치되고 가비지 컬렉션 프로세스가 완료된 후에 서비스가 예정된다. 이러한 유입된 요청은, 유입된 요청이 폭발적(bursty)일 때 스토리지 시스템(즉, SSD)의 성능을 저하시킬 수 있다. SSD의 동일한 볼륨에서의 가비지 컬렉션과 IO 동작의 연속적인(즉, 동시의) 동작은 IO 성능에 영향을 미친다.

[0004] 통상적으로, 데이터 분석(빅 데이터와 같은) 및 온라인 트랜잭션 처리(On-Line Transaction Processing; OLTP)는 동일한 스토리지 서버로부터 서비스되지 않는다. 둘은 서로 다른 두 종류의 애플리케이션을 제공하기 때문이다. OLTP 애플리케이션은 IOPS(초당 IO 작업 수) 속도가 더 빠른 서버를 사용하며 데이터 분석은 빠른 서버를 사용하지 않는다. 앞으로 더 빠른 데이터 분석에 대한 수요가 증가함에 따라 OLTP와 마찬가지로 고속 플래시 기반 스토리지 서버에서 빅 데이터/데이터 분석 애플리케이션을 호스팅하는 솔루션이 필요하다. 이는 가비지 컬렉션으로 인한 것이다.

[0005] 가비지 컬렉션은 SSD에서 피할 수 없는 프로세스이다. 현재, 플래시 어레이(예를 들어 SSD 어레이) 명령 또는 NVMe(Non-volatile memory express)/SAS(Serial Attached small computer system) 드라이버는 SSD 내부의 가비지 컬렉션에 대한 어떠한 제어도 하지 못한다. SSD 내의 가비지 컬렉션은 가비지 컬렉션 및 IO 작업이 동일한 블록에서 작동 중이거나 대기 중이므로 플래시 어레이에 대한 IO 작업 또는 데이터 연산(스마트 SSD의 경우)에 영향을 준다.

[0006] 고성능 블록 스토리지 및 데이터 분석을 모두 지원하는 현재의 메커니즘은 스토리지 시스템 및 데이터 분석이 주어진 시간 내에 실행될 수 있도록 각 SSD 내부 코어의 증가 및 SSD 내의 각 코어의 처리 용량 증가를 요구한다. 또한 SSD 내의 RAM(Random Access Memory) 요구량이 증가할 수 있다. 상기 모든 기술은 각각의 SSD 비용을 증가시키고 결과적으로 플래시 어레이 시스템 또는 SSD 어레이의 비용을 증가시키기 때문에 고비용을 요구한다. 때때로 IO 성능 요구 사항을 충족시키기 위하여 압축을 위한 특별한 하드웨어 지원이 필요하다.

[0007] 상기 정보는 읽는 사람으로 하여금 본 발명을 이해할 수 있도록 배경 지식으로 제공된다. 출원인은 위의 사항 중 어떠한 것도 본 출원과 관련하여 선행 기술로 적용될 수 있는지 여부에 대하여 아무런 결정도 내리지 않았으며 어떠한 주장도 하지 않았다.

발명의 내용

해결하려는 과제

[0008] 본 발명이 해결하고자 하는 기술적 과제는 SSD 어레이를 관리하기 위한 스트림 ID 기반 스토리지 시스템을 제공하는 것이다.

[0009] 본 발명의 기술적 과제들은 이상에서 언급한 기술적 과제로 제한되지 않으며, 언급되지 않은 또 다른 기술적 과제들은 아래의 기재로부터 당업자에게 명확하게 이해될 수 있을 것이다.

과제의 해결 수단

- [0010] 상기 기술적 과제를 달성하기 위한 본 발명의 일 실시예에 따른 스토리지 시스템은, 데이터를 저장하는 솔리드 스테이트 드라이브(SSD)의 어레이, 및 복수의 논리적 볼륨으로써 상기 SSD의 어레이를 관리하는 호스트 컨트롤러로, 각각의 상기 논리적 볼륨은 적어도 하나의 스트림 식별자(Identifier; ID)와 관련되고, 적어도 하나의 SSD 및 상기 호스트 컨트롤러는 스트림 ID 명령을 실행하여 상기 논리적 볼륨을 액세스하는 시스템 코드를 포함하는 호스트 컨트롤러를 포함한다.
- [0011] 본 발명의 몇몇 실시예에서, 상기 각각의 논리적 볼륨은 상기 SSD 각각의 메모리 영역에 전용(dedicate)되고, 상기 SSD 각각의 메모리 영역에 대응하는 상기 각각의 논리적 볼륨은 상기 적어도 하나의 스트림 ID를 이용하는 논리적 볼륨 그룹을 형성하도록 조합될 수 있다.
- [0012] 본 발명의 몇몇 실시예에서, 상기 호스트 컨트롤러는 상기 스트림 ID 명령을 이용하여 상기 적어도 하나의 스트림 ID를 이용하여 고유하게(uniquely) 식별되는 논리적 볼륨 그룹과 통신할 수 있다.
- [0013] 본 발명의 몇몇 실시예에서, 상기 스트림 ID 명령은 스트림 ID 활성화 명령, 스트림 ID 유희 명령, 스트림 ID 상태 획득(get) 명령 중 하나일 수 있다.
- [0014] 본 발명의 몇몇 실시예에서, 상기 스트림 ID 유희 명령은 적어도 하나의 파워 모드(power mode) 동작을 정의하는 동작 유형(operation type), 가비지 컬렉터(garbage collector) 동작, 데이터 연산 동작 및 사용자 지정(custom) 동작과 관련될 수 있다.
- [0015] 본 발명의 몇몇 실시예에서, 상기 적어도 하나의 SSD 및 상기 호스트 컨트롤러는 상기 시스템 코드를 실행(implement)하고 운영 체제(operating system)에 의하여 로드된 RAID 드라이버와 공존할 수 있다.
- [0016] 본 발명의 몇몇 실시예에서, 상기 호스트 컨트롤러는, 상기 복수의 논리적 볼륨으로부터 요청을 수신하여 논리적 볼륨의 LBA(Logical Block Addressing)에 액세스하고, 상기 논리적 볼륨에 대응하는 스트림 ID를 결정하고, 비휘발성 랜덤 액세스 메모리(Non-Volatile Random Access Memory; NVRAM)로부터 상기 논리적 볼륨에 대응하는 스트림 ID의 상태를 검출하되, 상기 상태는 활성화(active) 또는 유희(passive) 중의 하나이고, 상기 스트림 ID의 상태가 활성화로 검출된 경우, 상기 논리적 볼륨의 상기 스트림 ID에 대응하는 상기 복수의 SSD 중 적어도 하나의 SSD의 상기 LBA에 대한 액세스를 허용하고, 상기 스트림 ID의 상태가 유희로 검출된 경우, 상기 논리적 볼륨의 상기 스트림 ID에 대응하는 상기 복수의 SSD 중 적어도 하나의 SSD에 대하여 가비지 컬렉션 기능을 수행할 수 있다.
- [0017] 본 발명의 몇몇 실시예에서, 조율(orchestration) 로직에 기초한 패리티 그룹(parity group)으로부터 상기 LBA에 대한 상기 적어도 하나의 SSD가 선택되고, 상기 패리티 그룹은 상기 LBA에 기초하여 결정될 수 있다.
- [0018] 본 발명의 몇몇 실시예에서, 상기 복수의 SSD로부터의 상기 적어도 하나의 SSD에 대한 상기 가비지 컬렉션 기능은 조율 시퀀스에 기초하여 수행될 수 있다.
- [0019] 본 발명의 몇몇 실시예에서, 상기 요청은 읽기 요청, 쓰기 요청 및 LBA 연산 요청 중 하나일 수 있다.
- [0020] 본 발명의 몇몇 실시예에서, 상기 NVRAM은 상기 LBA에 액세스하기 위한 상기 요청을, 상기 스트림 ID에 대응하는 상기 논리적 볼륨으로 전송하기 전에 버퍼링할 수 있다.
- [0021] 본 발명의 몇몇 실시예에서, 상기 호스트 컨트롤러는 상기 스트림 ID 명령에 기초하여 상기 NVRAM 내의 상기 각각의 스트림 ID의 상태를 설정할 수 있다.
- [0022] 상기 기술적 과제를 달성하기 위한 본 발명의 일 실시예에 따른 호스트 장치는, 프로세서, 및 상기 프로세서 및 SSD 어레이에 연결되는 컨트롤러를 포함하되, 상기 호스트 컨트롤러는, 상기 SSD 어레이를 복수의 논리적 볼륨으로 관리하되, 상기 각각의 논리적 볼륨은 적어도 하나의 스트림 ID와 관련되고, 상기 호스트 컨트롤러는 상기 각각의 논리적 볼륨에 액세스하기 위한 스트림 ID 명령을 실행하는 시스템 코드를 포함한다.
- [0023] 상기 기술적 과제를 달성하기 위한 본 발명의 일 실시예에 따른 SSD 어레이의 관리 방법은, 호스트 컨트롤러에 의해, SSD 어레이로부터 논리적 볼륨의 LBA에 액세스하는 요청을 수신하고, 상기 호스트 컨트롤러에 의해, 상기 논리적 볼륨에 대응하는 스트림 ID를 결정하고, 상기 호스트 컨트롤러에 의해, 상기 논리적 볼륨에 대응하는 상기 스트림 ID의 상태를 검출하되, 상기 상태는 활성화 및 유희 중 하나이고, 상기 호스트 컨트롤러에 의해, 상기 스트림 ID의 상태가 활성화로 검출된 경우 상기 논리적 볼륨의 상기 스트림 ID에 대응하는 상기 복수의 SSD 중

적어도 하나의 SSD의 상기 LBA에 대한 액세스를 허용하고, 상기 스트림 ID의 상태가 유틸로 검출된 경우, 상기 논리적 볼륨의 상기 스트림 ID에 대응하는 상기 복수의 SSD 중 적어도 하나의 SSD에 대하여 가비지 컬렉션 기능을 수행하는 것을 포함한다.

[0024] 기타 실시예들의 구체적인 사항들은 상세한 설명 및 도면들에 포함되어 있다.

도면의 간단한 설명

[0025] 도 1은 본 발명의 일 실시예에 따른 스트림 ID 기반 SSD 어레이(300)의 관리를 위한 스토리지 시스템의 개략도이다.

도 2는 본 발명의 일 실시예에 따른 스트림 ID 기반 IO 동작의 관리를 위한 호스트 장치의 블록도이다.

도 3은 본 발명의 실시예에 따른 스트림 ID에 기초하여 IO 동작을 관리하는 호스트 장치에 의하여 수행되는 동작들을 설명하기 위한 순서도이다.

도 4는 본 발명의 실시예에 따른 스트림 ID에 기초하여 IO 및 가비지 컬렉션 동작을 관리하는 호스트 장치에 의해 수행되는 동작을 설명하기 위한 순서도이다.

도 5는 본 발명의 일 실시예에 따른 SSD 어레이에 관련된 볼륨에 대한 IO 동작의 관리를 설명하기 위한 시퀀스 다이어그램이다.

도 6은 본 발명의 일 실시예에 따른 스트림 ID 그룹에 기초하여 각각의 SSD가 수행하는 전력 관리를 설명하기 위한 시퀀스 다이어그램이다.

도 7은 본 발명의 일 실시예에 따른 스트림 ID 그룹에 기초한 글로벌 가비지 컬렉션의 관리를 설명하기 위한 시퀀스 다이어그램이다.

도 8은 본 발명의 일 실시예에 따른 SSD 어레이의 가비지 컬렉션을 수행하며 높은 IOPS를 유지하는 시스템을 설명하기 위한 예시적인 도면이다.

도 9는 본 발명의 일 실시예에 따른 스트림 ID 그룹에 기초하여 스토리지 내 연산을 수행하면서 높은 IOPS를 유지하는 시스템을 설명하기 위한 예시적인 도면이다.

도 10은 본 발명의 실시예에 따른 각 SSD/스트림 ID 그룹의 볼륨의 인-컴퓨팅(in-computing)을 설명하기 위한 시퀀스 다이어그램이다.

도 11은 본 발명의 실시예에 따른 SSD 어레이를 관리하기 위한 스트림 ID 기반 스토리지 시스템을 구현하는 컴퓨팅 환경을 도시한다.

발명을 실시하기 위한 구체적인 내용

[0026] 본 발명의 이점 및 특징, 그리고 그것들을 달성하는 방법은 첨부되는 도면과 함께 상세하게 후술되어 있는 실시예들을 참조하면 명확해질 것이다. 그러나 본 발명은 이하에서 개시되는 실시예들에 한정되는 것이 아니라 서로 다른 다양한 형태로 구현될 것이며, 단지 본 실시예들은 본 발명의 개시가 완전하도록 하며, 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자에게 발명의 범주를 완전하게 알려주기 위해 제공되는 것이며, 본 발명은 청구항의 범주에 의해 정의될 뿐이다. 도면에서 표시된 구성요소의 크기 및 상대적인 크기는 설명의 명료성을 위해 과장된 것일 수 있다. 명세서 전체에 걸쳐 동일 참조 부호는 동일 구성 요소를 지칭하며, "및/또는"은 언급된 아이템들의 각각 및 하나 이상의 모든 조합을 포함한다.

