



(19) 대한민국특허청(KR)  
(12) 등록특허공보(B1)

(45) 공고일자 2022년05월27일  
(11) 등록번호 10-2403266  
(24) 등록일자 2022년05월24일

(51) 국제특허분류(Int. Cl.)  
G06F 12/02 (2018.01) G06F 13/16 (2006.01)  
G06F 3/06 (2006.01)  
(52) CPC특허분류  
G06F 12/0292 (2013.01)  
G06F 13/1668 (2013.01)  
(21) 출원번호 10-2015-0088469  
(22) 출원일자 2015년06월22일  
심사청구일자 2020년04월29일  
(65) 공개번호 10-2016-0150451  
(43) 공개일자 2016년12월30일  
(56) 선행기술조사문헌  
US08601202 B  
(뒷면에 계속)

(73) 특허권자  
삼성전자주식회사  
경기도 수원시 영통구 삼성로 129 (매탄동)  
(72) 발명자  
강남옥  
경기도 화성시 동탄반석로 160, A동 513호 (반송동)  
(74) 대리인  
특허법인 고려

전체 청구항 수 : 총 20 항

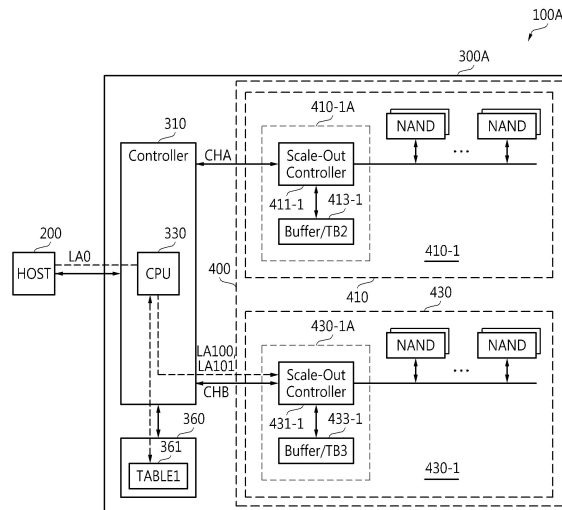
심사관 : 남윤권

(54) 발명의 명칭 데이터 저장 장치와 이를 포함하는 데이터 처리 시스템

(57) 요약

데이터 저장 장치가 공개된다. 상기 데이터 저장 장치는 제1휘발성 메모리, 제1불휘발성 메모리, 및 상기 제1휘발성 메모리와 상기 제1불휘발성 메모리를 제어하는 제1코어를 포함하는 제1멀티 칩 셋과, 제2휘발성 메모리, 제2불휘발성 메모리, 및 상기 제2휘발성 메모리와 상기 제2불휘발성 메모리를 제어하는 제2코어를 포함하는 제2멀티 칩 셋과, 상기 제1멀티 칩 셋과 상기 제2멀티 칩 셋과 접속되고, 상기 제1불휘발성 메모리의 제1저장 영역에 대한 제1논리 어드레스와 상기 제2불휘발성 메모리의 제2저장 영역에 대한 제2논리 어드레스를 스왑하는 컨트롤러를 포함한다.

대표도 - 도8b



(52) CPC특허분류

*G06F 3/0616* (2013.01)

*G06F 3/0685* (2013.01)

(56) 선행기술조사문헌

US20140143474 A1

US20150095546 A1

US20150149728 A1

US20100023800 A1

US20090157947 A1

---

## 명세서

### 청구범위

#### 청구항 1

제1휘발성 메모리, 제1불휘발성 메모리, 및 상기 제1휘발성 메모리와 상기 제1불휘발성 메모리를 제어하고, 상기 제1불휘발성 메모리에 포함된 저장 영역들 각각의 웨어 카운트와 기준 웨어-아웃 카운트를 비교하는 제1코어를 포함하는 제1멀티 칩 셋;

제2휘발성 메모리, 제2불휘발성 메모리, 및 상기 제2휘발성 메모리와 상기 제2불휘발성 메모리를 제어하는 제2코어를 포함하는 제2멀티 칩 셋;

제3휘발성 메모리, 제3불휘발성 메모리, 및 상기 제3휘발성 메모리와 상기 제3불휘발성 메모리를 제어하는 제3코어를 포함하는 제3멀티 칩 셋;

상기 제1멀티 칩 셋과 상기 제2멀티 칩 셋과 연결되고, 상기 비교 결과에 기반하여 상기 제1불휘발성 메모리에 포함된 제1저장 영역이 상기 기준 웨어-아웃 카운트보다 높은 웨어 카운트를 가지는 경우, 상기 제1불휘발성 메모리에 포함된 상기 제1저장 영역에 대한 제1논리 어드레스와 상기 제2불휘발성 메모리에 포함된 제2저장 영역에 대한 제2논리 어드레스를 스왑(swap)하는 컨트롤러; 및

상기 컨트롤러에 의하여 서로 스왑된 상기 제1논리 어드레스와 상기 제2논리 어드레스를 포함하는 테이블을 저장하는 메모리를 포함하되,

상기 컨트롤러는 호스트로부터 상기 제1논리 어드레스를 수신하고, 상기 테이블을 참조하여 상기 스왑된 제2논리 어드레스를 상기 제2멀티 칩 셋으로 전송하고,

상기 제1불휘발성 메모리 및 상기 제3불휘발성 메모리는 사용자 데이터를 저장하고, 상기 제2불휘발성 메모리는 오버프로비저닝(overprovisioning) 영역인 데이터 저장 장치.

#### 청구항 2

제1항에 있어서,

상기 컨트롤러는 상기 제1저장 영역에 저장된 제1데이터와 상기 제2저장 영역에 저장된 제2데이터를 스왑하는 데이터 저장 장치.

#### 청구항 3

제1항에 있어서,

상기 제1저장 영역의 웨어 카운트는 상기 제2저장 영역의 웨어 카운트보다 높은 데이터 저장 장치.

