



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2012-0094782
(43) 공개일자 2012년08월27일

(51) 국제특허분류(Int. Cl.)

G06F 12/02 (2006.01)

(21) 출원번호 10-2011-0014255

(22) 출원일자 2011년02월17일

심사청구일자 2011년02월17일

(71) 출원인

인하대학교 산학협력단

인천광역시 남구 인하로 100, 인하대학교 (용현동)

(72) 발명자

김덕환

서울특별시 양천구 목동로 230, 103동 1004호 (목동, 목동현대아파트)

(74) 대리인

특허법인 정안

전체 청구항 수 : 총 8 항

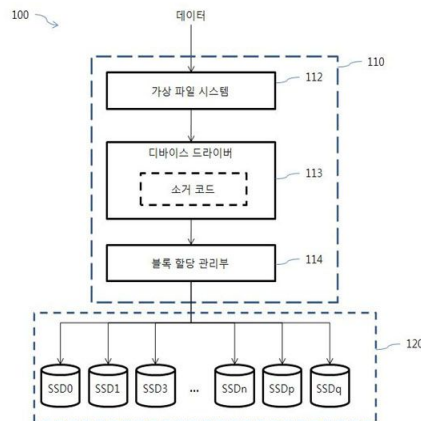
(54) 발명의 명칭 SSD 기반 저장장치 시스템

(57) 요약

본 발명은 SSD 기반 저장장치 시스템에 관한 것으로, 본 발명과 관련된 SSD 기반 저장장치 시스템은 다수의 디스크가 배열로 연결되어 구성되는 디스크 배열과, 상기 다수의 디스크로의 데이터 입출력 요구에 따라 소거 코드를 이용하여 데이터를 인코딩하여 상기 다수의 디스크에 분산하여 저장하는 중앙 제어부를 포함하며, 상기 디스크는 SSD(Solid State Device)와 같은 반도체 저장장치로 구현된다.

또한, 본 발명은 중앙 제어부의 파일 시스템 또는 디바이스 드라이버에 소거 코드를 적용하여 데이터를 인코딩하여 시스템의 입출력 성능을 개선시키며, 패리티 블록을 하나의 패리티 디스크에 집중 저장하여 SSD의 수명저하에 미치는 영향을 최소화하여 신뢰성을 향상시킬 수 있다.

대표도 - 도8



특허청구의 범위

청구항 1

다수의 디스크가 배열 구조로 연결되어 구성되는 디스크 배열과,

상기 다수의 디스크로의 데이터 입출력 요구에 따라 소거 코드를 이용하여 데이터를 인코딩하여 상기 다수의 디스크에 분산하여 저장하는 중앙 제어부를 포함하는 것을 특징으로 하는 SSD 기반 저장장치 시스템.

청구항 2

제1항에 있어서, 상기 디스크는,

SSD(Solid State Drive)와 같은 반도체 저장장치로 구현되는 것을 특징으로 하는 SSD 기반 저장 시스템.

청구항 3

제1항에 있어서, 상기 중앙 제어부는,

상기 디스크의 쓰기/삭제 연산 횟수를 측정하여 그 측정된 연산 횟수가 임계값 이상이면 디스크 교체 알림 메시지를 출력하는 것을 특징으로 하는 SSD 기반 저장 시스템.

청구항 4

제1항에 있어서, 상기 중앙 제어부는,

소거 코드를 이용하여 입력 데이터를 데이터 파일과 패리티 파일로 분할하는 파일 시스템과,

상기 파일 시스템과 연결되며 상기 분할된 파일을 수신하는 가상 파일 시스템과,

상기 가상 파일 시스템을 거쳐 수신되는 분할된 파일을 수신하는 디바이스 드라이버와,

상기 분할된 파일들의 메타데이터 정보에 근거하여 데이터 영역과 패리티 영역으로 분류하여 상기 디스크 배열에 저장하는 블록 할당 관리부를 포함하는 것을 특징으로 하는 SSD 기반 저장 시스템.

청구항 5

제4항에 있어서, 상기 블록 할당 관리부는,

분할된 파일들의 메타데이터 정보에 근거하여 섹터 단위로 데이터 영역과 패리티 영역으로 분류하는 것을 특징으로 하는 SSD 기반 저장 시스템.

청구항 6

제1항에 있어서, 상기 중앙 제어부는,

요청에 따라 데이터를 수신하는 가상 파일 시스템과,

상기 가상 파일 시스템으로부터 전송되는 데이터를 소거 코드를 통해 인코딩하는 디바이스 드라이버와,

상기 인코딩된 데이터를 데이터 영역과 패리티 영역으로 분할하여 상기 디스크 배열에 순차적으로 저장하는 블록 할당 관리부를 포함하는 것을 특징으로 하는 SSD 기반 저장 시스템.

청구항 7

제4항 또는 제6항에 있어서, 상기 블록 할당 관리부는,

갱신 횟수가 잦은 패리티 블록을 하나의 디스크에 저장하는 것을 특징으로 하는 SSD 기반 저장 시스템.