[0027] 소자(elements) 또는 층이 다른 소자 또는 층의 "위(on)" 또는 "상(on)"으로 지칭되는 것은 다른 소자 또는 층의 바로 위뿐만 아니라 중간에 다른 층 또는 다른 소자를 개재한 경우를 모두 포함한다. 반면, 소자가 "직접 위(directly on)" 또는 "바로 위"로 지칭되는 것은 중간에 다른 소자 또는 층을 개재하지 않은 것을 나타낸다.

[0028] 공간적으로 상대적인 용어인 "아래(below)", "아래(beneath)", "하부(lower)", "위(above)", "상부(upper)" 등은 도면에 도시되어 있는 바와 같이 하나의 소자 또는 구성 요소들과 다른 소자 또는 구성 요소들과의 상관관계를 용이하게 기술하기 위해 사용될 수 있다. 공간적으로 상대적인 용어는 도면에 도시되어 있는 방향에 더하여 사용시 또는 동작시 소자의 서로 다른 방향을 포함하는 용어로 이해되어야 한다. 예를 들면, 도면에 도시되어 있는 소자를 뒤집을 경우, 다른 소자의 "아래(below)" 또는 "아래(beneath)"로 기술된 소자는 다른 소자의 "위(above)"에 놓여질 수 있다. 따라서, 예시적인 용어인 "아래"는 아래와 위의 방향을 모두 포함할 수 있다. 소자

는 다른 방향으로도 배향될 수 있고, 이에 따라 공간적으로 상대적인 용어들은 배향에 따라 해석될 수 있다.

- [0029] 본 명세서에서 사용된 용어는 실시예들을 설명하기 위한 것이며 본 발명을 제한하고자 하는 것은 아니다. 본 명세서에서, 단수형은 문구에서 특별히 언급하지 않는 한 복수형도 포함한다. 명세서에서 사용되는 "포함한다(comprises)" 및/또는 "포함하는(comprising)"은 언급된 구성요소 외에 하나 이상의 다른 구성요소의 존재 또는 추가를 배제하지 않는다.
- [0030] 비록 제1, 제2 등이 다양한 소자나 구성요소들을 서술하기 위해서 사용되나, 이들 소자나 구성요소들은 이들 용어에 의해 제한되지 않음은 물론이다. 이들 용어들은 단지 하나의 소자나 구성요소를 다른 소자나 구성요소와 구별하기 위하여 사용하는 것이다. 따라서, 이하에서 언급되는 제1 소자나 구성요소는 본 발명의 기술적 사상 내에서 제2 소자나 구성요소 일 수도 있음은 물론이다.
- [0031] 다른 정의가 없다면, 본 명세서에서 사용되는 모든 용어(기술 및 과학적 용어를 포함)는 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자에게 공통적으로 이해될 수 있는 의미로 사용될 수 있을 것이다. 또 일반적으로 사용되는 사전에 정의되어 있는 용어들은 명백하게 특별히 정의되어 있지 않는 한 이상적으로 또는 과도하게 해석되지 않는다.
- [0032] 본 발명의 실시예들은 설명된 기능 또는 기능들을 수행하는 블록들에 의해 기술되고 예시될 수 있다. 본 실시예에서 사용되는 사용되는 '부' 또는 '모듈'이라는 용어는 소프트웨어 또는 FPGA또는 ASIC과 같은 하드웨어 구성요소를 의미하며, '부' 또는 '모듈'은 어떤 역할들을 수행한다. 그렇지만 '부' 또는 '모듈'은 소프트웨어 또는 하드웨어에 한정되는 의미는 아니다. '부' 또는 '모듈'은 어드레싱할 수 있는 저장 매체에 있도록 구성될 수도 있고 하나 또는 그 이상의 프로세서들을 재생시키도록 구성될 수도 있다. 따라서, 일 예로서 '부' 또는 '모듈'은 소프트웨어 구성요소들, 객체지향 소프트웨어 구성요소들, 클래스 구성요소들 및 태스크 구성요소들과 같은 구성요소들과, 프로세스들, 함수들, 속성들, 프로시저들, 서브루틴들, 프로그램 코드의 세그먼트들, 드라이버들, 펌웨어, 마이크로코드, 회로, 데이터, 데이터베이스, 데이터 구조들, 테이블들, 어레이들, 및 변수들을 포함할 수 있다. 구성요소들과 '부' 또는 '모듈'들 안에서 제공되는 기능은 더 작은 수의 구성요소들 및 '부' 또는 '모듈'들로 결합되거나 추가적인 구성요소들과 '부' 또는 '모듈'들로 더 분리될 수 있다.
- [0033] 본 발명의 실시예에 따르면, 스트림 식별자(ID)에 기반하여 SSD 어레이를 관리하는 호스트 장치가 제공된다. 상기 호스트 장치는 프로세서, 프로세서 및 SSD 어레이에 연결된 호스트 컨트롤러로, 상기 호스트 컨트롤러는 SSD 어레이를 복수의 논리적 볼륨으로 관리하되, 각각의 논리적 볼륨은 적어도 하나의 스트림 ID와 관련되며, 복수의 논리적 볼륨으로부터 논리적 볼륨의 LBA에 액세스하기 위한 요청을 수신하고, 논리적 볼륨에 대응하는 스트림 ID를 결정하고, NVRAM으로부터 논리적 볼륨에 대응하는 스트림 ID의 상태를 검출한다. 상기 상태는 활성(active) 또는 유틸(passive) 중 하나이고, 스트림 ID의 상태가 활성으로 검출된 경우 논리적 볼륨의 스트림 ID에 대응하는 복수의 SSD로부터 적어도 하나의 SSD의 LBA에 대한 액세스를 허용하고, 스트림 ID의 상태가 유틸로 검출된 경우, 논리적 볼륨의 스트림 ID에 대응하는 복수의 SSD로부터 적어도 하나의 SSD에 대하여 가비지 컬렉션 기능을 수행한다.
- [0034] 본 발명의 실시예에 따르면, 스트림 ID에 기반하여 SSD 어레이를 관리하기 위한 스토리지 시스템이 제공된다. 스토리지 시스템은 각각이 데이터를 저장하는 SSD의 어레이와, SSD 어레이를 복수의 논리적 볼륨으로 관리하는 호스트 컨트롤러를 포함하고, 각각의 논리적 볼륨은 적어도 하나 이상의 스트림 ID와 관련되고, 각각의 SSD는 스트림 ID 명령을 실행하여 각각의 논리적 볼륨에 액세스한다.
- [0035] 종래의 스토리지 시스템(예를 들어 SSD)과는 달리, 제안된 스토리지 시스템은 SSD의 성능을 향상시키기 위하여 가비지 컬렉션을 관리하기 위한 로그 구조 어레이(log-structured array; LSA) RAID 및 SSD를 조정한다.
- [0036] 종래의 메커니즘에서, SSD 내부의 가비지 컬렉션은 플래시 어레이 시스템(예를 들어 SSD 어레이)의 IO 동작 또는 데이터 연산(스마트 SSD의 경우)에 영향을 미친다. 그러나, 제안된 메커니즘은 가비지 컬렉션 및 IO 동작 사이의 간섭을 피하기 위하여 호스트 컨트롤러 및 SSD의 NVRAM을 조정한다.
- [0037] 종래의 스토리지 시스템과는 달리, 스트림 IO 기반 볼륨 또는 논리 유닛 관리(Logical Unit Management; LUM)를 갖는 제안된 스토리지 시스템은 SSD 볼륨(즉, SSD 어레이)에 대한 기업 스토리지 관리에서의 IO 동작의 성능을 향상시킬 수 있다.
- [0038] 종래의 스토리지 시스템과 달리, 제안된 저장 시스템은 가비지 컬렉션 및 IO 동작의 중첩(overlapping)을 피하고 SSD 어레이의 각 볼륨/볼륨 그룹에 걸친 전력 관리를 구현한다.

- [0039] 종래의 스토리지 시스템과 달리, 제안된 스토리지 시스템은 성능을 손상시키지 않으면서 단일 스토리지 시스템 (즉, 단일 스토리지 서버) 상에서 OLTP 및 데이터 분석을 지원한다.
- [0040] 종래의 스토리지 시스템과 달리, 제안된 스토리지 시스템은 스트림 ID 명령을 사용하여 가비지 컬렉션 단계와 쓰기 단계를 동기화하는 쉬운 방법을 제안한다. 종래의 스토리지 시스템과 달리 제안된 스토리지 시스템은 IO 작업 중 쓰기 레이턴시를 감소시킨다.
- [0041] 종래의 스토리지 시스템과 달리, 제안된 스토리지 시스템은 모든 SSD로부터 최고의 성능을 얻기 위해 조율될 수 있는 활성(active) 및 유ힴ(idle) 스트림 ID를 제공한다.
- [0042] 종래의 스토리지 시스템과 달리, 제안된 스토리지 시스템은 SSD의 전력 관리를 향상시킬 수 있다. SSD 어레이는 활성 상태 및 유ힴ 상태에 있는 SSD 어레이의 볼륨/볼륨 그룹의 전원 상태를 조율할 수 있다.
- [0043] 종래의 스토리지 시스템과 달리, 제안된 스토리지 시스템은 종래의 스택(stack)에 비해 데이터 분석의 성능을 향상시킨다.
- [0044] 종래의 스토리지 시스템과 달리, 제안된 스토리지 시스템은 SSD 어레이 내의 SSD를 조정하기 위한 스트림 ID 기반 스토리지 시스템 관리이며, 기존 SSD 하드웨어에서 어떠한 변경도 요구하지 않는다. 종래의 메커니즘이 가비지 컬렉션으로 인하여 성능이 뒤떨어지는 반면, 제안된 메커니즘은 두 개의 가장 널리 사용되는 케이스, 즉 동일한 서버에서의 데이터 분석과 OLTP에서 성능을 손상시키지 않고 이를 수용할 수 있다. 제안된 SSD 어레이에서의 스트림 ID 기반 스토리지 시스템은 SSD의 내구성을 증가시킨다.
- [0045] 종래의 스토리지 시스템과 달리, 제안된 스토리지 시스템은 시스템의 쓰루풋(through-put)을 높게 유지하며 스토리지 시스템의 연산 용량을 향상시킨다.
- [0046] 도면, 더욱 구체적으로 도1 내지 도 4를 참고하면, 도면에 걸쳐 동일한 도면 부호는 동일한 구성 요소를 지칭하여 실시예들이 도시된다.
- [0047] 도 1은 본 발명의 일 실시예에 따른 스트림 ID 기반 SSD 어레이(300)의 관리를 위한 스토리지 시스템(100)의 개략도이다.
- [0048] 스토리지 시스템(100)은 인터페이스를 통해 연결된 호스트 장치(200) 및 SSD 어레이를 포함한다. 호스트 장치(200)는 들어오는(incoming) 요청(즉, 특정 애플리케이션에 대한 읽기/쓰기 요청)을 관리하도록 구성된다. 인터페이스는 호스트 장치(200)가 SSD 어레이(SSD)와 통신할 수 있도록 한다. 예를 들어, 인터페이스는 SAS 확장기 및 다른 SAS 및/또는 PCI(Peripheral Component Interface) 장치를 사용하는 SAS 토폴로지에서 발견되는 스위칭 패브릭(switched fabric)일 수 있다.
- [0049] SSD 어레이(300)는 복수의 SSD(예를 들어 SSD1, SSD2 및 SSD3)를 포함하는 데이터 스토리지 시스템이다. 또한, SSD 어레이는 복수의 논리적 볼륨(예를 들어 볼륨 1, 볼륨 2 및 볼륨 3)으로 분할된다. 각각의 논리적 볼륨은 SSD의 논리적 볼륨을 고유하게 식별하기 위하여 적어도 하나의 스트림 ID와 연관된다. 예를 들어, 볼륨 1에 대한 스트림 ID 1, 볼륨 2에 대한 스트림 ID 2, 볼륨 3에 대한 스트림 ID 3. 복수의 논리적 볼륨은 함께 논리적 볼륨 그룹을 형성한다. 논리적 볼륨 그룹과 설정된 스트림 ID는 맵핑 테이블(표 1에 도시된) 또는 트리 기반 데이터 구조의 형태로 유지될 수 있다.
- [0050] 본 발명의 일 실시예에서, 각각의 논리적 볼륨은 각각의 SSD 메모리 영역에 전용되고(dedicated), 각각의 SSD의 메모리 영역에 대응하는 논리적 볼륨 각각은 적어도 하나의 스트림 ID를 이용하여 논리적 볼륨 그룹을 형성하도록 결합된다. 호스트 장치(200)는 스트림 ID 명령을 이용하여, 적어도 하나의 스트림 ID를 이용하여 고유하게 식별되는 논리적 볼륨 그룹과 통신한다.