#### 청구항 4

제3항에 있어서,

상기 웨어 카운트는 이레이즈 카운트, 프로그램 카운트, 리드 카운트, 배드 블록 발생 횟수, 및 리드 에러 발생 횟수 중 적어도 하나에 기반하는 데이터 저장 장치.

#### 청구항 5

제1항에 있어서,

상기 제1저장 영역은 핫 데이터 및 콜드 데이터 중에서 어느 하나를 저장하기 위한 영역이고,

상기 제2저장 영역은 상기 핫 데이터 및 상기 콜드 데이터 중에서 다른 하나를 저장하는 영역인 데이터 저장 장치.

#### 청구항 6

제1항에 있어서,

상기 제1휘발성 메모리 및 상기 제1코어는 제1반도체 패키지에 포함되고, 상기 제2휘발성 메모리 및 상기 제2코어는 제2반도체 패키지에 포함되는 데이터 저장 장치.

#### 청구항 7

제1항에 있어서,

상기 컨트롤러는 상기 컨트롤러의 제1요청에 따라 상기 제1코어로부터 제1채널을 통하여 제1논리 어드레스를 수신하고, 상기 컨트롤러의 제2요청에 따라 상기 제2코어로부터 제2채널을 통하여 제2논리 어드레스를 수신하는 데이터 저장 장치.

#### 청구항 8

제1항에 있어서,

상기 제1휘발성 메모리와 상기 제2휘발성 메모리 각각은 DRAM이고,  
상기 제1불휘발성 메모리와 상기 제2불휘발성 메모리 각각은 플래시 메모리이고,  
상기 데이터 저장 장치는 솔리드 스테이트 드라이브(SSD)인 데이터 저장 장치.

#### 청구항 9

제1항에 있어서,

상기 컨트롤러는 상기 비교 결과에 기반하여 상기 제3불휘발성 메모리의 제3저장 영역이 상기 기준 웨어-아웃 카운트보다 높은 웨어 카운트를 가지는 경우, 상기 제3불휘발성 메모리의 상기 제3저장 영역의 제3논리 어드레스와 상기 제2불휘발성 메모리의 상기 제2저장 영역의 상기 제2논리 어드레스를 스왑하는 데이터 저장 장치.

#### 청구항 10

제1휘발성 메모리, 제1불휘발성 메모리, 및 상기 제1휘발성 메모리와 상기 제1불휘발성 메모리를 제어하는 제1코어를 포함하는 제1멀티 칩 셋 및 제2휘발성 메모리, 제2불휘발성 메모리, 및 상기 제2휘발성 메모리와 상기 제2불휘발성 메모리를 제어하는 제2코어를 포함하는 제2멀티 칩 셋을 포함하는 데이터 저장 장치의 동작 방법에 있어서:

호스트로부터 제1논리 어드레스 및 데이터를 가지는 프로그램 요청을 수신하는 단계;

상기 제1코어에 의하여, 상기 제1불휘발성 메모리에 포함된 저장 영역들 각각의 웨어 카운트와 기준 웨어-아웃 카운트를 비교하고, 상기 저장 영역들 각각의 상기 웨어 카운트가 상기 기준 웨어-아웃 카운트보다 높은지 여부를 결정하는 단계; 및

컨트롤러에 의하여, 상기 비교 결과에 기반하여 상기 제1불휘발성 메모리에 포함된 제1저장 영역이 상기 기준 웨어-아웃 카운트보다 높은 웨어 카운트를 가지는 경우, 상기 제1불휘발성 메모리에 포함된 상기 제1저장 영역의 제1로직 어드레스와 상기 제2불휘발성 메모리에 포함된 제2저장 영역의 제2로직 어드레스를 스왑(swap)하는 단계를 포함하는 데이터 저장 장치의 동작 방법.

#### 청구항 11

제10항에 있어서,

상기 컨트롤러는 상기 제2불휘발성 메모리에 포함된 상기 제2저장 영역에 데이터를 제공하는 데이터 저장 장치의 동작 방법.

#### 청구항 12

제10항에 있어서,

상기 컨트롤러에 의하여 스왑된 상기 제1로직 어드레스 및 상기 제2로직 어드레스를 포함하는 테이블을 메모리에 저장하는 단계를 더 포함하는 데이터 저장 장치의 동작 방법.

### 청구항 13

제12항에 있어서,

상기 컨트롤러는 상기 테이블을 참조하여 스왑된 상기 제2로직 어드레스를 상기 제 2 멀티 칩 셋으로 전송하는 데이터 저장 장치의 동작 방법.

### 청구항 14

제10항에 있어서,

상기 웨어 카운트는 이레이즈 카운트, 프로그램 카운트, 리드 카운트, 배드 블록 발생 횟수, 및 리드 에러 발생 횟수 중 적어도 하나에 기반하는 데이터 저장 장치의 동작 방법.

### 청구항 15

제10항에 있어서,

상기 제1저장 영역은 핫 데이터 및 콜드 데이터 중에서 어느 하나를 저장하기 위한 영역이고,

상기 제2저장 영역은 상기 핫 데이터 및 상기 콜드 데이터 중에서 다른 하나를 저장하는 영역이고,

상기 핫 데이터는 상기 콜드 데이터에 비하여 더 빈번하게 액세스되는 데이터 저장 장치의 동작 방법.

### 청구항 16

제10항에 있어서,

상기 제1휘발성 메모리 및 상기 제1코어는 제1반도체 패키지에 포함되고,

상기 제2휘발성 메모리 및 상기 제2코어는 제2반도체 패키지에 포함되는 데이터 저장 장치의 동작 방법.

### 청구항 17

제16항에 있어서,

상기 제1반도체 패키지 및 제2반도체 패키지 각각은 멀티-칩 패키지(MCP)인 데이터 저장 장치의 동작 방법.

### 청구항 18

제10항에 있어서,

상기 제1멀티 칩 셋 및 상기 컨트롤러는 하나의 반도체 패키지로 구현되는 데이터 저장 장치의 동작 방법.

### 청구항 19

제18항에 있어서,

상기 반도체 패키지는 임베디드 패키지 온 패키지(embedded Package On Package)인 데이터 저장 장치의 동작 방법.