청구항 8

제1항에 있어서, 상기 소거 코드는,

리드-솔몬 코드, EVENODD 코드, 분리 코드(Liberation code) 중 어느 하나인 것을 특징으로 하는 SSD 기반 저

장 시스템.

명세서

기술분야

[0001] 본 발명은 신뢰성 및 입출력 성능 개선을 위한 소거 코드(erasure code)를 적용하는 SSD 기반 저장장치 시스템에 관한 것이다.

배경기술

[0002] 일반적으로, 대용량 서버급 환경에서는 여러 개의 디스크를 배열 구조로 연결한 RAID(Redundant Array of Independent Disks) 시스템과 같은 저장장치 시스템(storage System)이 사용되고 있다. RAID 시스템은 여러 개의 디스크들을 배열 구조로 연결하여 액세스의 병렬성을 높이면서 사용자에게 하나의 디스크로 보이게 하는 기술이다.

[0003] 이러한 RAID 시스템은 다음과 같은 세 가지 작업을 수행할 수 있다. 첫째, 여러 디스크가 동시에 동작해 단일 데이터 흐름을 공급하거나 받아들이게 허용하는 방식으로 다수의 디스크에 존재하는 스트라이핑(striping) 기법으로 성능을 향상시킬 수 있다. 둘째, 다수의 디스크에 존재하는 데이터를 복사하는 미러링(mirroring) 기법이 있어 단일 디스크의 파손과 관련된 위험을 줄일 수 있다. 셋째, 파손된 단일 디스크의 데이터를 복구하기 위한 패리티(parity) 기법이 존재한다. 패리티 기법을 수행하는 경우 적용하는 소거 코드(Erasure code)에 따라 저장장치의 성능이 변경되므로 소거 코드가 저장장치의 성능에 미치는 영향을 고려해야 한다.

[0004] RAID 시스템에 주로 하드디스크를 사용하였으나, 하드디스크는 기계적인 동작이 추가적으로 발생하기 때문에 디스크의 수가 증가할수록 전력소모가 증가하게 될 뿐만 아니라 외부의 충격으로 인해 쉽게 데이터가 손실될 가능성이 높다. 또한, 하드디스크는 소음이 많고 발열이 쉽게 발생하는 단점을 가지고 있다.

[0005] 최근에는 하드디스크와 달리 기계적인 동작이 없고 전력소모가 낮으며 외부의 충격에 강한 반도체 기억장치(Solid State Disk: SSD)가 사용되고 있다. SSD는 다수의 낸드 플래시 메모리(NAND Flash Memory)로 구성되어 있으며, 다음과 같은 두 가지의 특성들을 가진다. 첫째, SSD는 마모도(Wear leveling)라고 불리는 블록 갱신 횟수 제한을 고려한다. 싱글 레벨 셀(Single Level Cell)에서는 약 10만 번 그리고 멀티 레벨 셀(Multi Level Cell)에서는 약 1만 번 정도의 수명을 가지고 있다. SSD는 멀티 레벨 셀로 구성되어 있기 때문에 싱글 레벨 셀 보다 수명의 제한이 낮다. 둘째, 기존의 하드디스크와 달리 SSD는 읽기/쓰기/삭제 3개의 연산의 동작으로 구분할 수 있으며, 읽기/쓰기 연산은 페이지(512B~4KB) 단위로 수행되고 삭제 연산은 블록(16KB~256KB) 단위로 수행된다.

[0006] 상기한 바와 같이 (590)SSD는 고성능 및 저전력의 장점을 가지고 있어 휴대기기에서부터 엔터프라이즈 서버와 같은 고성능 시스템에 이르기까지 하드디스크를 대체할 저장장치로 각광받고 있다.

[0007] 기존의 대용량 저장장치 시스템에서 사용되는 패리티 기반의 하드디스크 RAID 시스템과 SSD의 수명 증가를 위해 연구된 Diff-RAID(Differential RAID) 시스템에 대해 설명한다.

[0008] 하드디스크 기반의 대표적인 패리티 RAID 시스템으로는 도 1과 같이 패리티 정보를 그룹내 모든 디스크들로 분산 저장하는 방식을 이용하는 RAID-5시스템과 2차원 패리티 방식을 사용하는 RAID-6시스템이 있다.

[0009] RAID-5 시스템은 최소 3 개 이상의 디스크가 있어야만 구성이 가능하며, 보통 5 개 이상의 디스크로 구성된다. 또한 디스크에 오류가 발생했을 때에는 패리티(Parity)를 이용해서 데이터를 복구할 수 있기 때문에 패리티 RAID 시스템이라고도 불린다.

[0010] RAID-5 시스템과 같은 저장장치 시스템은 서로 독립적으로 동작하므로, 여러 개의 소규모 어플리케이션 요청 사항을 동시에 처리할 수 있으며, 대용량 입/출력 요청을 처리해야 하는 경우에는 모든 디스크가 활성화 되어 하나의 작업을 처리할 수 있다.