표 1

Volume ID	SSD, Stream ID
Vol-1	(SSD-0, Stream-ID-1), (SSD-1, Stream-ID-1)··· (SSD-10, Stream-ID-1)
Vol-2	(SSD-0, Stream-ID-2), (SSD-1, Stream-ID-2)··· (SSD-10, Stream-ID-2)
Vol-3	(SSD-0, Stream-ID-3), (SSD-1, Stream-ID-3)··· (SSD-10, Stream-ID-3)

- [0052] 본 발명의 일 실시예에서, 논리적 볼륨 그룹 및 설정된 스트림 ID의 맵핑 테이블은 데이터가 중복으로 저장됨을

나타낸다(예를 들어 패리티 그룹).

- [0053] 스트림 ID는 SSD의 각 논리적 볼륨에 할당된 고유 ID이다. SSD는 내부적으로 고유한 스트림 ID를 컨텍스트로써 사용하여, 동일한 기대 수명(lifetime-expectancy)을 갖는 데이터를 관리한다. 데이터의 기대 수명에 따라 각 SSD와 관련된 볼륨의 스트림 ID에 데이터가 쓰여진다. 스트림 ID는 실질적으로 동일한 시간 동안 SSD에 저장될 것으로 예상되는 관련 파일 그룹을 기록하기 위한 기대 수명 값으로 사용된다.
- [0054] SSD는 스트림 ID에 속하는 모든 데이터를 함께 저장하고 다른 스트림으로부터의 데이터를 혼합하지 않는다. 호스트 장치(200)는 SSD의 논리적 볼륨의 스트림 ID에 대응하는 데이터를 오픈/판독/기록할 수 있다. 논리적 볼륨에 속하는 각각의 SSD는 복수의 큐(queue)를 가질 수 있다(예를 들어, Q1, Q2 및 Q3 등). SSD 어레이(300)와 관련된 큐들은 호스트 장치(200)로부터 수신되어 들어오는 IO 요청들을 저장하도록 구성된다.
- [0055] 호스트 장치(200)는 스트림 ID 명령을 사용하여, 적어도 하나의 스트림 ID를 사용하여 고유하게 식별되는 논리적 볼륨 그룹과 통신한다. 스트림 ID 명령은 다음과 같이 주어진다:
- [0056] 스트림 ID 활성화 명령(Stream ID): 적어도 하나의 SSD 및 호스트 장치(200)에 존재하는 시스템 코드는 특정한 스트림 ID를 활성화로 설정하기 위하여 스트림 ID 활성화 명령을 호출(involve)할 수 있으며, 여기서 스트림 ID는 SSD 어레이(300)로부터 각각의 SSD의 논리적 볼륨과 관련된다. 스트림 ID 활성화는 IO가 진행 중임을 나타낸다(또는, 예를 들어 IO가 예상될 수 있음). 시스템 코드는 스트림 ID 명령을 실행하는 명령 또는 프로토콜의 세트, 펌웨어일 수 있다.
- [0057] 스트림 ID 유희 명령(stream ID, int TYPE): 적어도 하나의 SSD 및 호스트 장치(200)에 존재하는 시스템 코드는 스트림 ID 유희 명령을 호출하여 특정한 스트림 ID를 유희로 설정할 수 있다. 스트림 ID 유희는 블록/스트라이프와 관련된 스트림 ID를 가비지 컬렉션 또는 관리(maintenance) 활동에 적용할 수 있음을 의미한다. SSD의 유희 스트림 ID에서의 전원 관리는 저전력 모드로 유지될 수 있다.
- [0058] 스트림 ID 유희 명령의 유형(TYPE)은 실행될 유희 명령을 결정한다.
- [0059] 예를 들어:
- [0060] - POWER DOWN - POWER DOWN/Power mode ON 유형을 갖는 유희 명령은 저전력 상태를 나타낸다.
- [0061] - 데이터 연산 - 데이터 커뮤테이션(commutation) 유형을 갖는 유희 명령은 관련 데이터를 실행하여 스트림 ID에서 프로그램을 연산한다.
- [0062] - 글로벌 GC - 글로벌 가비지 컬렉션 유형을 갖는 유희 명령은 스트림 ID에서 가비지 컬렉션을 실행할 수 있다.
- [0063] - 사용자 지정(custom) - 사용자 지정 타입의 유희 명령은 확장 목적을 위해 특정한 명령을 실행할 수 있다.
- [0064] 스트림 ID 상태 획득 명령(stream ID): 적어도 하나의 SSD 및 호스트 장치(200)에 존재하는 시스템 코드는 이 명령을 호출하여 IO 동작의 현재 사이클에서의 스트림 ID의 상태 정보를 결정한다.
- [0065] 스트림 ID 명령은 RAID 명령을 이용함으로써 SSD 드라이버 또는 SSD 그룹(SSD 어레이(300))에 의하여 단일 SSD에 대하여 사용될 수 있다. 스트림 ID 명령은 호스트 가비지 컬렉션을 수행하는 RAID와 호스트 가비지 컬렉션을 수행하지 않는 RAID에 의하여 사용될 수 있다. 스트림 ID 명령은 시스템 코드 및 SSD 모두에 의하여 유지 영역(즉, 가비지 컬렉션 영역)을 식별하는 것을 도울 수 있으므로, 상기 영역으로의 쓰기를 피하여 가비지 컬렉션과 쓰기 사이의 중첩을 제거할 수 있다.
- [0066] 종래의 스토리지 시스템과는 달리, 제안된 스토리지 시스템(100)은 활성화 및 유희 스트림 ID를 조율하여 모든 SSD로부터 최적의 성능을 얻을 수 있다.
- [0067] 도 2는 본 발명의 일 실시예에 따른 스트림 ID 기반 IO 동작의 관리를 위한 호스트 장치(200)의 블록도이다.
- [0068] 본 발명의 일 실시예에서, 호스트 장치(200)는 프로세서(202), 애플리케이션 구성 요소(204), 호스트 컨트롤러(206), IO 컨트롤러(208), 메모리(210), 가비지 컬렉터(212) 및 통신기(214)를 포함한다. 프로세서(202)와 SSD 어레이(300)에 연결된 호스트 컨트롤러(206)는 SSD 어레이(300)를 복수의 논리적 볼륨으로 관리하도록 구성된다. 논리적 볼륨 각각은 적어도 하나의 스트림 ID와 관련된다. IO 컨트롤러(208)에 연결된 애플리케이션 구성 요소(204)는 복수의 논리적 볼륨으로부터 논리적 볼륨의 LBA에 액세스 하기 위한 IO 요청을 수신하도록 구성될 수 있다.
- [0069] 호스트 장치(200)의 메모리(210)는, 맵핑 테이블의 형태로 유지되어 SSD 어레이(300)의 논리적 볼륨을 고유하게

식별할 수 있는 논리적 볼륨 그룹과 설정된 스트림 ID의 관계를 저장하도록 구성된다.

- [0070] 또한, 호스트 컨트롤러(206)는 논리적 볼륨에 대응하는 스트림 ID를 결정하도록 구성된다. 또한, 호스트 컨트롤러(206)는 NVRAM으로부터 논리적 볼륨에 대응하는 스트림 ID의 상태를 검출하도록 구성된다. 상기 상태는 활성 또는 유휴 상태 중 하나이다. 또한, IO 컨트롤러(208)는 스트림 ID의 상태가 활성 상태로 검출된 경우, 논리적 볼륨의 스트림 ID에 대응하는 복수의 SSD 중 적어도 하나의 SSD의 LBA에 액세스하는 것을 허용한다. 호스트 컨트롤러(206)가 유휴 상태의 스트림 ID를 검출한 경우, 가비지 컬렉터(212)는 논리적 볼륨의 스트림 ID에 대응하는 복수의 SSD 중 적어도 하나에 대하여 가비지 컬렉션 기능을 수행할 수 있다.
- [0071] 통신기(214)는 인터페이스를 통하여 호스트 장치(200)와 SSD 어레이(300) 사이의 통신을 설정하여 IO 요청을 서비스하도록 구성된다.
- [0072] 도 2는 호스트 장치(200)의 개요를 도시하지만, 본 발명이 이에 제한되는 것은 아니다. 본 발명의 다른 실시예에서, 호스트 장치(200)는 더 적거나 많은 수의 장치 및 구성 요소를 포함할 수 있다. 또한, 구성 요소들의 레이블 또는 명칭은 단지 예시적인 목적으로 사용된 것이며, 본 발명의 범위를 제한하지 않는다. 하나 이상의 구성 요소가 호스트 장치(200)에서 동일하거나 실질적으로 유사한 기능을 수행하도록 함께 결합될 수도 있다.
- [0073] 도 3은 본 발명의 실시예에 따른 스트림 ID에 기초하여 IO 동작을 관리하는 호스트 장치(200)에 의하여 수행되는 동작들을 설명하기 위한 순서도(S300)이다.
- [0074] 단계 S302에서, 호스트 컨트롤러(206)는 SSD의 복수의 논리적 볼륨으로부터 논리적 볼륨의 LBA에 액세스 하기 위한 요청(즉, 애플리케이션으로부터 읽기/쓰기 요청)을 수신할 수 있다.
- [0075] 단계 S304에서, 호스트 컨트롤러(206)는 SSD의 논리적 볼륨에 대응하는 스트림 ID를 결정할 수 있다.
- [0076] 단계 S306에서, 호스트 컨트롤러(306)는 조율 로직(orchestration logic)을 트리거링하여 요청의 유형(읽기/쓰기 요청, 연산 요청 또는 가비지 컬렉션 요청)을 결정할 수 있다.
- [0077] 단계 S308에서, 조율 로직에 연결된 호스트 컨트롤러가 스트림 ID 활성 상태에 기초하여 읽기/쓰기 요청인지를 결정할 수 있다.
- [0078] 단계 S310에서, 호스트 컨트롤러(206)는 스트림 ID가 NVRAM에 캐시되었는지 여부를 검출할 수 있다.
- [0079] 만약 호스트 컨트롤러(206)가 스트림 ID가 NVRAM에 캐시된 것으로 결정하면, 단계 S312에서 호스트 컨트롤러(206)는 NVRAM으로부터 SSD의 논리적 볼륨으로 IO 동작을 계속하도록 요청(읽기/쓰기)을 허용할 수 있다. 만약 호스트 컨트롤러(206)가 스트림 ID가 NVRAM에 캐시되지 않은 것으로 결정한다면, 단계 S316에서 호스트 컨트롤러(206)는 조율 시퀀스에 기초하여 입출력(IO) 동작을 플러싱(flush)할 수 있다.
- [0080] 단계 S312에서, 호스트 컨트롤러(206)는 조율 로직에 기초하여 요청 유형이 연산 요청(SSD, 스트림 ID)인지 여부를 결정하고, 호스트 컨트롤러(206)는 SSD에 대한 LBA와 스트림 ID를 연산할 수 있다. 또한, 조율 로직에 기초하여 호스트 컨트롤러(206)는 스트림 ID가 활성/유휴 상태인지 여부를 식별한다. 만약 호스트 컨트롤러(206)가 스트림 ID가 활성 상태에 있는 것으로 결정하면, 호스트 컨트롤러(206)는 조율 로직에 기초하여 IO를 플러싱한다. 만약 호스트 컨트롤러(206)가 스트림 ID가 유휴 상태에 있는 것으로 결정하면, 호스트 컨트롤러(206)는 가비지 컬렉션을 수행한다.
- [0081] 단계 S314에서, 호스트 컨트롤러(206)는 조율 로직에 기초하여 스트림 ID가 유휴 상태에 있는 것으로 결정할 수 있고, 호스트 컨트롤러(206)는 가비지 컬렉션을 수행한다.
- [0082] 본 발명의 일 실시예에서, 조율 로직은 이하의 예시와 같이 설명될 수 있다.
- [0083] 입력:
- [0084] i0: Read/Write Volume Set = {V0_io, V1_io?}
- [0085] i1: Compute Set = {V0_c, V1_c?} One of both sets can be null set
- [0086] 출력:
- [0087] o0: Read/Write Set = {(OD0_io, OS0_io), (OD1_io, OS1_io)?}
- [0088] o1: Compute Set = {(OD0_c, OS0_c), (OD1_c, OS1_c)?}