### 청구항 20

제1휘발성 메모리, 제1불휘발성 메모리, 및 상기 제1휘발성 메모리와 상기 제1불휘발성 메모리를 제어하고, 상기 제1불휘발성 메모리에 포함된 저장 영역들 각각의 웨어 카운트와 기준 웨어-아웃 카운트를 비교하는 제1코어를 포함하는 제1멀티 칩 셋;

제2휘발성 메모리, 제2불휘발성 메모리, 및 상기 제2휘발성 메모리와 상기 제2불휘발성 메모리를 제어하고, 상기 제2불휘발성 메모리에 포함된 저장 영역들 각각의 웨어 카운트와 상기 기준 웨어-아웃 카운트를 비교하는 제2코어를 포함하는 제2멀티 칩 셋;

제3휘발성 메모리, 제3불휘발성 메모리, 및 상기 제3휘발성 메모리와 상기 제3불휘발성 메모리를 제어하는 제3코어를 포함하는 제3멀티 칩 셋;

상기 제1멀티 칩 셋, 상기 제2멀티 칩 셋 및 상기 제3멀티 칩 셋과 연결되고, 상기 비교 결과에 기반하여 상기

제1불휘발성 메모리에 포함된 제1저장 영역이 상기 기준 웨어-아웃 카운트보다 높은 웨어 카운트를 가지고 상기 제2불휘발성 메모리에 포함된 제2저장 영역이 상기 기준 웨어-아웃 카운트보다 낮은 웨어 카운트를 가지는 경우, 상기 제1불휘발성 메모리에 포함된 상기 제1저장 영역의 제1논리 어드레스와 상기 제2불휘발성 메모리에 포함된 상기 제2저장 영역의 제2논리 어드레스를 스왑(swap)하는 컨트롤러; 및

상기 컨트롤러에 의하여 서로 스왑된 상기 제1논리 어드레스와 상기 제2논리 어드레스를 포함하는 테이블을 저장하는 메모리를 포함하되,

상기 컨트롤러는 호스트로부터 제1논리 어드레스를 수신하고, 상기 테이블을 참조하여 상기 스왑된 제2논리 어드레스를 상기 제2멀티 칩 셋으로 전송하고,

상기 제1불휘발성 메모리 및 제2불휘발성 메모리는 사용자 데이터를 저장하고, 상기 제3불휘발성 메모리는 오버프로비저닝(overprovisioning) 영역인 데이터 저장 장치.

## 발명의 설명

### 기술 분야

[0001] 본 발명의 개념에 따른 실시 예는 데이터 저장 장치에 관한 것으로, 특히 데이터 저장 장치의 수명과 성능을 제어하기 위해 스왑 매핑 테이블을 이용하여 소스 영역의 논리 어드레스를 목적지 영역의 논리 어드레스를 변경할 수 있는 데이터 저장 장치와 이를 포함하는 데이터 처리 시스템에 관한 것이다.

### 배경 기술

[0002] 메모리 장치는 데이터를 저장하는데 사용되고, 상기 메모리 장치는 휘발성 메모리 장치와 불휘발성 메모리 장치로 구분된다. 플래시 메모리 장치는 복수의 메모리 셀들이 한 번의 프로그램 작동으로 이레이즈(erase) 또는 프로그램(program)되는 EEPROM의 일 예이다. 예컨대, 프로그램 또는 리드는 페이지 단위로 수행되고, 이레이즈는 블록 단위로 수행된다. 블록은 복수의 페이지들을 포함할 수 있다.

[0003] 플래시 메모리 장치는 복수의 플래시 메모리 칩들을 포함할 수 있다. 상기 플래시 메모리 장치 전체의 수명과 성능은 상기 복수의 플래시 메모리 칩들 중에서 어느 하나에 따라 결정될 수 있다. 예컨대, 상기 복수의 플래시 메모리 칩들 중에서 어느 하나라도 닳거나(wear-out) 성능이 저하되면, 상기 복수의 플래시 메모리 칩들을 포함하는 플래시 메모리 장치 전체의 수명과 성능은 저하될 수 있다.

## 발명의 내용

### 해결하려는 과제

[0004] 본 발명이 이루고자 하는 기술적인 과제는 데이터 저장 장치의 수명과 성능을 제어하기 위해 스왑 매핑 테이블을 이용하여 소스 영역의 논리 어드레스를 목적지 영역의 논리 어드레스를 변경할 수 있는 데이터 저장 장치와 이를 포함하는 데이터 처리 시스템을 제공하는 것이다.

### 과제의 해결 수단

[0005] 본 발명의 실시 예에 따른 데이터 저장 장치는 제1휘발성 메모리, 제1불휘발성 메모리, 및 상기 제1휘발성 메모리와 상기 제1불휘발성 메모리를 제어하는 제1코어를 포함하는 제1멀티 칩 셋과, 제2휘발성 메모리, 제2불휘발성 메모리, 및 상기 제2휘발성 메모리와 상기 제2불휘발성 메모리를 제어하는 제2코어를 포함하는 제2멀티 칩 셋과, 상기 제1멀티 칩 셋과 상기 제2멀티 칩 셋과 접속되고, 상기 제1불휘발성 메모리의 제1저장 영역에 대한 제1논리 어드레스와 상기 제2불휘발성 메모리의 제2저장 영역에 대한 제2논리 어드레스를 스왑(swap)하는 컨트롤러를 포함한다.

[0006] 상기 컨트롤러는 상기 제1저장 영역에 저장된 제1데이터와 상기 제2저장 영역에 저장된 제2데이터를 스왑한다.

[0007] 상기 데이터 저장 장치는 상기 컨트롤러에 의해 스왑된 상기 제1논리 어드레스와 상기 제2논리 어드레스를 포함하는 테이블을 저장하는 메모리를 더 포함한다. 상기 컨트롤러는 호스트로부터 상기 제1논리 어드레스를 수신하고, 상기 테이블을 참조하여, 스왑된 상기 제2논리 어드레스를 상기 제2멀티 칩 셋으로 전송한다.