[0011] 그러나, 상기와 같이 하드디스크 기반으로 구성된 RAID 시스템은 하드디스크의 기계적 특성 때문에 소모되는 전력이 많이 필요하며, 디스크의 공간이 협조해지는 경우 성능이 저하되는 문제점을 갖는다.

[0012] RAID-6 시스템은 2 개의 패리티를 사용함으로써 공간 효율은 RAID-5 보다 떨어지지만, 2 개의 디스크가 동시에 고장 나도 데이터에는 이상이 없도록 하는 방식이다. 따라서 RAID-6 시스템은 최소 4 개의 디스크로 도2와 같이 구성된다.

- [0013] 따라서 RAID-6 시스템의 장점으로 RAID-5 시스템 보다 데이터에 대한 신뢰도는 좀 더 높아지는 효과를 갖는다. 하지만 RAID-6 시스템의 단점으로 RAID-5 시스템은 패리티를 1 개만 생성하면 되지만, RAID-6 시스템은 패리티를 2 개를 생성해야 하므로, 내부적으로 소거 코드를 사용하여 인코딩 할 때, 쓰기(Write) 연산이 복잡해 RAID-5 보다 성능이 저하된다.
- [0014] SSD는 삭제 연산이 10,000 번으로 제한되어 있다. 따라서 SSD의 수명 저하로 인한 데이터 손실 및 디스크 오류에 치명적인 단점을 가지게 된다. 최근에는 SSD 기반 RAID 시스템의 연구로 SSD의 수명제한을 고려한 Diff-RAID(Differential RAID) 방식이 연구되고 있다.
- [0015] 도 3은 Diff-RAID시스템의 패리티 분산 방법을 도시한다. 4개의 SSD(SSD0~SSD3) 로 RAID 시스템이 구성 되어 있고, 각각의 SSD로 70 %, 10 %, 10 %, 10 %의 비율로 패리티 분산(Parity Distribution)되어 저장된다. SSD0의 패리티 분산률이 70%로 가장 높기 때문에 SSD0로 쓰기/삭제 연산이 집중된다. 따라서, SSD0가 나머지 SSD(SSD1~SSD3)보다 낮은 수명률을 보이게 된다.
- [0016] 그리고, SSD0의 수명저하로 인해 디스크에 오류가 발생하면 하나의 SSD(SSD4)를 추가하여 기존의 RAID 시스템을 계속 유지한다. SSD1은 SSD0 대신 패리티 분산률이 10%에서 70%로 증가하게 되고 SSD1로 데이터의 쓰기/삭제 연산이 집중되는 시스템을 보여주게 된다.
- [0017] 상기와 같은 Diff-RAID 시스템은 디스크의 수명을 고려한 것으로, SSD 기반의 RAID 시스템의 수명이 효율적으로 관리되지만 소거 코드의 연산이 SSD에 미치는 영향에 대해서는 고려되지 않았다.

발명의 내용

해결하려는 과제

- [0018] 본 발명은 상기와 같은 문제점을 해결하기 위해 안출된 것으로서, 본 발명은 소거 코드를 파일 시스템 레벨이나 디바이스 드라이버 레벨에 적용하는 SSD 기반 저장장치 시스템을 제공하기 위한 것이다.
- [0019] 또한, 본 발명은 갱신 횟수가 잦은 패리티 블록을 하나의 디스크로 집중되게 저장하는 데이터 할당 기법을 적용하여 디스크의 수명 저하를 최소화하는 SSD 기반 저장장치 시스템을 제공하기 위한 것이다.

과제의 해결 수단

- [0020] 상기한 과제를 실현하기 위한 본 발명의 일 예와 관련된 SSD 기반 저장장치 시스템은 다수의 디스크가 배열로 연결되어 구성되는 디스크 배열과, 상기 다수의 디스크로의 데이터 입출력 요구에 따라 소거 코드를 이용하여 데이터를 인코딩하여 상기 다수의 디스크에 분산하여 저장하는 중앙 제어부를 포함한다.
- [0021] 또한, 상기 디스크는 SSD(Solid State Device)와 같은 반도체 저장장치로 구현된다.
- [0022] 또한, 상기 중앙 제어부는 소거 코드를 이용하여 입력 데이터를 데이터 파일과 패리티 파일로 분할하는 파일 시스템과, 상기 파일 시스템과 연결되며 상기 분할된 파일을 수신하는 가상 파일 시스템과, 상기 가상 파일 시스템을 거쳐 수신되는 분할된 파일을 수신하는 디바이스 드라이버와, 상기 분할된 파일들의 메타데이터 정보에 근거하여 데이터 영역과 패리티 영역으로 분류하여 상기 디스크 배열에 저장하는 블록 할당 관리부를 포함한다.
- [0023] 또한, 상기 중앙 제어부는 요청에 따라 데이터를 수신하는 가상 파일 시스템과, 상기 가상 파일 시스템으로부터 전송되는 데이터를 소거 코드를 통해 인코딩하는 디바이스 드라이버와, 상기 인코딩된 데이터를 데이터 영역과 패리티 영역으로 분할하여 상기 디스크 배열에 순차적으로 저장하는 블록 할당 관리부를 포함한다.