- [0089] o2: Garbage collection Set = {(OD0_g, OS0_g), (OD1_g, OS1_g)?}
- [0090] o3: Power down set = {(OD0_p, OS0_p), (OD1_p, OS1_p)?}
- [0091] 상술한 예시에 대한 조율 로직의 상세한 흐름은 이하의 단계에서 자세하게 설명된다.
- [0092] 1. i0의 각 볼륨에 대하여: 볼륨-스트림 ID 맵으로부터 디스크(D_io, S_io)를 검색한다.
- [0093] 2. o0 세트에 (D_io, S_io)를 더한다: $o0 = o0 \cup \{(D_io, S_io)\dots\}$
- [0094] 3. i1의 각 볼륨에 대하여: 볼륨-스트림 ID 맵으로부터 디스크(D_c, S_c)를 검색한다.
- [0095] 4. o1 세트에 (D_c, S_c)를 더한다: $o1 = o1 \cup \{(D_c, S_c)\dots\}$
- [0096] 5. o0과 o1 사이에서 공통의 (Disk, Stream ID)를 검색하고 이를 o0과 o1 세트 모두에서 이를 제거한다.
- [0097] 6. o0과 o1 사이의 공통 요소를 분산시킨다. 이 단계는 주어진 볼륨에 대하여, 일부는 IO 요청을 가지며 다른 일부는 동시에 LBA에 대한 연산 요청을 갖는 것을 보장한다. IO 요청의 분배 및 연산 요청은 활성 카운트(count) 또는 IOPS 카운트를 이용할 수 있다.
- [0098] 7. 볼륨-스트림 ID 맵으로부터, 빈 페이지(free page) 카운트가 낮고 o0에 존재하지 않는 디스크를 검색한다. 검색된 디스크에 속하는 모든 스트림들을 o2(가비지 컬렉션 세트)에 더한다. 만약 디스크에 빈 페이지가 적고 o0에도 존재한다면, o0에서 검색된 디스크의 모든 스트림을 제거하고 가비지 컬렉션이 종료된 이후 o0에 더한다.
- [0099] 8. 볼륨-스트림 ID 맵으로부터, 활성 스트림 카운트가 0인 스트림 ID를 검색하고, 검색된 디스크에 속하는 모든 스트림을 o3(파워 다운 세트)에 더한다.
- [0100] 순서도 (S300)의 방법에 포함된 동작, 행동, 블록, 단계 등은 기술된 순서나, 이와 다른 순서 또는 동시에 수행될 수 있다. 또한, 본 발명의 몇몇 실시예에서, 이러한 동작, 행동, 블록, 단계 등은 본 발명의 기술적 사상의 범위 내에서 생략, 부가, 변경, 스킵될 수 있다.
- [0101] 도 4는 본 발명의 실시예에 따른 스트림 ID에 기초하여 IO 및 가비지 컬렉션 동작을 관리하는 호스트 장치(200)에 의해 수행되는 동작을 설명하기 위한 순서도(S400)이다.
- [0102] 단계 S402에서, 방법은 복수의 논리적 볼륨으로부터 논리적 볼륨의 LBA에 액세스하기 위한 요청을 수신하는 단계를 포함한다. 본 발명의 일 실시예에서, 방법은 호스트 컨트롤러(206)가 복수의 논리적 볼륨으로부터 논리적 볼륨에 액세스하기 위한 요청을 수신하도록 한다.
- [0103] 단계 S404에서, 방법은 논리적 볼륨에 대응하는 스트림 ID를 결정하는 것을 포함한다. 본 발명의 일 실시예에서, 방법은 호스트 컨트롤러(206)가 논리적 볼륨에 대응하는 스트림 ID를 결정하도록 한다.
- [0104] 단계 S406에서, 방법은 NVRAM으로부터 논리적 볼륨에 대응하는 스트림 ID의 상태를 검출하는 것을 포함하고, 여기서 검출된 상태는 활성 상태 및 유틸 상태 중의 하나이다. 본 발명의 일 실시예에서, 방법은 호스트 컨트롤러(206)가 NVRAM으로부터 논리적 볼륨에 대응하는 스트림 ID의 상태를 검출하는 것을 포함하고, 여기서 검출된 상태는 활성 상태 및 유틸 상태 중의 하나이다.
- [0105] 단계 S408에서, 방법은 스트림 ID의 상태가 활성 상태로 검출된 경우 논리적 볼륨의 스트림 ID에 대응하는 복수의 SSD 중 적어도 하나의 SSD의 LBA의 액세스를 허용하고, 스트림 ID의 상태가 유틸 상태로 검출된 경우 논리적 볼륨의 스트림 ID에 대응하는 복수의 SSD 중 적어도 하나의 SSD에 가비지 컬렉션 기능을 수행한다. 본 발명의 일 실시예에서, 방법은, 스트림 ID의 상태가 활성 상태로 검출된 경우 논리적 볼륨의 호스트 컨트롤러(206)가 스트림 ID에 대응하는 복수의 SSD 중 적어도 하나의 SSD의 LBA의 액세스를 허용하도록 하고, 스트림 ID의 상태가 유틸 상태로 검출된 경우 호스트 컨트롤러(206)가 논리적 볼륨의 스트림 ID에 대응하는 복수의 SSD 중 적어도 하나의 SSD에 가비지 컬렉션 기능을 수행하도록 한다.
- [0106] 순서도(S400)의 방법에 포함된 동작, 행동, 블록, 단계 등은 기술된 순서나, 이와 다른 순서 또는 동시에 수행될 수 있다. 또한, 본 발명의 몇몇 실시예에서, 이러한 동작, 행동, 블록, 단계 등은 본 발명의 기술적 사상의 범위 내에서 생략, 부가, 변경, 스킵될 수 있다.
- [0107] 도 5는 본 발명의 일 실시예에 따른 SSD 어레이(300)에 관련된 볼륨에 대한 IO 동작의 관리를 설명하기 위한 시퀀스 다이어그램이다.

- [0108] 애플리케이션 구성 요소(204)는 복수의 논리적 볼륨으로부터 논리적 볼륨의 LBA에 액세스하기 위하여 IO 요청(즉, 읽기/쓰기)을 개시한다. 또한, 호스트 컨트롤러(206)는 복수의 논리적 볼륨으로부터 논리적 볼륨의 LBA에 액세스하기 위하여 애플리케이션 구성 요소(204)로부터의 IO 요청을 수신한다. 또한, 호스트 컨트롤러(206)는 SSD 어레이(300)와 관련된 논리적 볼륨에 대한 스트림 ID 맵(메모리(210)에 저장되고 표 1에 기술된)을 탐색(lookup)(504)하도록 구성된다. 맵핑 테이블은 SSD 어레이(300)의 논리적 볼륨 그룹과 스트림 ID의 관계를 나타낸다.
- [0109] 또한, 호스트 컨트롤러(206)의 NVRAM은 스트림 ID에 대응하는 논리적 볼륨에 요청을 보내기 전에 요청을 버퍼링한다. 호스트 컨트롤러(206)는 SSD 어레이(300)와 관련된 복수의 논리적 볼륨으로부터 논리적 볼륨의 LBA에 대한 패리티 그룹을 결정(506)할 수 있다. 호스트 컨트롤러(206)는 또한 조율 로직에 기초하여 결정된 패리티 그룹(즉, SSD 어레이(300)와 관련된 볼륨 그룹)으로부터 대응하는 LBA에 대한 SSD를 선택한다(508). SSD의 플래시 어레이는 각각의 스트림 ID 그룹에 관한 이하의 정보를 유지한다: 스트림 ID 그룹에 대한 증가(increment) 활성화 카운트, 유틸 상태 스트림 ID 그룹 및 QoS(Quality of Service)에 대한 스트림 ID 그룹의 IOPS(Input Output Per Second).
- [0110] 호스트 컨트롤러(206)는 NVRAM으로부터 논리적 볼륨의 LBA에 대응하는 스트림 ID 또는 스트림 ID 그룹의 상태(즉, 활성화)(stream ID, power mode OFF)를 검출(510)하며, 파워 모드 OFF는 스트림 ID 활성화 중의 전력 모드를 나타낸다. SSD 어레이(300)는 SSD 어레이(300)와 관련된 논리적 볼륨의 LBA에서 결정된 활성화 상태에 기초하여 호스트 컨트롤러(206)로 성공적인 응답(acknowledgment)을 제공한다. 논리적 볼륨과 관련된 스트림 ID의 활성화 상태는 고성능 상태를 가리킨다(514).
- [0111] 또한, 호스트 컨트롤러(206)는 애플리케이션 요청(즉, 읽기/쓰기)을 통해 복수의 SSD 중 논리적 볼륨의 스트림 ID에 대응하는 SSD의 LBA에 액세스하고(516), 스트림 ID 그룹에 대한 활성화 카운트를 감소시킨다. 또한, SSD 어레이는 호스트 컨트롤러(206)의 요청에 성공적으로 응답하고(518), 호스트 컨트롤러는 애플리케이션 구성 요소(204)의 요청에 성공적으로 응답한다(520).
- [0112] 도 6은 본 발명의 일 실시예에 따른 스트림 ID 그룹에 기초하여 각각의 SSD가 수행하는 전력 관리를 설명하기 위한 시퀀스 다이어그램이다.
- [0113] 애플리케이션 구성 요소(204)는 복수의 논리적 볼륨으로부터 논리적 볼륨의 LBA에 액세스하기 위한 LBA 연산 요청을 개시한다. 또한, 호스트 컨트롤러(206)는 애플리케이션 구성 요소(204)로부터 연산 요청을 수신하고(602), SSD 어레이와 관련된 논리적 볼륨의 LBA를 연산한다. 또한, 호스트 컨트롤러(206)는 SSD 어레이(300)와 관련된 논리적 볼륨에 대한 스트림 ID 맵을 탐색한다(604).
- [0114] 호스트 컨트롤러(206)는 SSD 어레이(300)와 관련된 복수의 논리적 볼륨으로부터 논리적 볼륨의 LBA에 대한 패리티 그룹을 결정한다(606). 호스트 컨트롤러(206)는 조율 로직에 기초하여 결정된 패리티 그룹(즉, SSD 어레이(300)와 관련된 볼륨 그룹)으로부터 LBA에 대한 SSD를 선택한다(608). 호스트 컨트롤러(206)는 NVRAM으로부터 논리적 볼륨의 LBA에 대응하는 스트림 ID 또는 스트림 ID 그룹의 상태(즉, 유틸(Stream ID, power mode ON)를 검출한다(610). 논리적 볼륨과 관련된 스트림 ID의 유틸 상태는 스트림 ID 그룹이 절전 모드에 있는 것을 표시하고(612), SSD는 가비지 컬렉션과 같은 내부 유지 보수 활동을 수행할 수 있다.
- [0115] SSD 어레이(300)는 호스트 컨트롤러(206)로 SSD 어레이(300)와 관련된, 결정된 스트림 ID 그룹의 유틸 상태의 결과에 따라 또한, 호스트 컨트롤러(206)는 애플리케이션 구성 요소(204)의 요청에 성공적으로 응답한다.
- [0116] 도 7은 본 발명의 일 실시예에 따른 스트림 ID 그룹에 기초한 글로벌 가비지 컬렉션의 관리를 설명하기 위한 시퀀스 다이어그램이다.
- [0117] 애플리케이션 구성 요소(204)는 복수의 논리적 볼륨으로부터 논리적 볼륨의 LBA에 액세스하기 위한 연산 LBA 요청을 개시한다. 또한, 호스트 컨트롤러(206)는 애플리케이션 구성 요소(204)로부터 연산 요청을 수신하고(702), SSD 어레이(300)와 관련된 볼륨의 LBA를 연산한다. 또한, 호스트 컨트롤러(206)는 SSD 어레이(300)와 관련된 논리적 볼륨에 대한 스트림 ID 맵을 탐색한다(704). 호스트 컨트롤러(206)는 SSD 어레이(300)와 관련된 복수의 논리적 볼륨으로부터 논리적 볼륨의 LBA에 대한 패리티 그룹을 결정한다(706). 호스트 컨트롤러(206)는 또한 조율 로직에 기초하여 결정된 패리티 그룹(즉, SSD 어레이(300)와 관련된 볼륨 그룹)으로부터 LBA에 대한 SSD를 선택한다(708). 호스트 컨트롤러(206)는 NVRAM으로부터 논리적 볼륨의 LB에 대응하는 스트림 ID 또는 스트림 ID 그룹의 상태(즉, 유틸(Stream ID, Power mode ON)를 검출한다(710). 논리적 볼륨과 관련된 스트림 ID의 유틸 상태는 스트림 ID 그룹이 절전 모드에 있는 것을 표시한다. 호스트 컨트롤러(206)는 호스트 장치(200) 측에서 가

비지 컬렉션을 수행할 수 있고(712)/SSD는 가비지 컬렉션과 같은 내부 유지 보수 활동을 수행할 수 있다. SSD 어레이(300)는 SSD 어레이(300)와 관련된 스트림 ID 그룹의 결정된 유희 상태의 통합된(716) 결과들을 호스트 컨트롤러(206)에 성공적으로 응답한다(714). 또한, 호스트 컨트롤러(206)는 애플리케이션 구성 요소(204)에 대한 요청에 응답한다(718).