[0008] 실시 예에 따라, 상기 제1저장 영역은 사용자 데이터를 저장하는 영역이고, 상기 제2저장 영역은 오버프로비저

닝(overprovisioning) 영역이다.

- [0009] 상기 제1저장 영역에 대한 웨어 카운트(wear count)는 상기 제2저장 영역에 대한 웨어 카운트보다 높다. 상기 웨어 카운트는 이레이즈 카운트, 프로그램 카운트, 리드 카운트, 배드 블록 발생 횟수, 또는 리드 에러 발생 횟수일 수 있다.
- [0010] 실시 예에 따라, 상기 제1저장 영역은 핫 데이터와 콜드 데이터 중에서 어느 하나를 저장하는 영역이고, 상기 제2저장 영역은 상기 핫 데이터와 상기 콜드 데이터 중에서 다른 하나를 저장하는 영역이다.
- [0011] 실시 예에 따라, 상기 제1휘발성 메모리와 상기 제1코어는 제1반도체 패키지에 포함되고, 상기 제2휘발성 메모리와 상기 제2코어는 제2반도체 패키지에 포함된다. 실시 예에 따라, 상기 제1반도체 패키지와 상기 제2반도체 패키지 각각은 멀티 칩 패키지(multi chip package)이다.
- [0012] 실시 예에 따라, 상기 제1멀티 칩 셋과 상기 컨트롤러는 하나의 반도체 패키지로 구현된다. 상기 반도체 패키지는 임베디드 패키지 온 패키지(embedded package on package(ePOP))이다.
- [0013] 상기 제1논리 어드레스는 상기 컨트롤러의 제1요청에 따라 상기 제1코어로부터 제1채널을 통해 상기 컨트롤러로 전송되고, 상기 제2논리 어드레스는 상기 컨트롤러의 제2요청에 따라 상기 제2코어로부터 제2채널을 통해 상기 컨트롤러로 전송된다.
- [0014] 실시 예에 따라, 상기 제1코어는, 상기 제1요청에 따라, 상기 제1불휘발성 메모리에 포함된 저장 영역들 각각의 웨어 카운트와 기준 웨어 카운트를 비교하고, 비교 결과에 따라 상기 제1저장 영역에 대한 상기 제1논리 어드레스를 상기 제1채널을 통해 상기 컨트롤러로 전송하고, 상기 제2코어는, 상기 제2요청에 따라, 상기 제2불휘발성 메모리에 포함된 저장 영역들 각각의 웨어 카운트와 상기 기준 웨어 카운트를 비교하고, 비교 결과에 따라 상기 제2저장 영역에 대한 상기 제2논리 어드레스를 상기 제2채널을 통해 상기 컨트롤러로 전송하고, 상기 제1저장 영역에 대한 제1웨어 카운트는 상기 제2저장 영역에 대한 제2웨어 카운트보다 크다.
- [0015] 실시 예에 따라, 상기 제1코어는 상기 제1불휘발성 메모리에 포함된 저장 영역들 각각의 웨어 카운트와 기준 웨어 카운트를 비교하고, 비교 결과에 따라 상기 제1저장 영역에 대한 상기 제1논리 어드레스를 제1채널을 통해 상기 컨트롤러로 전송하고, 상기 제2코어는 상기 제2불휘발성 메모리에 포함된 저장 영역들 각각의 웨어 카운트와 상기 기준 웨어 카운트를 비교하고, 비교 결과에 따라 상기 제2저장 영역에 대한 상기 제2논리 어드레스를 상기 제2채널을 통해 상기 컨트롤러로 전송하고, 상기 제1저장 영역에 대한 제1웨어 카운트는 상기 제2저장 영역에 대한 제2웨어 카운트보다 크다.
- [0016] 상기 제1저장 영역과 상기 제2저장 영역 각각은 블록 또는 페이지이다.
- [0017] 상기 제1휘발성 메모리와 상기 제2휘발성 메모리 각각은 DRAM이고, 상기 제1불휘발성 메모리와 상기 제2불휘발성 메모리 각각은 플래시 메모리이고, 상기 데이터 저장 장치는 솔리드 스테이트 드라이브이다.
- [0018] 본 발명의 실시 예에 따른 데이터 처리 시스템은 호스트와, 상기 호스트에 접속된 데이터 저장 장치를 포함한다. 상기 데이터 저장 장치는 제1휘발성 메모리, 제1불휘발성 메모리, 및 상기 제1휘발성 메모리와 상기 제1불휘발성 메모리를 제어하는 제1코어를 포함하는 제1멀티 칩 셋과, 제2휘발성 메모리, 제2불휘발성 메모리, 및 상기 제2휘발성 메모리와 상기 제2불휘발성 메모리를 제어하는 제2코어를 포함하는 제2멀티 칩 셋과, 상기 제1멀티 칩 셋과 상기 제2멀티 칩 셋과 접속되고, 상기 제1불휘발성 메모리의 제1저장 영역에 대한 제1논리 어드레스와 상기 제2불휘발성 메모리의 제2저장 영역에 대한 제2논리 어드레스를 스왑하는 컨트롤러를 포함한다. 상기 컨트롤러는 상기 제1저장 영역에 저장된 제1데이터와 상기 제2저장 영역에 저장된 제2데이터를 스왑한다.

### 발명의 효과

- [0019] 본 발명의 실시 예에 따른 스케일-아웃 저장 장치들을 포함하는 데이터 저장 장치의 전체 수명과 전체 성능을 향상시키는 효과가 있다.
- [0020] 상기 스케일-아웃 저장 장치들 각각에 포함된 메모리 칩, 또는 상기 메모리 칩에 포함된 블록들 각각에 대한 웨어-카운트에 기초하여 소스 블록의 논리 어드레스와 목적지 블록의 논리 어드레스를 스왑할 수 있으므로, 상기 데이터 저장 장치는 상기 스케일-아웃 저장 장치들 사이에서 글로벌 웨어 레벨링을 수행할 수 있는 효과가 있다.
- [0021] 상기 스케일-아웃 저장 장치들 각각에 포함된 메모리 칩, 또는 상기 메모리 칩에 포함된 블록들 각각에 대한 웨어-카운트에 기초하여 소스 블록의 논리 어드레스와 목적지 블록의 논리 어드레스를 스왑할 수 있으므로, 상기

데이터 저장 장치는 상기 스케일-아웃 저장 장치들 각각의 성능과 수명을 균일하게 제어할 수 있는 효과가 있다.