발명의 효과

- [0024] 상기와 같이 구성되는 본 발명의 적어도 하나의 실시 예에 관련된 본 발명은 하드디스크를 SSD(solid state disk)로 대체하고 소거 코드를 파일 시스템 레벨 또는 디바이스 드라이버 레벨에 적용하므로, 저장장치 시스템의 입출력 성능을 향상시킬 수 있다.
- [0025] 또한, 본 발명은 패리티 블록을 하나의 디스크로 집중 저장되도록 데이터 할당을 수행하므로, 디스크의 수명 저하를 개선하여 신뢰성을 향상시킬 수 있다.

도면의 간단한 설명

- [0026] 도 1은 일반적인 RAID-5 시스템의 구성.
 도 2는 종래의 하드디스크 기반의 RAID-6 시스템의 구성.
 도 3은 종래의 Differential RAID의 패리티 분산에 따른 수명률을 도시한 도표.
 도 4는 본 발명과 관련된 리드-솔로몬 코드가 적용된 SSD기반의 RAID-6 시스템.
 도 5는 본 발명과 관련된 EVENODD 코드가 적용된 SSD기반의 RAID-6 시스템.
 도 6는 본 발명과 관련된 Liberation 코드가 적용된 SSD기반의 RAID-6 시스템.
 도 7는 본 발명의 일 실시 예와 관련된 SSD 기반 저장장치 시스템의 구조를 도시한 블록도.
 도 8은 본 발명의 다른 실시 예와 관련된 SSD 기반 저장장치 시스템의 구조를 도시한 블록도.

발명을 실시하기 위한 구체적인 내용

- [0027] 이하, 첨부한 도면을 참조하여 본 발명의 실시 예들을 좀 더 상세하게 설명한다.
- [0028] 본 발명은 대용량 저장장치 시스템으로 사용되는 기존의 하드디스크 기반의 RAID(Redundant Array of Independent Disks) 시스템에서 하드디스크를 SSD(solid state device)로 대체한 RAID 시스템을 제안한다. 그리고, 본 발명에서 사용되는 RAID 시스템은 2 개의 패리티(parity)를 사용하는 RAID-6 시스템이고, 이 RAID-6 시스템에 소거 코드(Erasure code)를 적용한다.
- [0029] 소거 코드들은 인코딩 기법에 따라 연산 방법이 달라지기 때문에 SSD에 미치는 영향에 따라 성능 및 전력소모가 달라질 수 있다. 따라서, SSD 환경에 적합한 소거 코드를 사용해야 한다.
- [0030] SSD 환경에 적용 가능한 소거 코드는 다음과 같다.
- [0031] 첫째, 리드-솔로몬(Reed-Solomon) 코드는 가장 강력하고 보편적인 코딩 기법이다. 리드-솔로몬 코딩 기법은 여러 개의 오류를 복잡한 매트릭스 곱으로 복원 할 수 있다. 도 4에 도시된 바와 같이, 분산 매트릭스 B는 매트릭스 I와 하나의 매트릭스로 구성되어 있고 데이터 D와의 곱셈 연산으로 데이터와 코딩 정보를 생성할 수 있다.
- [0032] 하지만 분산 매트릭스 기반의 리드-솔로몬 코드 연산은 계산하는데 많은 비용을 필요로 한다. 따라서 SSD 기반 환경에서는 비트 갱신 횟수가 증가함에 따라 SSD 마모도에 미치는 영향이 높아져 시스템의 전체적인 성능이 저하 될 수 있다.
- [0033] 둘째, EVENODD 코드 기법은 도 5에 도시된 바와 같이 독립적인 두 개의 패리티 열을 가지고 있으며, 두 개의 패리티 열은 각각 데이터들의 체크섬으로 계산 할 수 있다.
- [0034] 첫 번째 패리티($C_{0,0} \sim C_{0,3}$)는 디스크 행렬의 각 행을 모두 XOR 한 값이며 다음과 같은 식으로 표현 할 수 있다. 식 1 를 살펴보면, j는 열을 의미하며 $0 \leq j \leq 3$ 모든 열의 XOR 계산 된다.

[0035]
$$C_{0,j} = a_{0,j} \oplus a_{1,j} \oplus a_{2,j} \oplus a_{3,j} \oplus a_{4,j} \quad (1)$$

- [0036] 두 번째 패리티는 EVENODD 코드의 특수한 대각 패턴과 신드롬(Syndrome)을 함께 XOR 연산하여 계산 된 결과로 식 2 과 같이 표현 할 수 있다.

$$S = a_{4,0} \oplus a_{3,1} \oplus a_{2,2} \oplus a_{2,3}$$

$$C_{1,0} = S \oplus a_{0,0} \oplus a_{2,3} \oplus a_{3,2} \oplus a_{4,1}$$

$$C_{1,1} = S \oplus a_{1,0} \oplus a_{0,1} \oplus a_{3,3} \oplus a_{4,2} \quad (2)$$

$$C_{1,2} = S \oplus a_{2,0} \oplus a_{1,1} \oplus a_{0,2} \oplus a_{4,3}$$

$$C_{1,3} = S \oplus a_{3,0} \oplus a_{2,1} \oplus a_{1,2} \oplus a_{0,3}$$

[0037]

[0038]

위와 같은 계산방법으로 EVENODD 코드가 SSD에 미치는 영향은 단순한 XOR 연산으로 리드-솔로몬 코드보다 연산이 단순하기 때문에 SSD 수명저하에 미치는 영향이 적을 것이다.