- [0118] 본 발명의 일 실시예에서, 인-컴퓨팅(In-compute) 스토리지를 위한 조율 로직은 아래와 같이 설명된다:
- [0119] 입력: 주어진 LBA에 대응하는 (SSD, 스트림 ID)의 세트를 포함하는 패리티 그룹. 예를 들어, 두 개의 SSD에서의 리던던시(redundancy)는 다음의 세트와 같다. {(SSD0, S0), (SSD1, S1) ...}
- [0120] 출력:
- [0121] 1. IO가 지시하는 (SSD0, StreamID0)의 세트(하나 이상일 수 있음)
- [0122] 2. 리듀스(Reduce)가 지시하는 (SSD1, StreamID1)
- [0123] 상술한 예시에 따른 조율 로직의 상세한 흐름은 다음 단계들에서 자세하게 설명된다.
- [0124] 1. 주어진 세트에서 최소 어카운트(account)를 갖는 스트림 ID를 찾는다. (Dmin, Smin)을 SSD, 스트림 ID 튜플(tuple)이라고 한다.
- [0125] 2. 감소 작업(reduce operation)에 대하여, Dmin, Smin을 반환한다.
- [0126] 3. 입력 패리티 그룹에서 Dmin, Smin을 제거하여 새로운 그룹을 만든다.
- [0127] 4. IO 동작에서, IO 연산을 수행할 새로운 그룹을 반환한다.
- [0128] 5. 감소 작업이 완료되면, (Dmin, Smin)에 대하여 IO를 수행한다. 본 발명의 다른 실시예에서, IOPS (Stream ID)는 활성 카운트 대신에 사용될 수 있다. 또한, 최소 감소 카운트(minimum reduce count)를 갖는 SSD를 반환한다. 이는 또한 활성 카운트, 감소 카운트 또는 IOPS의 조합과도 작동할 수 있다.
- [0129] 도 8은 본 발명의 일 실시예에 따른 SSD 어레이의 가비지 컬렉션을 수행하며 높은 IOPS를 유지하는 시스템을 설명하기 위한 예시적인 도면이다.
- [0130] 호스트 컨트롤러(206)는 서로 다른 애플리케이션(예를 들어, 애플리케이션 1, 애플리케이션 2, 애플리케이션 3 및 애플리케이션 n)으로부터 복수의 요청을 수신하여 SSD의 복수의 논리적 볼륨의 LBA를 액세스하도록 구성된다. 호스트 컨트롤러(206)는 또한 SSD 어레이(300)와 관련된 복수의 논리적 볼륨에 대응하는 스트림 ID (예를 들어 볼륨 1 과 스트림 ID 1, 볼륨 2와 스트림 ID 2, 볼륨 3과 스트림 ID 3 및 볼륨 n과 스트림 ID n)를 결정하도록 구성된다.
- [0131] 본 발명의 일 실시예에서, 각각의 논리적 볼륨은 스트림 ID 또는 스트림 ID의 세트와 관련된다. 스트림 ID는 SSD 어레이(300)의 개별 SSD로부터 얻어진다. 예를 들어, 각각의 스트림 ID는 전용 큐(dedicated queue)이거나 스트림 ID의 그룹이 큐를 공유할 수 있다. 스트림 ID는 애플리케이션 요청의 우선 순위와 관련될 수도 있다. 호스트 컨트롤러(206)는 각각의 스트림 ID에 스레드를 할당하여 논리적 볼륨에 대응하는 서로 다른 애플리케이션과 관련된 스트림 ID의 상태를 검출한다. 호스트 컨트롤러(206)의 NVRAM은 들어오는 애플리케이션 요청을 스트림 ID에 기초하여 SSD의 적절한 볼륨으로 전송하기 전에 버퍼링할 수 있다. 스트림 ID의 버퍼링은 호스트 컨트롤러(206)의 스레드가 스트림 ID가 활성 상태(IO 동작) 또는 유희 상태(가비지 컬렉션)에 있는지 여부를 결정하도록 한다.
- [0132] 만약 호스트 컨트롤러(206)의 스레드가 스트림 ID가 활성 상태임을 검출하면, 호스트 컨트롤러(206)는 애플리케이션 구성 요소(204)가 복수의 SSD 중 하나의 SSD의 LBA에 액세스하도록 한다. 또한, 호스트 컨트롤러(206)의 스레드가 스트림 ID가 유희 상태임을 검출하면, 호스트 컨트롤러(206)는 가비지 컬렉터(212)가 검출된 논리적 볼륨의 스트림 ID에 대하여 가비지 컬렉션 기능을 수행하도록 한다.
- [0133] 도 9는 본 발명의 일 실시예에 따른 스트림 ID 그룹에 기초하여 스토리지 내 연산을 수행하면서 높은 IOPS를 유지하는 시스템을 설명하기 위한 예시적인 도면이다.
- [0134] 빅 데이터/데이터 분석(big data/data analysis)과 OLTP는 두 개의 서로 다른 스토리지 서버에 의하여 서비스된다. 빅 데이터 분석은 비즈니스 보고에 사용되며 OLTP는 빠른 거래 세부사항을 제공하기 위해 ATM(Automated Teller Machine)과 같은 온라인 거래 처리 애플리케이션에 사용된다. OLTP 애플리케이션은 더 빠른 IOPS 서버를 이용하고, 빅 데이터/데이터 분석은 더 빠른 서버를 사용하지 않는다. 그 결과 빅 데이터/데이터 분석과 OLTP는

두 개의 서로 다른 스토리지 서버가 사용된다. 그러나, 제안된 본 발명의 스토리지 서버에 따르면 빅 데이터/데이터 분석과 OLTP는 단일 스토리지 서버로부터 서비스될 수 있다.

- [0135] 호스트 컨트롤러(206)는 빅 데이터/데이터 분석과 OLTP와 같은 두 개의 서로 다른 정보 시스템으로부터 복수의 요청을 받도록 구성된다. 예를 들어, 애플리케이션 1, 애플리케이션 2, 애플리케이션 3과 애플리케이션-n은 SSD 어레이(300)의 복수의 논리적 볼륨의 LBA에 액세스하기 위하여 호스트 장치(200)에 요청을 전송할 수 있다. 호스트 컨트롤러(206)는 또한 빅 데이터/데이터 분석으로부터의 요청(즉, 쿼리)을 수신한다.
- [0136] 호스트 컨트롤러(206)는 또한 SSD 어레이(300)와 관련된 복수의 논리적 볼륨에 대응하는 서로 다른 OLTP 애플리케이션에 대한 스트림 ID(예를 들어 볼륨 1과 스트림 ID 1, 볼륨 2와 스트림 ID 2, 볼륨 3과 스트림 ID3 및 볼륨 n과 스트림 ID n)를 결정한다. 호스트 컨트롤러(206)는 각각의 스트림 ID에 스레드를 할당하여 논리적 볼륨에 대응하는 서로 다른 OLTP 애플리케이션과 관련된 스트림 ID의 상태를 검출한다.
- [0137] 호스트 컨트롤러(206)의 NVRAM은 들어오는 애플리케이션 요청을 적절한 스트림 ID로 전송하기 전에 버퍼링할 수 있다. 스트림 ID의 버퍼링은 호스트 컨트롤러(206)의 스레드가, 스트림 ID가 활성 상태 또는 유휴 상태에 있는지 여부를 결정하도록 한다. 만약 호스트 컨트롤러(206)의 스레드가 스트림 ID가 활성 상태에 있는 것을 검출하면, 호스트 컨트롤러(206)의 스레드는 애플리케이션이 논리적 볼륨의 스트림 ID에 대응하는 복수의 SSD 중 적어도 하나의 SSD의 LBA에 액세스하도록 한다. 또한, 호스트 컨트롤러(206)의 스레드가 스트림 ID가 유휴 상태에 있는 것을 검출하면, 호스트 컨트롤러(206)의 스레드는 가비지 컬렉터가 논리적 볼륨의 검출된 스트림 ID에 대하여 가비지 컬렉션을 수행하도록 한다.
- [0138] OLTP 애플리케이션은 더 빠른 IOPS를 요구하며, 따라서 OLTP 애플리케이션은 더 높은 우선 순위 큐에 배치되고 빅 데이터 분석 쿼리 요청은 더 낮은 우선 순위 큐에 배치될 수 있다. SSD 어레이(300)는 복수의 SSD를 알고 있으므로, IO 작업이 진행되는 SSD와 데이터 분석이 발생하는 SSD를 조율할 수 있다. 이 동안 SSD의 논리적 볼륨에 단 하나의 복사본만이 쓰여지기 때문에 애플리케이션 데이터는 완전히 보호되지 않으며 이는 다른 SSD로의 쓰기를 지연함으로써 해결될 수 있다. 그 동안 전력 상실(power loss)이 발생한다면, 데이터는 NVRAM으로부터 복구될 수 있다.
- [0139] 제안된 스토리지 시스템은 인-컴퓨팅 스토리지 및 빅 데이터에 적용될 수 있다. 일련의 스트림 ID는 높은 IOPS를 요구하는 OLTP 기반 애플리케이션에 전용된(dedicated) 높은 우선 순위 큐와 관련될 수 있다. 낮은 우선 순위 큐와 관련된 스트림 ID는 데이터 분석 작업의 서비스에 전용될 수 있다. 데이터 분석 쿼리는 특수 모듈(리듀스 모듈 또는 맵 리듀스(Map Reduce))로 불리는 모듈의 도움으로 스트라이핑 로직에 따라 분할된다. 이러한 하위 쿼리(sub query)는 각각의 SSD에 대하여 각 스트림 ID로 전송된다.
- [0140] SSD 어레이(300)는 중복 방식으로 데이터를 저장한다; 패리티 그룹을 형성하는 SSD의 스트림 ID와 관련된 논리적 볼륨의 그룹에 대하여 동일한 데이터가 이용 가능하다. SSD 어레이(300)는 주어진 LBA가 저장되는 블록과 그 복사본이 복사되는지(또는 그것이 어떻게 얻어질 수 있는지)를 완벽하게 알 수 있다. 고성능 블록 스토리지와 데이터 분석을 지원하기 위해, 다음의 접근 방법이 사용된다:
- [0141] $P(lba) = \{ \text{parity group of LBA } (lba): \text{ set of SSD and locations where } lba \text{ is stored } \}$
- [0142] Request (IO && data-analysis on lba)
- [0143] {
- [0144] For (p in P (lba)): do IO operation on p: Stream ID (p, lba)
- [0145] For (~p in P(lba)): do data-analysis on ~p: Stream ID(~p, lba)
- [0146] Enqueue (~p, lba) to a lazy-parity queue
- [0147] }
- [0148] $P(lba) = \{ \text{LBA}(lba) \text{의 패리티 그룹: SSD 세트 및 } lba \text{가 저장되는 위치} \}$
- [0149] 요청 (IO && lba에 대한 데이터 분석)
- [0150] {
- [0151] $P(lba)$ 의 p에 대하여: p에 대한 IO 작업: 스트림 ID(p, lba)