### 도면의 간단한 설명

[0022]

본 발명의 상세한 설명에서 인용되는 도면을 보다 충분히 이해하기 위하여 각 도면의 상세한 설명이 제공된다.

도 1은 본 발명의 일 실시 예에 따른 데이터 처리 시스템을 나타내는 블록도이다.

도 2는 도 1에 도시된 데이터 저장 장치의 구체적인 블록도이다.

도 3은 도 2에 도시된 제1클러스터의 실시 예를 나타내는 블록도이다.

도 4는 도 3에 도시된 제1스케일-아웃 장치의 개략도이다.

도 5는 도 4에 도시된 제1스케일-아웃 장치의 반도체 패키지의 실시 예이다.

도 6은 도 1에 도시된 컨트롤러와 제1스케일-아웃 저장 장치를 포함하는 반도체 패키지의 실시 예이다.

도 7은 도 4에 도시된 제1스케일-아웃 컨트롤러의 블록도이다.

도 8a는 도 1에 도시된 데이터 처리 시스템의 작동을 설명하기 위한 블록도이다.

도 8b는 논리 어드레스 스왑 작동을 수행하는 도 1에 도시된 데이터 처리 시스템의 작동을 설명하기 위한 블록도이다.

도 9는 본 발명의 일 실시 예에 따른 스왑 매핑 테이블 업데이트 작동을 설명하기 위한 플로우이다.

도 10은 본 발명의 다른 실시 예에 따른 스왑 매핑 테이블 업데이트 작동을 설명하기 위한 플로우이다.

도 11은 논리 어드레스 스왑 작동을 위한 스왑 매핑 테이블의 일 실시 예이다.

도 12는 논리 어드레스, 물리 어드레스, 및 글로벌 논리 어드레스의 관계를 설명하는 개념도이다.

도 13은 도 8b에 도시된 데이터 저장 장치의 데이터 스왑 작동을 설명하는 개념도이다.

도 14는 논리 어드레스 스왑 작동을 위한 스왑 매핑 테이블의 다른 실시 예이다.

도 15는 스왑 매핑 테이블이 스케일-아웃 저장 장치에 포함될 때, 데이터 저장 장치의 작동을 설명하는 개념도이다.

도 16은 본 발명의 다른 실시 예에 따른 데이터 처리 시스템을 나타내는 블록도이다.

도 17은 본 발명의 또 다른 실시 예에 따른 데이터 처리 시스템을 나타내는 블록도이다.

도 18은 도 8a 또는 도 8b에 도시된 데이터 저장 장치의 작동을 설명하는 플로우 차트이다.

### 발명을 실시하기 위한 구체적인 내용

[0023]

본 명세서에 개시되어 있는 본 발명의 개념에 따른 실시 예들에 대해서 특정한 구조적 또는 기능적 설명은 단지 본 발명의 개념에 따른 실시 예들을 설명하기 위한 목적으로 예시된 것으로서, 본 발명의 개념에 따른 실시 예들은 다양한 형태들로 실시될 수 있으며 본 명세서에 설명된 실시 예들에 한정되지 않는다.

[0024]

본 발명의 개념에 따른 실시 예들은 다양한 변경들을 가할 수 있고 여러 가지 형태들을 가질 수 있으므로 실시 예들을 도면에 예시하고 본 명세서에서 상세하게 설명하고자 한다. 그러나, 이는 본 발명의 개념에 따른 실시 예들을 특정한 개시 형태들에 대해 한정하려는 것이 아니며, 본 발명의 사상 및 기술 범위에 포함되는 모든 변경, 균등물, 또는 대체물을 포함한다.

[0025]

제1 또는 제2 등의 용어는 다양한 구성 요소들을 설명하는데 사용될 수 있지만, 상기 구성 요소들은 상기 용어들에 의해 한정되어서는 안 된다. 상기 용어들은 하나의 구성 요소를 다른 구성 요소로부터 구별하는 목적으로만, 예컨대 본 발명의 개념에 따른 권리 범위로부터 벗어나지 않은 채, 제1구성 요소는 제2구성 요소로 명명될 수 있고 유사하게 제2구성 요소는 제1구성 요소로도 명명될 수 있다.

[0026]

어떤 구성 요소가 다른 구성 요소에 "연결되어" 있다거나 "접속되어" 있다고 언급된 때에는, 그 다른 구성 요소에 직접적으로 연결되어 있거나 또는 접속되어 있을 수도 있지만, 중간에 다른 구성 요소가 존재할 수도 있다고



이해되어야 할 것이다. 반면에, 어떤 구성 요소가 다른 구성 요소에 "직접 연결되어" 있다거나 "직접 접속되어" 있다고 언급된 때에는 중간에 다른 구성 요소가 존재하지 않는 것으로 이해되어야 할 것이다. 구성 요소들 간의 관계를 설명하는 다른 표현들, 즉 "~사이에"와 "바로 ~사이에" 또는 "~에 이웃하는"과 "~에 직접 이웃하는" 등도 마찬가지로 해석되어야 한다.

[0027] 본 명세서에서 사용한 용어는 단지 특정한 실시 예를 설명하기 위해 사용된 것으로서, 본 발명을 한정하려는 의도가 아니다. 단수의 표현은 문맥상 명백하게 다르게 뜻하지 않는 한, 복수의 표현을 포함한다. 본 명세서에서, "포함하다" 또는 "가지다" 등의 용어는 본 명세서에 기재된 특징, 숫자, 단계, 동작, 구성 요소, 부분품 또는 이들을 조합한 것이 존재함을 지정하려는 것이지, 하나 또는 그 이상의 다른 특징들이나 숫자, 단계, 동작, 구성 요소, 부분품 또는 이들을 조합한 것들의 존재 또는 부가 가능성을 미리 배제하지 않는 것으로 이해되어야 한다.