[0039]

셋째, Liberation 코드 기법은 리드-솔로몬 코드와 매우 유사하게 계산되며, 리드-솔로몬 코드에서 사용되는 분산 매트릭스와 달리 이진 비트 매트릭스로 분할하여 데이터와 결합하는 코딩 기법을 사용한다.

[0040]

일반적으로 Liberation 코드는 도 6과 같이 k개의 데이터 디바이스와 m개의 코딩 디바이스, 그리고 각각의 w로 고정된 비트로 구성된다. 또한 모든 요소들이 바이너리(binary)로 구성되어 있기 때문에 BDM(Binary Distribution Matrix)이라 불리며, 리드-솔로몬 코딩 보다 빠른 연산이 가능하다.

[0041]

Liberation 코드는 리드-솔로몬 코드와 연산 방법이 매우 유사하지만, 이진 매트릭스를 사용하여 계산되기 때문에 계산 복잡도가 낮으며 빠른 연산을 할 수 있다. 따라서 SSD 기반 환경에 적용 하였을 때 리드-솔로몬 보다 낮은 비트 갱신 횟수로 인해 SSD 마모도에 미치는 영향과 SSD 수명저하에 미치는 영향이 적다는 특징을 가지고 있다.

[0042]

도 7는 본 발명의 일 실시 예와 관련된 SSD 기반 저장장치 시스템의 구조를 도시한 블록도이다. 본 실시 예는 소거 코드(Erasure code)를 파일 시스템에 적용한 경우 쓰기 동작을 수행하는 것을 예로 들어 설명한다.

[0043]

SSD 기반 저장장치 시스템(100)은 중앙 제어부(110)와 디스크 배열(disk array)(120)를 포함한다.

[0044]

중앙 제어부(110)는 저장장치 시스템에서 외부입력에 의해 시스템 호출이 요청되거나 하드웨어로부터 인터럽트가 요청되면 그 요청에 대응하는 처리를 수행한다. 그리고, 중앙 제어부(110)는 디스크 배열(120)을 구성하는 다수의 디스크들로 데이터를 입출력하는 경우 소거 코드를 이용하여 데이터를 인코딩(encoding)하거나 디코딩(decoding)한다.

[0045]

중앙 제어부(110)는 파일 시스템(111)과, 가상 파일 시스템(Virtual File System: VFS)(112)와 디바이스 드라이버(device driver)(113)과, 블록 할당 관리부(114)를 포함한다.

[0046]

파일 시스템(111)은 파일의 생성 및 접근 제어, 디렉토리 관리 등을 서비스한다. 파일 시스템(111)은 쓰기 요청이 수신되면 하나의 데이터를 소거 코드를 이용하여 데이터 파일(파일0~파일3)과 패리티(parity) 파일(파일p, 파일q)로 분할한다. 다시 말해서, 파일 시스템(111)은 소거 코드를 통해 데이터를 인코딩하여 파일 단위로 생성한다.

[0047]

가상 파일 시스템(112)은 사용자로부터 특정 이벤트가 들어오면 파일시스템에 대한 인터페이스를 제공하는 역할을 한다. 가상 파일 시스템(112)는 상기 분할된 파일들을 디바이스 드라이버(113)로 전송한다.

[0048]

디바이스 드라이버(113)은 디스크, 터미널, 네트워크 카드 등의 주변 장치를 구동하는 드라이버들로 구성된다. 디바이스 드라이버(113)은 디스크 배열(120)로 데이터 입출력을 제어하며, 상기 가상 파일 시스템(112)을 거쳐 수신되는 상기 분할된 파일들을 블록 할당 관리부(114)로 전송한다.

[0049]

상기 블록 할당 관리부(114)는 내부에 데이터 영역(data area)과 패리티 영역(parity area)을 구비한다. 상기 블록 할당 관리부(114)는 상기 분할된 파일들의 메타데이터 정보에 근거하여 섹터(sector) 단위(512B~4KB)로 데이터 영역과 패리티 영역을 분류하여 디스크 배열(120)로 전송한다. 다시 말해서, 블록 할당 관리부(114)는 상기 파일 시스템(111)에 의해 인코딩된 데이터의 메타데이터 정보를 가지고 데이터를 분류한다. 상기 디스크 배열(120)로 전송된 섹터들은 디스크 배열(120)의 각 디스크에 순차적으로 페이지(page) 단위(512B~4KB)로 저장된다.