- [0152] P(lba)의 ~p에 대하여: ~p에 대한 데이터 분석: 스트림 ID(~p, lba)
- [0153] 레이저-패리티(lazy-parity) 큐 추가(~p, lba)
- [0154] }
- [0155] 쓰기-플러싱 동안 SSD로의 완전한 스트라이프:
- [0156] (~p, lba) = 레이저-패리티 큐의 큐 해제
- [0157] SSD는 lba에 대하여 ~p 완전한 패리티에 기록
- [0158] 패리티 그룹은 다음과 같이 주어진다. SSD의 중복 세트: 호스트 컨트롤러가 볼륨으로부터 주어진 lba를 읽거나 도출할 수 있는 물리적 위치
- [0159] 도 10은 본 발명의 실시예에 따른 각 SSD/스트림 ID 그룹의 볼륨의 인-컴퓨팅(in-computing)을 설명하기 위한 시퀀스 다이어그램이다.
- [0160] 애플리케이션 구성 요소(204)는 복수의 논리적 볼륨으로부터 논리적 볼륨의 BLA에 액세스하기 위한 연산 LBA 요청을 개시한다. 또한, 호스트 컨트롤러(206)는 애플리케이션 구성 요소(204)로부터 연산 요청을 수신하고 (1002), SSD 어레이(300)와 관련된 볼륨의 LBA를 연산한다.
- [0161] 또한, 호스트 컨트롤러(206)는 SSD 어레이(300)와 관련된 논리적 볼륨에 대한 스트림 ID 맵을 검색한다(1004). 호스트 컨트롤러(206)는 SSD 어레이(300)와 관련된 복수의 논리적 볼륨으로부터 논리적 볼륨의 LBA의 패리티 그룹을 결정한다(1006). 호스트 컨트롤러(206)는 또한 조율 로직에 기초하여 결정된 패리티 그룹(즉, SSD 어레이(300)와 관련된 볼륨 그룹)으로부터 LBA에 대한 SSD를 선택한다. 호스트 컨트롤러(206)는 논리적 볼륨의 LBA에 대응하는 스트림 ID 또는 스트림 ID 그룹의 스트림 ID 그룹 상태(즉, 유휴(스트림 ID, 데이터 연산, 파워 모드 OFF))를 NVRAM으로부터 검색한다(1010). 논리적 볼륨의 스트림 ID 그룹 상태는 데이터 분석을 위한 읽기 전용 데이터 연산을 지시한다(1012).
- [0162] SSD 어레이(300)는 SSD 어레이(300)와 관련된 스트림 ID 그룹의 결정된 유휴 상태의 통합된 결과(1016)와 함께 호스트 컨트롤러(206)에 성공적인 응답을 제공한다(1014). 또한, 호스트 컨트롤러(206)는 애플리케이션 구성 요소(204)에 대한 요청을 성공적으로 응답한다(1018).
- [0163] 도 11은 본 발명의 실시예에 따른 SSD 어레이(300)를 관리하기 위한 스트림 ID 기반 스토리지 시스템을 구현하는 컴퓨팅 환경을 도시한다.
- [0164] 도시된 것과 같이 컴퓨팅 환경(1102)은 컨트롤 유닛(1004) 및 산술 논리 유닛(Arithmetic Logic Unit, 1106)을 포함하는 적어도 하나의 프로세싱 유닛(1108), 메모리(1110), 스토리지 유닛(1112), 복수의 네트워킹 장치(1116) 및 복수의 입출력(IO) 장치(1114)를 포함한다. 프로세싱 유닛(1108)은 알고리즘의 명령들을 처리하는 역할을 한다. 프로세싱 유닛(1108)은 처리를 수행하기 위해 컨트롤 유닛(1108)으로부터 명령을 수신한다. 또한, 명령들을 실행과 관련된 임의의 논리 및 산술 연산들은 ALU(1106)의 도움으로 연산된다.
- [0165] 본 명세서에 기재된 실시예들은 적어도 하나의 하드웨어 장치 상에서 실행되는 적어도 하나의 소프트웨어 프로그램을 통해 구현될 수 있고, 구성 요소를 제어하기 위해 네트워크 관리 기능을 수행할 수 있다. 도 1 내지 도 11에 도시된 구성 요소들은 하드웨어 장치와 소프트웨어 모듈의 조합 중 적어도 하나일 수 있는 블록을 포함한다.
- [0166] 이상 첨부된 도면을 참조하여 본 발명의 실시예들을 설명하였으나, 본 발명은 상기 실시예들에 한정되는 것이 아니라 서로 다른 다양한 형태로 제조될 수 있으며, 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자는 본 발명의 기술적 사상이나 필수적인 특징을 변경하지 않고서 다른 구체적인 형태로 실시될 수 있다는 것을 이해할 수 있을 것이다. 그러므로 이상에서 기술한 실시예들은 모든 면에서 예시적인 것이며 한정적이 아닌 것으로 이해해야만 한다.