[0028] 다르게 정의되지 않는 한, 기술적이거나 과학적인 용어를 포함해서 여기서 사용되는 모든 용어들은 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자에 의해 일반적으로 이해되는 것과 동일한 의미를 나타낸다. 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.

[0029] 이하, 본 명세서에 첨부된 도면들을 참조하여 본 발명의 실시 예들을 상세히 설명한다.

[0030] 데이터 저장 장치, 예컨대, 데이터 서버의 데이터 처리 능력을 증가하는 방법으로서 스케일-업(scale-up)과 스케일-아웃(scale-out)이 있다. 상기 스케일-업은 상기 데이터 서버 그 자체의 용량을 증가시켜 상기 데이터 서버의 데이터 처리 능력을 향상시키는 것으로서 수직 스케일링(vertical scaling)으로 불리기도 한다. 상기 스케일-아웃은 접속된 데이터 서버의 개수를 늘려 상기 데이터 서버를 포함하는 시스템의 데이터 처리 능력을 향상시키는 것으로서 수평 스케일링(horizontal scaling)으로 불리기도 한다.

[0031] 스케일-업(scale-up) 구조는 갖는 데이터 저장 장치에서, 상기 데이터 저장 장치에 포함된 불휘발성 메모리 장치들의 개수가 증가하면, 상기 불휘발성 메모리 장치들을 제어하는 컨트롤러와 상기 불휘발성 메모리 장치들 사이에 구현된 인터페이스의 로딩 커패시턴스(load capacitance)가 증가할 수 있다. 이에 따라, 상기 데이터 저장 장치의 데이터 처리 속도가 저하될 수 있다. 따라서, 본 명세서에서 설명된 데이터 저장 장치는 스케일-아웃 구조를 갖는 데이터 저장 장치이다.

[0032] 스케일-아웃 구조를 갖는 데이터 저장 장치는 스케일-아웃 저장 장치, 상기 스케일-아웃 저장 장치에 포함된 반도체 칩, 및/또는 상기 반도체 칩의 메모리 셀 어레이에 포함된 블록의 내구성(endurance)을 대표할 수 있는 대표값(이하, "웨어-카운트(wear-count)"라 한다)를 이용하여 논리 어드레스(logical address) 또는 글로벌(global) 논리 어드레스를 변경 또는 스왑할 수 있다.

[0033] 예컨대, 웨어-카운트는 해당 블록이 몇 번 이레이즈되었는지를 나타내는 이레이즈 카운트(erase count), 해당 블록이 몇 번 프로그램되었는지를 나타내는 프로그램(program) 카운트, 해당 블록이 몇 번 리드되었는지를 나타내는 리드(read) 카운트, 해당 블록에 대한 배드(bad) 블록 발생 횟수, 또는 해당 블록에서 리드 에러(read error)의 발생 횟수를 포함할 수 있으나 웨어-카운트가 이들에 한정되는 것은 아니다. 여기서, 이레이즈 카운트는 P/E 사이클(program-erase(P/E) cycle)을 포함하는 개념으로 이해될 수 있다.

[0034] 또한, 스케일-아웃 저장 장치(scale-out storage device)는 휘발성 메모리, 적어도 하나의 불휘발성 메모리, 및 상기 휘발성 메모리와 상기 적어도 하나의 불휘발성 메모리를 제어하는 스케일-아웃 컨트롤러를 포함할 수 있다. 상기 휘발성 메모리, 상기 적어도 하나의 불휘발성 메모리, 및 상기 스케일-아웃 컨트롤러 각각이 반도체 칩 또는 반도체 패키지로 구현될 때, 상기 스케일-아웃 저장 장치는 멀티 칩 셋(multi chip set)으로 불릴 수 있다. 스케일-아웃 컨트롤러는 코어(core)를 의미할 수 있다.

[0035] 도 1은 본 발명의 일 실시 예에 따른 데이터 처리 시스템을 나타내는 블록도이다. 도 1을 참조하면, 데이터 처리 시스템(100)은 호스트(200)와, 인터페이스 (110)를 통해 호스트(200)와 명령 및/또는 데이터를 주거나 받을 수 있는 데이터 저장 장치(300)를 포함할 수 있다.

[0036] 데이터 저장 장치(300)에 포함된 메모리 클러스터(400)의 구조가 도 2에 예시적으로 도시된 구조를 가질 때, 데이터 저장 장치(300)는 스케일-아웃(scale-out) 구조로 구현될 수 있다. 예컨대, 데이터 처리 시스템(100)은 메모리 시스템을 의미할 수 있으나 이에 한정되는 것은 아니다.