- [0050] 도 8은 본 발명의 다른 실시 예와 관련된 SSD 기반 저장장치 시스템의 구조를 도시한 블록도이다. 본 실시 예는 소거 코드를 디바이스 드라이브에 적용하여 쓰기 동작을 수행하는 것을 예로 들어 설명한다.
- [0051] 도 8을 참조하면, 가상 파일 시스템(112)은 쓰기 요청이 있으면 데이터를 수신하면 디바이스 드라이버(113)로 전송한다.
- [0052] 디바이스 드라이버(113)는 소거 코드를 사용하여 가상 파일 시스템(112)으로부터 전송되는 데이터를 섹터 단위로 인코딩한다. 여기서, 소거 코드는 리드-솔로몬(Reed-Solomon) 코드, EVENODD 코드, Liberation 코드 중 어느 하나이다. 그리고, 디바이스 드라이버(113)는 상기 인코딩된 데이터를 블록 할당 관리부(114)로 전송한다.
- [0053] 블록 할당 관리부(114)는 디바이스 드라이버(113)로부터 전송되는 인코딩된 데이터를 데이터 영역과 패리티 영역으로 분할하여 디스크 배열(120)의 디스크들에 순차적으로 저장한다.
- [0054] 상기 블록 할당 관리부(114)는 인코딩된 데이터의 메타데이터 정보에 의해 데이터를 분류할 수 있다. 예를 들어, 인코딩된 데이터의 메타데이터 정보가 데이터 블록이면 블록 할당 관리부(114)는 인코딩된 데이터를 데이터 블록 할당 관리부(114) 내의 데이터 영역으로 분류하여 디스크 배열(120)의 데이터 디스크(SSD0~SSDn)에 저장한다. 그리고, 인코딩된 데이터의 메타데이터 정보가 패리티블록_p이거나 또는 패리티 블록_q이면, 블록 할당 관리부(114)는 인코딩된 데이터를 블록 할당 관리부(114) 내의 패리티 영역으로 분류하여 디스크 배열(120)의 패리티 디스크(SSDp, SSDq)에 저장한다.
- [0055] 또한, 블록 할당 관리부(114)는 갱신 횟수가 잦은 패리티 블록을 하나의 디스크로 집중하여 저장한다. 예를 들어, 4개의 데이터 디스크와 2개의 패리티 디스크로 디스크 배열(120)이 구성된 경우, 블록 할당 관리부(114)는 데이터를 4개의 데이터 디스크로 분산하여 저장하고, 패리티p와 패리티q는 2개의 패리티 디스크 중 하나의 디스크에 집중되게 저장한다. 따라서, 2개의 패리티 디스크 중 다른 하나의 디스크 수명저하를 최소화할 수 있다.
- [0056] 디스크 배열(120)은 적어도 둘 이상의 데이터 디스크들과 두 개의 패리티 디스크들이 배열로 연결된다. 여기서, 디스크는 SSD(solid state disk)와 같은 반도체 저장장치로 구현된다. 각 디스크는 쓰기/삭제 연산 횟수를 측정하며, 그 측정된 쓰기/삭제 연산 횟수가 임계값 이상이면 디스크 교체가 필요함을 중앙 제어부(110)에 알린다. 상기 중앙 제어부(110)은 디스크의 교체 알림 메시지를 출력한다. 예를 들어, SSD 수명이 10,000회로 제한되므로, SSD의 쓰기/삭제 연산 횟수가 9,000회가 되면 이를 감지하여 중앙 제어부(110)는 사용자가 인지할 수 있도록 디스크 교체 알림을 출력한다. 상기 디스크 교체 알림은 메시지 및/또는 경고음 등으로 출력된다.
- [0057] 상기 디스크 배열(130)은 제어 명령에 따라 디스크 배열(120)을 구성하는 디스크들 중 적어도 하나 이상의 디스크 동작을 제어하는 디스크 제어부를 더 포함할 수 있다. 디스크 제어부는 각 디스크들의 상태를 감지하여 그 연결 상태를 파악하고, 상기 파악된 디스크 상태 정보를 중앙 제어부(110)로 전달한다. 상기 디스크 제어부는 각 디스크의 쓰기/삭제 연산 횟수를 측정하여 그 측정된 쓰기/삭제 연산 횟수가 임계값 이상이면 중앙 제어부(110)에 디스크의 교체 필요 여부를 알리는 정보를 전송한다.