부호의 설명

- [0167] 100: 스토리지 시스템 200: 호스트 장치
- 202: 프로세서 204: 애플리케이션 구성 요소

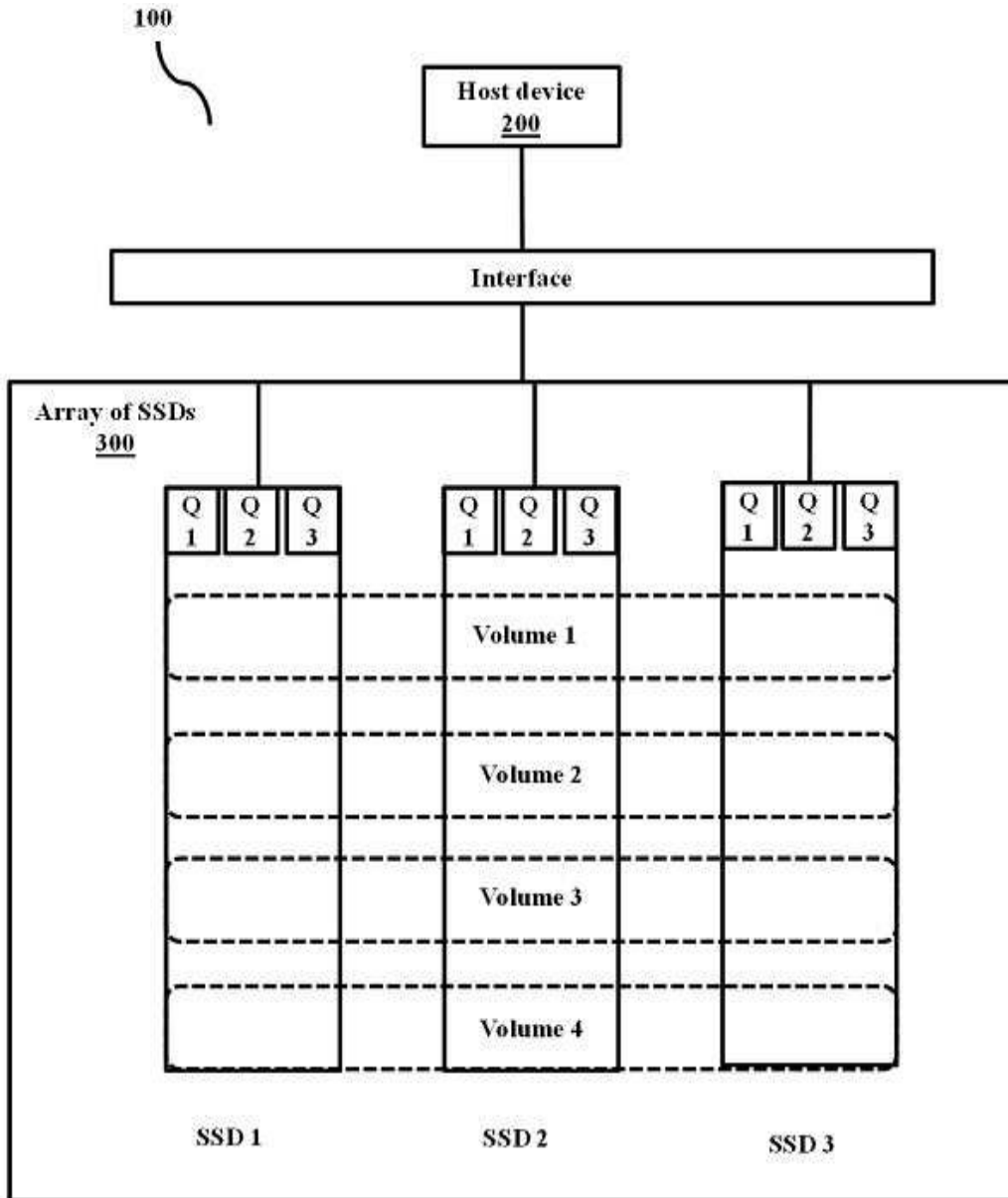
206: 호스트 컨트롤러 208: IO 컨트롤러

210: 메모리 212: 가비지 컬렉터

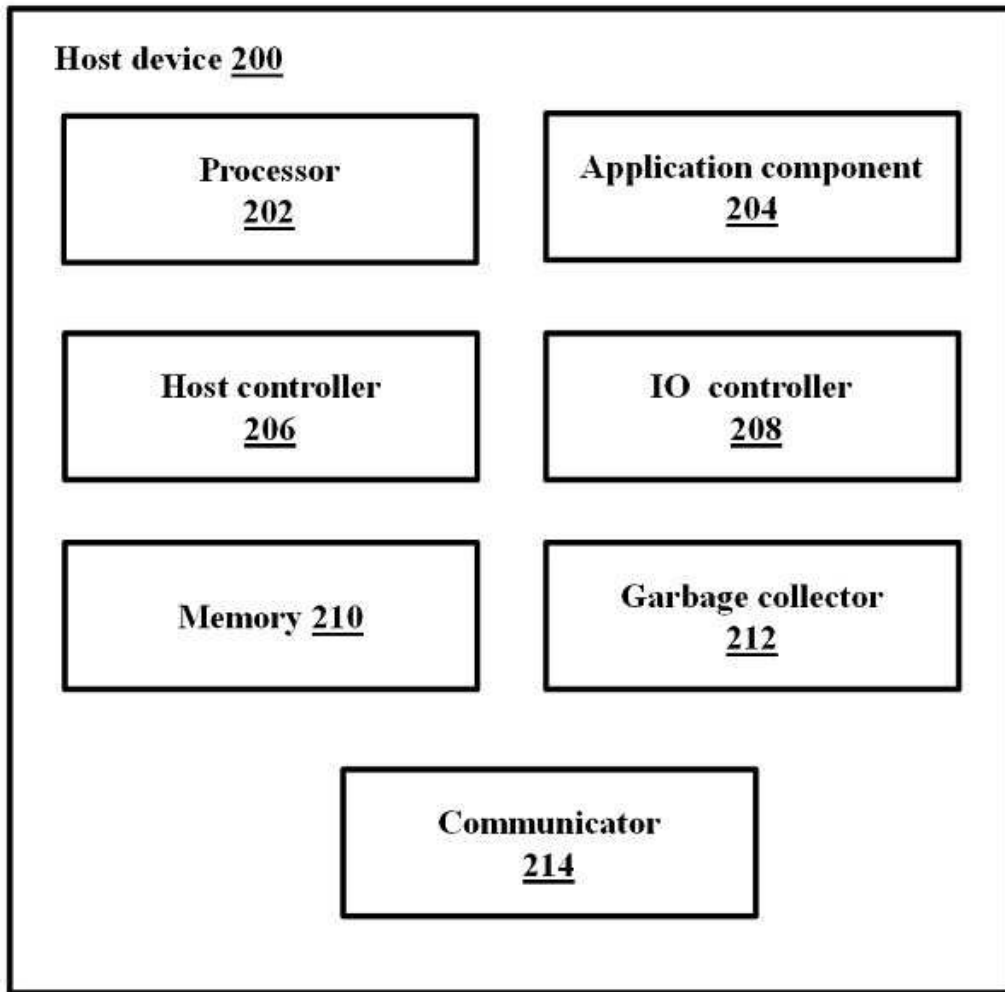
214: 통신기 300: SSD 어레이

도면

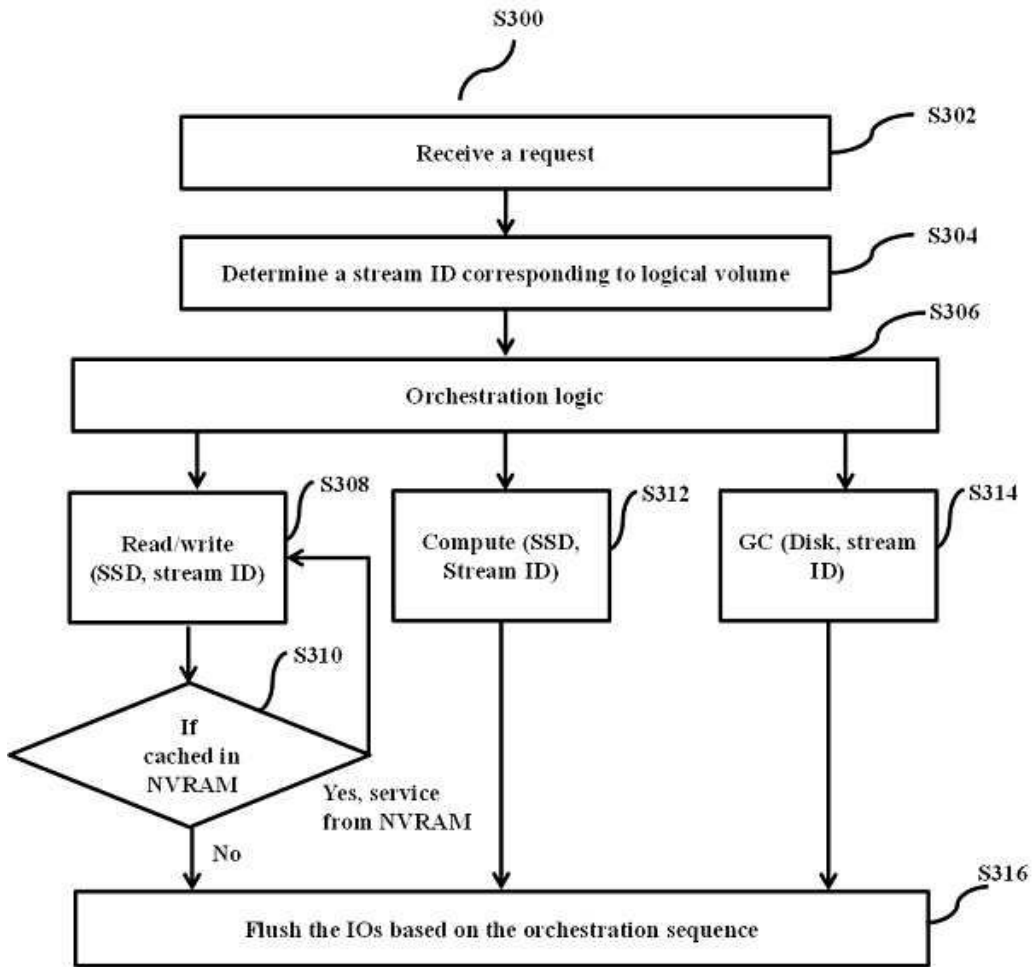
도면1



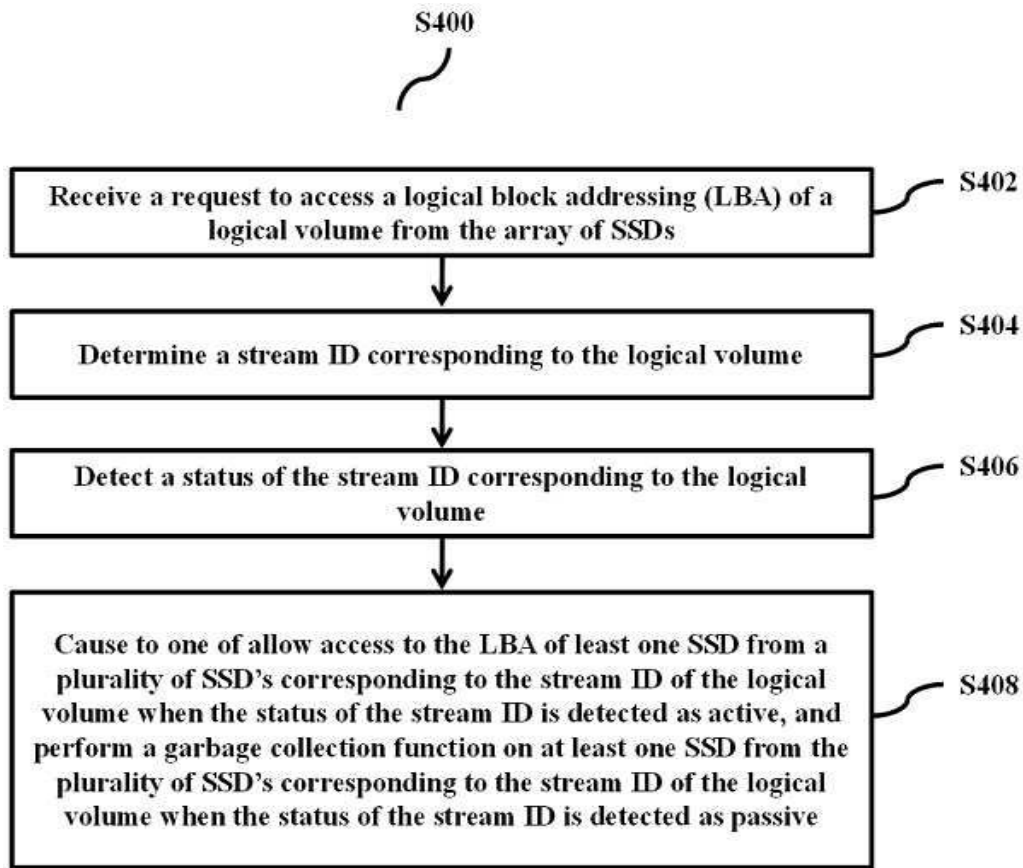