- [0037] 실시 예들에 따라, 데이터 처리 시스템(100)은 PC(personal computer), 워크스테이션(workstation), 데이터 센터, 인터넷 데이터 센터(internet data center(IDC)), SAN(storage area network), NAS(network attached storage) 또는 모바일 컴퓨팅 장치로 구현될 수 있으나 이에 한정되는 것은 아니다.
- [0038] 메모리 클러스터(400)를 포함하는 모바일 컴퓨팅 장치는 랩탑 컴퓨터(laptop computer), 이동 전화기, 스마트폰(smart phone), 태블릿(tablet) PC, PDA (personal digital assistant), EDA(enterprise digital assistant), 디지털 스틸 카메라(digital still camera), 디지털 비디오 카메라(digital video camera), PMP (portable multimedia player), PND(personal navigation device 또는 portable navigation device), 휴대용 게임 콘솔(handheld game console), 모바일 인터넷 장치(mobile internet device(MID)), 웨어러블 컴퓨터, 사물 인터넷(internet of things(IoT)) 장치, 만물 인터넷(internet of everything(IoE)) 장치, 드론 (drone), 또는 e-북 (e-book)으로 구현될 수 있으나 이에 한정되는 것은 아니다.
- [0039] 인터페이스(110)는 SATA(serial advanced technology attachment) 인터페이스, SATAe(SATA express) 인터페이스, SAS(serial attached small computer system interface(SCSI)) 인터페이스, PCIe(peripheral component interconnect express) 인터페이스, NVMe(non-volatile memory Express) 인터페이스, AHCI(advanced host controller interface) 인터페이스, 또는 멀티미디어 카드(multimedia card(MMC)) 인터페이스로 구현될 수 있으나 이에 한정되는 것은 아니다. 실시 예들에 따라, 인터페이스(110)는 전기 신호들 또는 광신호들을 전송할 수 있다.
- [0040] 호스트(200)는, 인터페이스(110)를 통해, 데이터 저장 장치(300)의 데이터 처리 작동(예컨대, 라이트(또는 프로그램) 작동 또는 리드 작동 등)을 제어할 수 있다. 호스트 (200)는 호스트 컨트롤러를 의미할 수 있다.
- [0041] CPU(220)와 제1인터페이스(230)는, 버스 구조(210)를 통해, 명령 및/또는 데이터를 서로 주거나 받을 수 있다. 비록, 도 1에는 버스 구조(210), CPU(220), 및 제1인터페이스(230)를 포함하는 호스트(200)가 도시되어 있으나, 본 발명의 기술적 사상이 도 1에 도시된 호스트(200)의 블록도에 한정되는 것은 아니다.
- [0042] 실시 예에 따라, 호스트(200)는 집적 회로(integrated circuit(IC)), 마더보드(motherboard), 시스템 온 칩(system on chip(SoC)), 애플리케이션 프로세서 (application processor(AP)), 모바일(mobile) AP, 웹 서버(web server), 데이터 서버, 또는 데이터베이스 서버로 구현될 수 있으나 이에 한정되는 것은 아니다.
- [0043] 예컨대, 버스 구조(210)는 AMBA(advanced microcontroller bus architecture), AHB(advanced high-performance bus), APB(advanced peripheral bus), AXI(advanced eXtensible interface), ASB(advanced system bus), ACE(AXI Coherency Extensions), 또는 이들의 결합으로 구현될 수 있으나 이에 한정되는 것은 아니다.
- [0044] CPU(220)는 데이터 저장 장치(300)의 라이트 작동(write operation)을 제어할 수 있는 라이트 요청, 또는 데이터 저장 장치(300)의 리드 작동(read operation)을 제어할 수 있는 리드 요청을 생성할 수 있다. 상기 라이트 요청은 라이트 어드레스(예컨대, 논리 어드레스)를 포함할 수 있고, 상기 리드 요청은 리드 어드레스(예컨대, 논리 어드레스)를 포함할 수 있다. 예컨대, CPU(220)는 하나 또는 그 이상의 코어들을 포함할 수 있다. 요청(request)은 명령(command)을 의미할 수 있다.
- [0045] 제1인터페이스(230)는 데이터 저장 장치(300)로 전송될 명령 및/또는 데이터의 포맷(format)을 변경하고, 변경된 포맷을 갖는 명령 및/또는 변경된 포맷을 갖는 데이터를 인터페이스(110)를 통해 데이터 저장 장치(300)로 전송할 수 있다.
- [0046] 또한, 제1인터페이스(230)는 데이터 저장 장치(300)로부터 전송된 응답 (response) 및/또는 데이터의 포맷을 변경하고, 변경된 포맷을 갖는 응답 및/또는 변경된 포맷을 갖는 데이터를 버스 구조(210)를 통해 CPU(220)로 전송할 수 있다. 실시 예에 따라, 제1인터페이스(230)는 명령 및/또는 데이터를 송수신할 수 있는 송수신기를 포함할 수 있다. 제1인터페이스(230)의 구조와 작동은 인터페이스(110)의 구조와 작동에 적합하게(compatible) 구현될 수 있다.
- [0047] 데이터 저장 장치(300)는 컨트롤러(310), 버퍼(360), 전력 관리(power management) IC(370), 및 메모리 클러스터(400)를 포함할 수 있다. 여기서, 메모리 클러스터(400)는 메모리 장치들의 집합을 의미할 수 있다. 또한, 메모리 클러스터 (400)는, 도 2에 도시된 바와 같이, 스케일-아웃 장치들과 메모리 장치들을 포함할 수 있다.
- [0048] 데이터 저장 장치(300)는 플래시 메모리-기반 메모리 장치로 구현될 수 있으나 이에 한정되는 것은 아니다. 예컨대, 데이터 저장 장치(300)는 SSD(solid-state drive or solid-state disk), 임베디드 SSD(embedded

SSD(eSSD)), 유니버설 플래시 스토리지(universal flash storage(UFS)), 멀티미디어 카드(multimedia card(MMC)), 임베디드 MMC(embedded MMC(eMMC)) 또는 매니지드(managed) NAND로 구현될 수 있으나 이에 한정되는 것은 아니다. 예컨대, 플래시 메모리-기반 메모리 장치는 NAND-타입 플래시 메모리 장치 또는 NOR-타입 플래시 메모리 장치로 구현될 수 있다.

[0049] 플래시 메모리-기반 메모리 장치는 메모리 셀 어레이를 포함할 수 있다. 상기 메모리 셀 어레이는 복수의 메모리 셀들을 포함할 수 있다. 예컨대, 상기 메모리 셀 어레이는 2차원 메모리 셀 어레이 또는 3차원 메모리 셀 어레이를 포함할 수 있다.

[0050] 3차원 메모리 셀 어레이는 실리콘 기판 위(on or above)에 배치된 액티브 영역(active region)을 갖는 메모리 셀들의 어레이의 하나 또는 그 이상의 물리적인 레벨들 내에서 모노리식하게(monolithically) 형성되고, 상기 메모리 셀들의 작동에 관련된 회로를 포함할 수 있다. 상기 회로는 상기 기판의 내부 또는 위(on or above)에 형성될 수 있다. 모노리식(monolithic)이라는 용어는 어레이의 각 레벨의 레이어들(layers)이 상기 어레이의 각 하부 레벨(each underlying level)의 레이어들에 직접 증착 (directly deposited)되는 것을 의미한다. 3차원 메모리 셀 어레이는 적어도 하나의 메모리 셀이 다른 메모리 셀의 위에 위치하도록 수직으로 배향되는(vertically oriented) 수직 NAND 스트링을 포함할 수 있다. 상기 적어도 하나의 메모리 셀은 전하 트랩 레이어(charge trap layer)를 포함할 수 있다.