부호의 설명

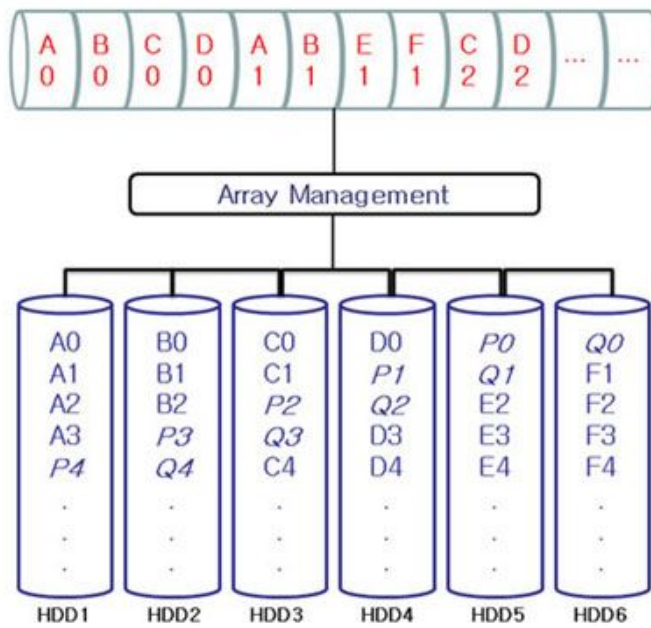
- [0058]
- | | |
|----------------|----------------|
| 100: 저장장치 시스템 | 110: 중앙 제어부 |
| 111: 파일 시스템 | 112: 가상 파일 시스템 |
| 113: 디바이스 드라이버 | 114: 블록 할당 관리부 |
| 120: 디스크 배열 | |

도면

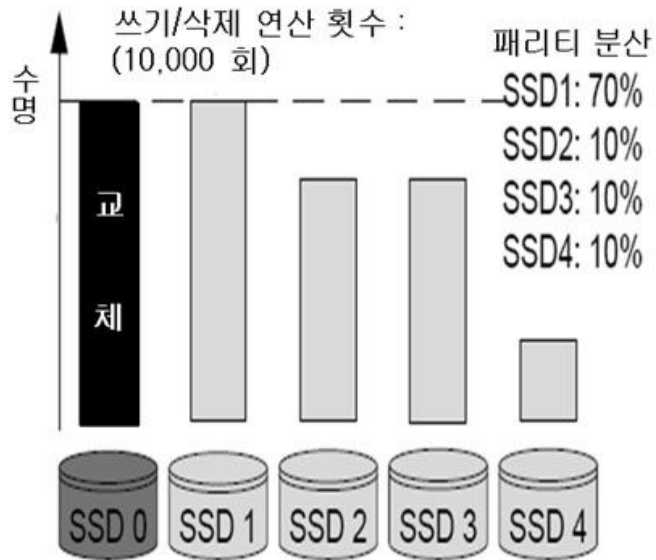
도면1

	1	2	3	4	5	디스크 번호
섹터						
1	D ₁	D ₂	D ₃	D ₄	P	
2	D ₅	D ₆	D ₇	P	D ₈	
3	D ₉	D ₁₀	P	D ₁₁	D ₁₂	
4	D ₁₃	P	D ₁₄	D ₁₅	D ₁₆	
5	P	D ₁₇	D ₁₈	D ₁₉	D ₂₀	