도면2



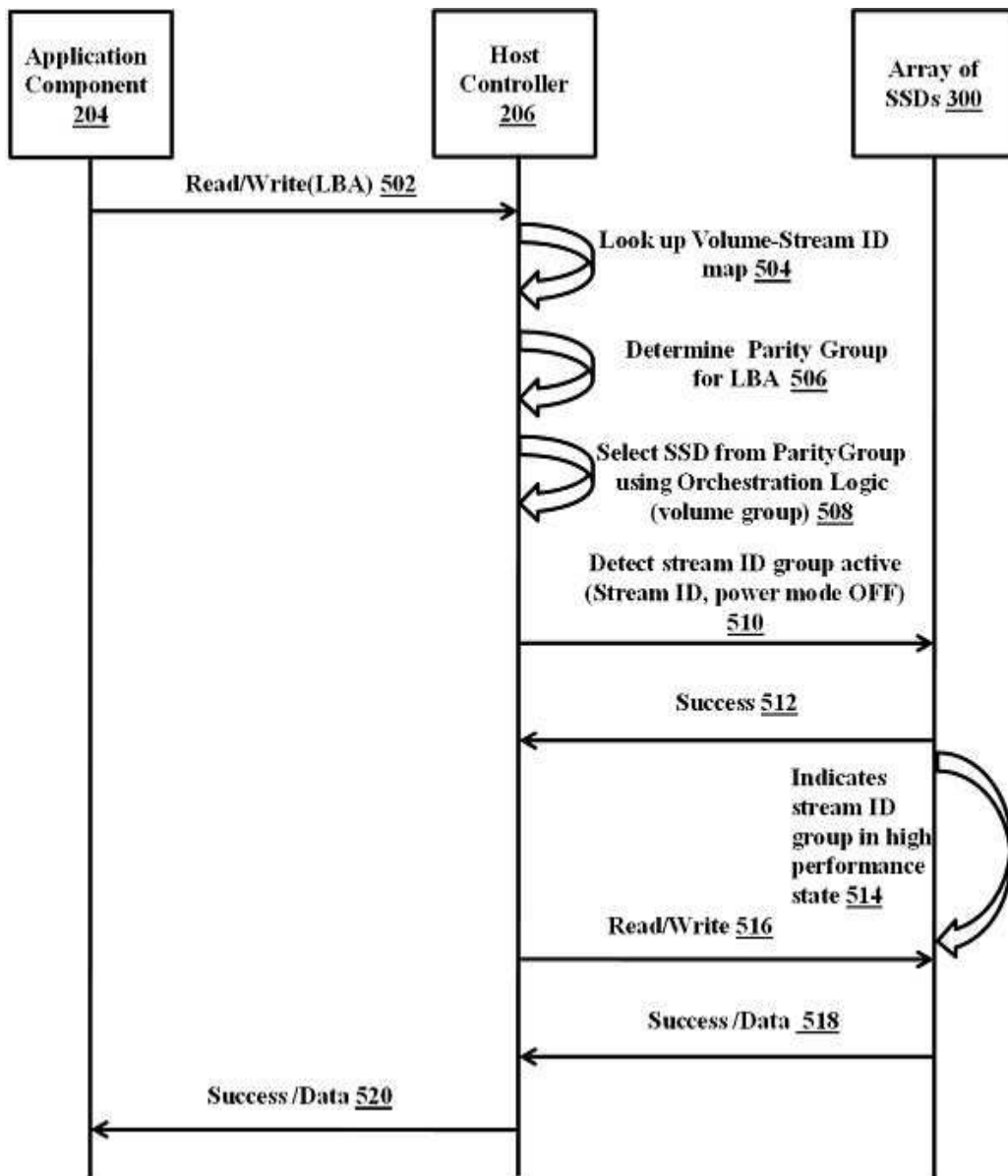
도면3



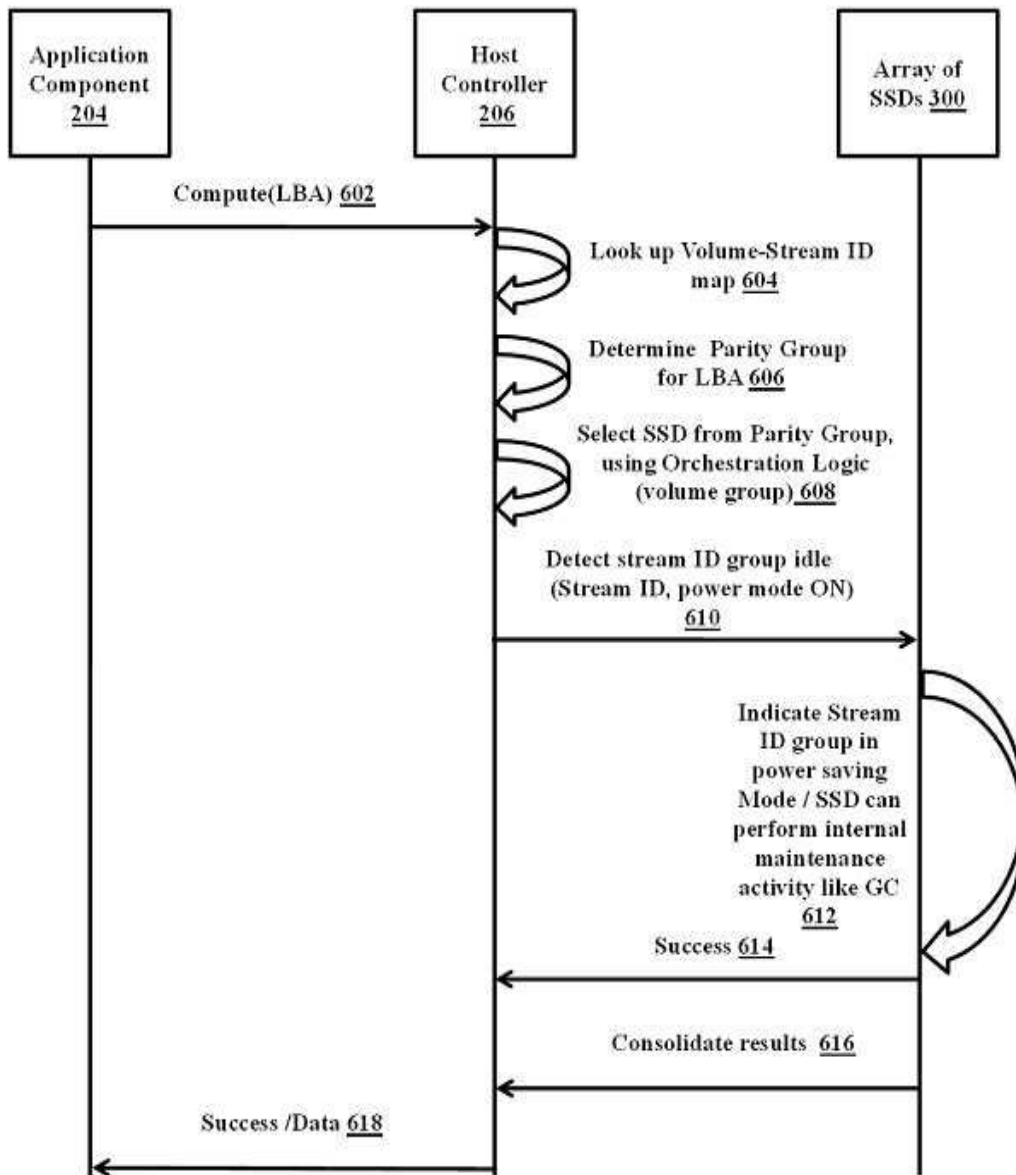
도면4



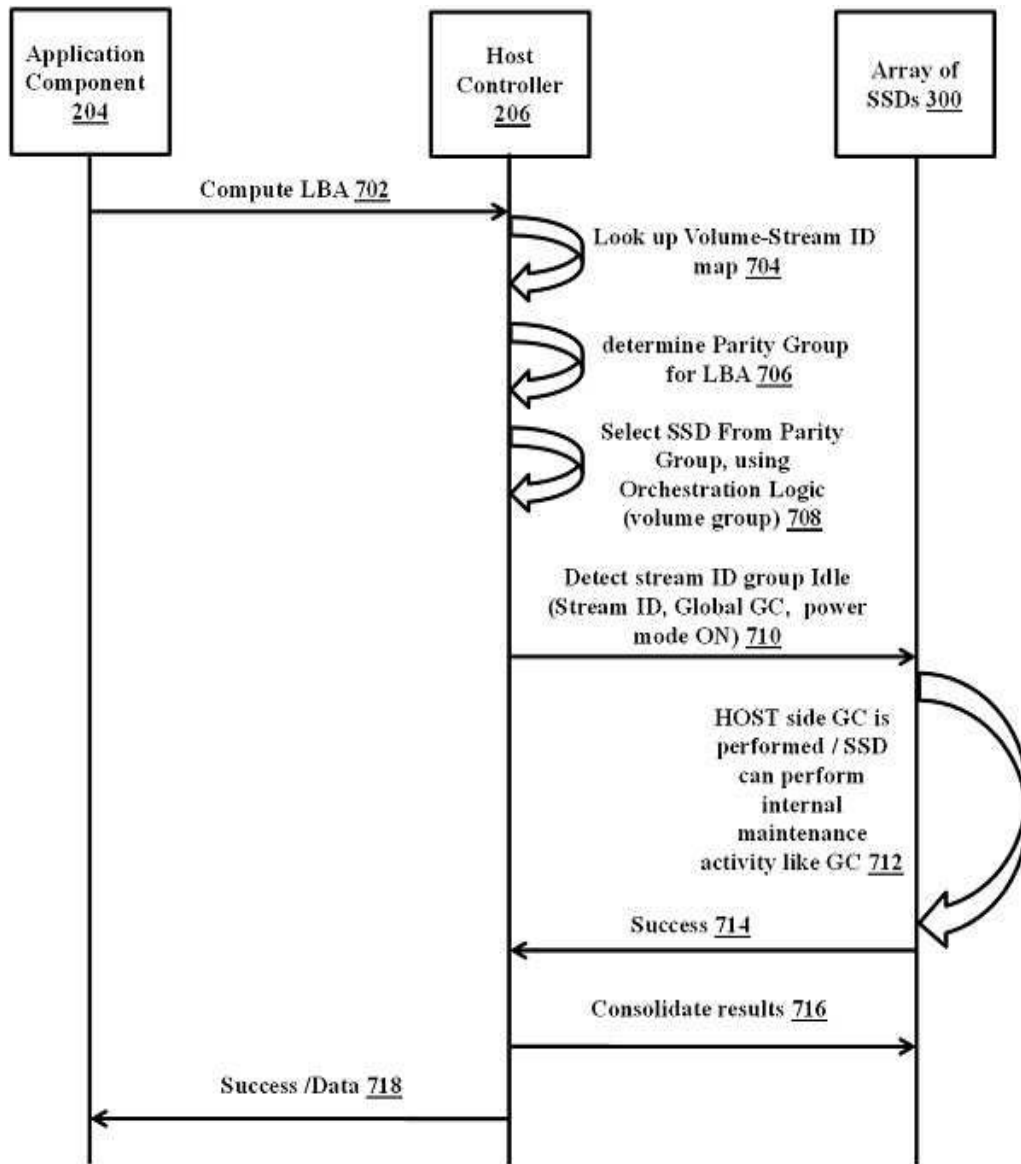
도면5



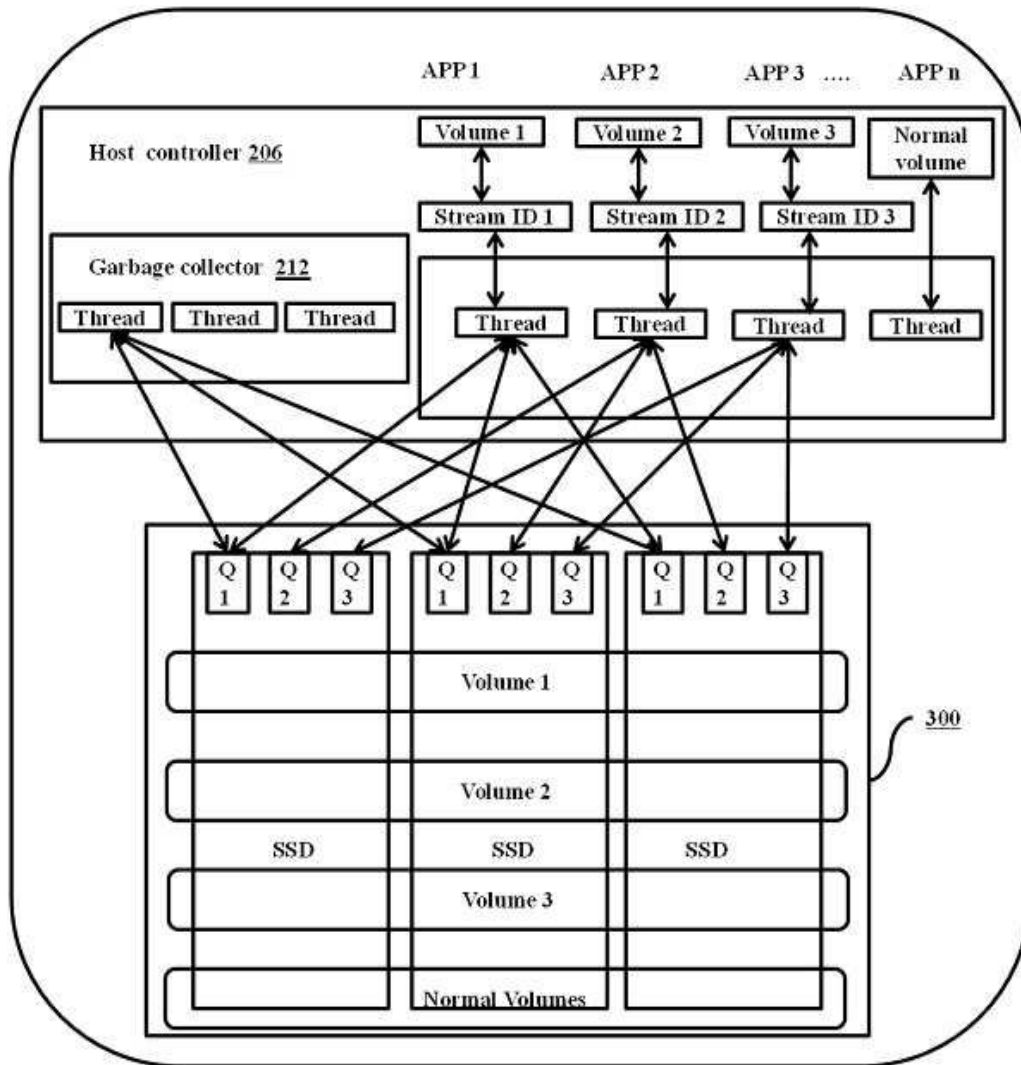
도면6



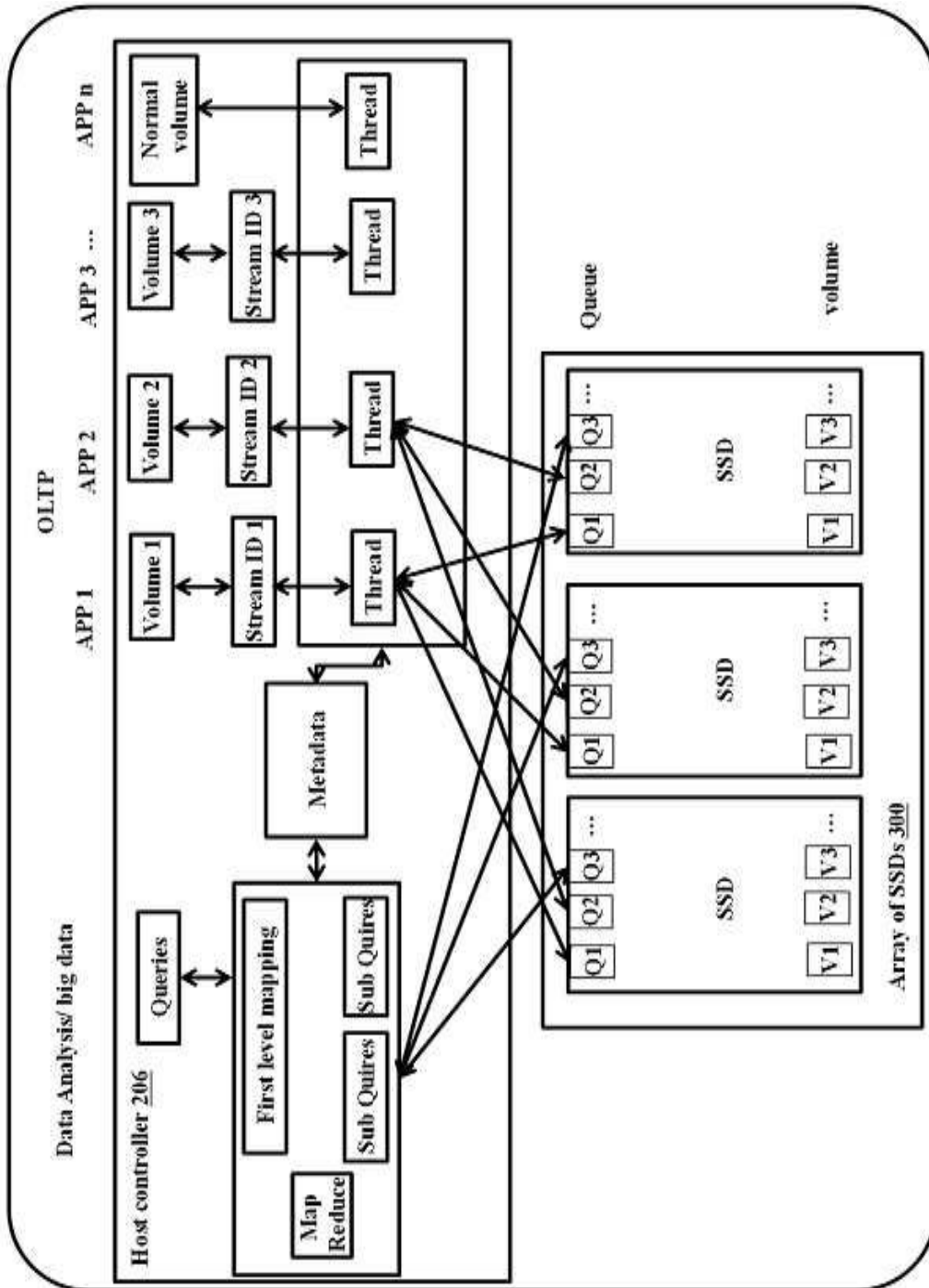
도면7



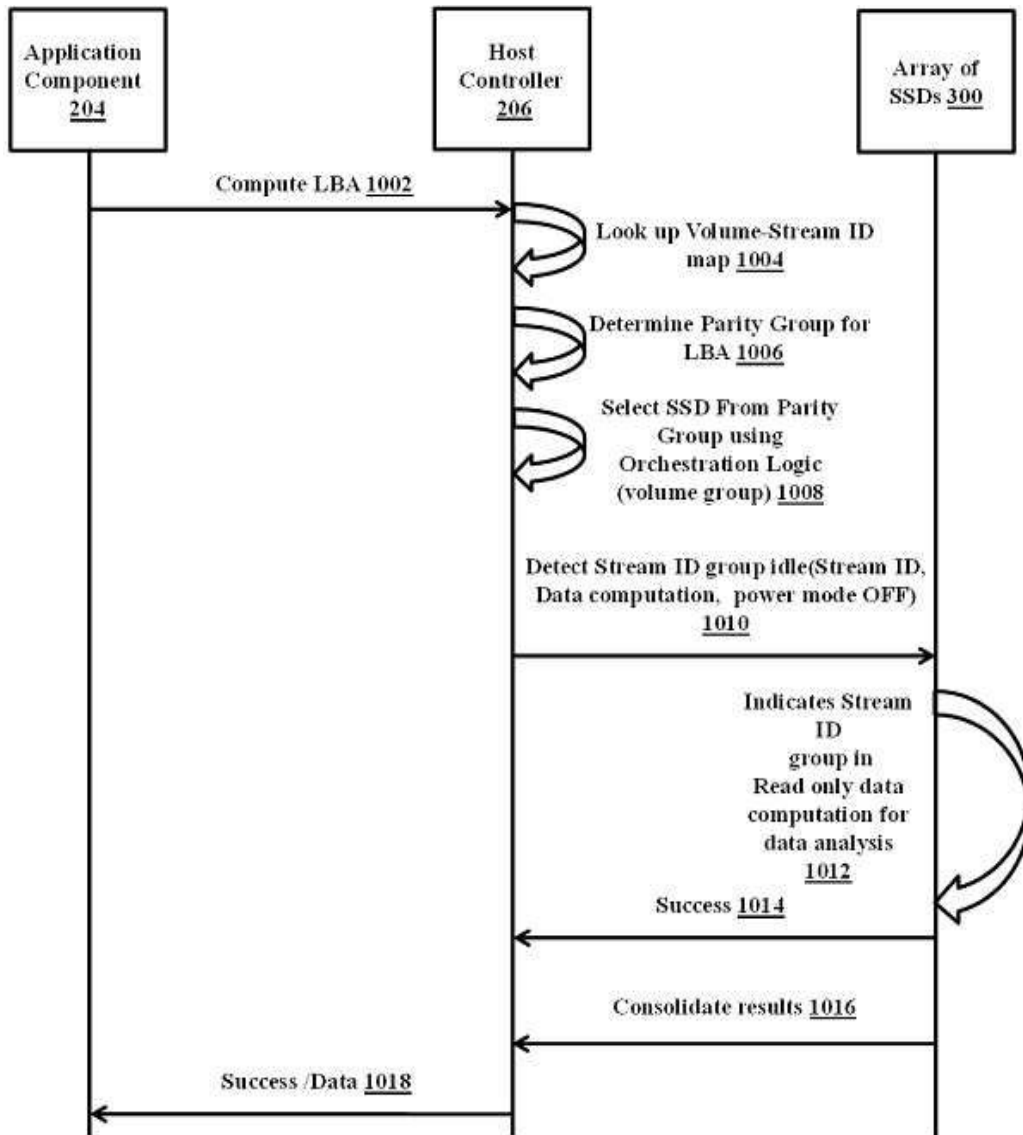
도면8



도면9



도면10



도면11