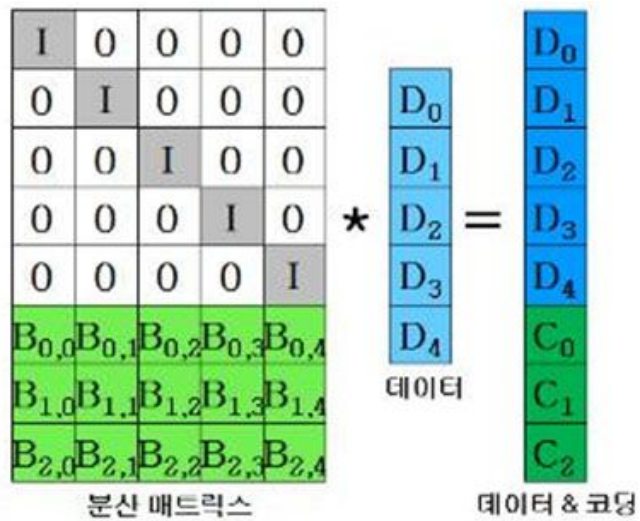
도면2



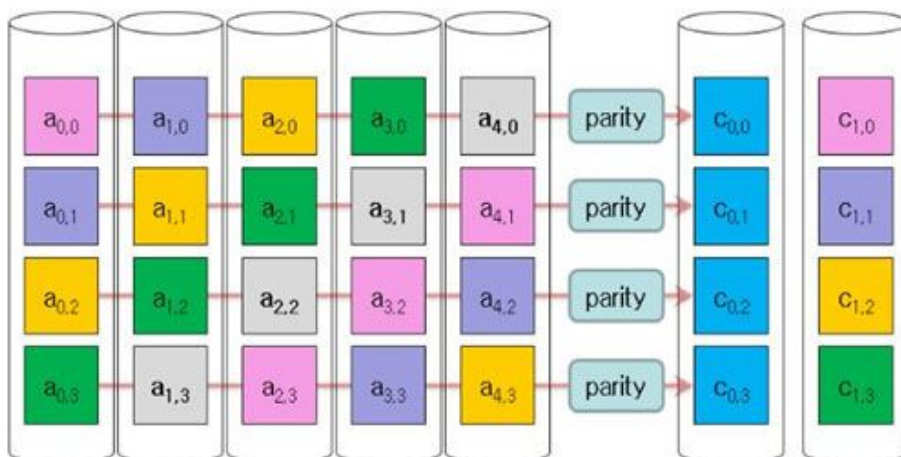
도면3



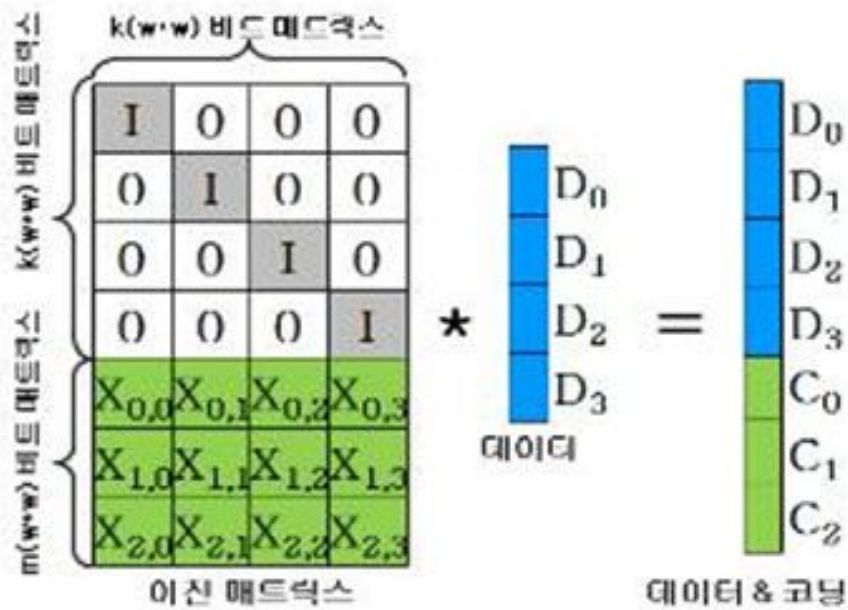
도면4



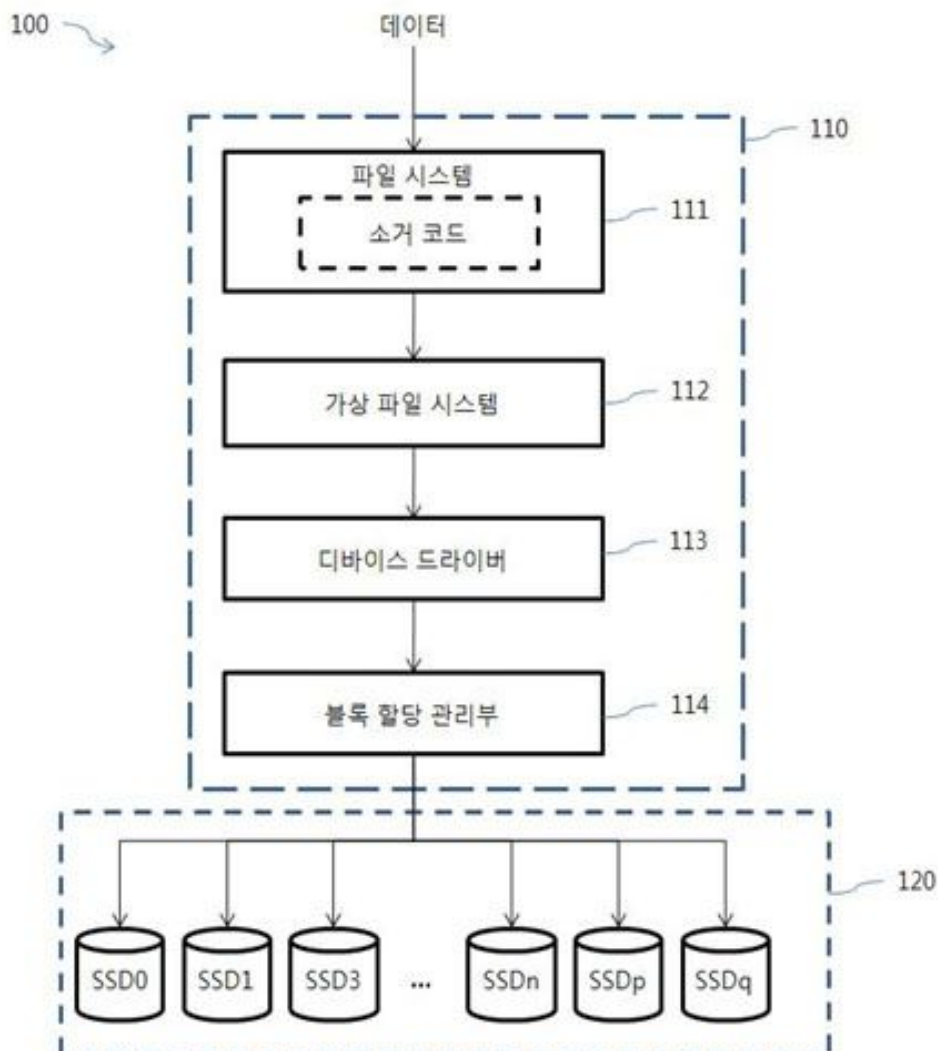
도면5



도면6



도면7



도면8