[0051] 실시 예들에 따라, 데이터 저장 장치(300)는 하드 디스크 드라이브(hard disk drive(HDD)), PRAM(phase change RAM) 장치, MRAM(magnetoresistive RAM) 장치, STT-MRAM(spin-transfer torque MRAM) 장치, FRAM(ferroelectric RAM) 장치 또는 RRAM(resistive RAM) 장치로 구현될 수 있으나 이에 한정되는 것은 아니다.

[0052] 컨트롤러(310)는 호스트(200), 버퍼(360), 및 메모리 클러스터(400) 사이에서 주거나 받는 명령 및/또는 데이터의 전송(또는 처리)을 제어할 수 있다. 실시 예들에 따라, 컨트롤러(310)는 IC, SoC, 또는 칩들의 집합으로 구현될 수 있다.

[0053] 컨트롤러(310)는 버스 구조(311), 내부 메모리(315), 제2인터페이스(320), 적어도 하나의 CPU(330 및/또는 331), 버퍼 컨트롤러(340), 및 제3인터페이스(350)를 포함할 수 있다.

[0054] 버스 구조(311)는 AMBA, AHB, APB, AXI, ASB, ACE 또는 이들의 결합으로 구현될 수 있으나 이에 한정되는 것은 아니다.

[0055] 내부 메모리(315)는 컨트롤러(310)의 작동에 필요한 데이터 또는 컨트롤러 (310)에 의해 수행된 데이터 처리 작동(예컨대, 라이트 작동 또는 리드 작동)에 의해 생성된 데이터를 저장할 수 있다. 예컨대, 내부 메모리(315)는 CPU(330 및/또는 331)에 의해 실행될 수 있는 제1플래시 변환 계층(flash translation layer(FTL)) 코드를 저장할 수 있다. FTL 코드는 간단히 FTL이라고 불릴 수 있다. 예컨대, 데이터 저장 장치(300)가 부팅(booting)될 때, 제1FTL 코드는 메모리 클러스터(400)로부터 내부 메모리(315)로 로드되고 상기 제1FTL 코드는 CPU(330 및/또는 331)에 의해 실행될 수 있다.

[0056] 실시 예들에 따라, 내부 메모리(315)는 RAM(random access memory), DRAM (dynamic RAM), SRAM(static RAM), 버퍼(buffer), 버퍼 메모리, 캐시(cache), 또는 강하게 결합된 메모리(tightly coupled memory (TCM))로 구현될 수 있으나 내부 메모리(315)의 종류가 이들에 한정되는 것은 아니다.

[0057] 제2인터페이스(320)는 호스트(200)로 전송될 응답 및/또는 데이터의 포맷을 변경하고, 변경된 포맷을 갖는 응답 및/또는 변경된 포맷을 갖는 데이터를 인터페이스(110)를 통해 호스트(200)로 전송할 수 있다. 또한, 제2인터페이스(320)는 호스트(200)로부터 전송된 명령 및/또는 데이터를 수신하고, 수신된 명령 및/또는 데이터의 포맷을 변경하고, 변경된 포맷을 갖는 명령 및/또는 변경된 포맷을 갖는 데이터를 적어도 하나의 CPU(330 및/또는 331) 및/또는 버퍼 컨트롤러(340)로 전송할 수 있다. 실시 예에 따라, 제2인터페이스(320)는 신호 및/또는 데이터를 송수신할 수 있는 송수신기를 포함할 수 있다.

[0058] 제2인터페이스(320)의 구조와 작동은 인터페이스(110)의 구조와 작동에 적합하게 구현될 수 있다. 예컨대, 제2 인터페이스(320)는 SATA 인터페이스, SATAe 인터페이스, SAS 인터페이스, PCIe 인터페이스, NVMe 인터페이스, AHCI 인터페이스, MMC 인터페이스, NAND-타입 플래시 메모리 인터페이스, 또는 NOR-타입 플래시 메모리 인터페이스로 구현될 수 있으나 이에 한정되는 것은 아니다.

[0059] 하나 또는 그 이상의 CPU(330 및/또는 331)는, 버스 구조(311)를 통해, 내부 메모리(315), 제2인터페이스(320),

버퍼 컨트롤러(340), 및 제3인터페이스(350)를 제어할 수 있다. 각 CPU(330과 331)는 하나 또는 그 이상의 코어들을 포함할 수 있다. 실시 예에 따라, 하나 또는 그 이상의 CPU(330 및/또는 331)는 PMIC(370)를 제어할 수 있다.

- [0060] 예컨대, 제1CPU(330)는 제2인터페이스(320)를 통해 호스트(200)와 데이터를 주거나 받을 수 있다. 제2CPU(331)는 제3인터페이스(350)를 통해 메모리 클러스터 (400)와 데이터를 주거나 받을 수 있다. 실시 예에 따라, 제1CPU(330)와 제2CPU (331)는 멀티-CPU를 구성할 수 있다. 실시 예에 따라, 제1CPU(330)는 제2CPU(331)를 제어할 수 있으나 이에 한정되는 것은 아니다.
- [0061] CPU(330 및/또는 331)는 스왑될 논리 어드레스들에 대한 스왑 매핑 테이블을 생성하고, 상기 스왑 매핑 테이블을 버퍼 컨트롤러(340)를 이용하여 버퍼(360)에 저장할 수 있다.
- [0062] 버퍼 컨트롤러(340)는, 제1CPU(330) 또는 제2CPU(331)의 제어에 따라, 버퍼 (360)에 데이터(또는 스왑 매핑 테이블)를 라이트하거나 버퍼(360)로부터 데이터(또는 스왑 매핑 테이블)를 리드할 수 있다. 버퍼 컨트롤러(340)는 버퍼(360)에 대한 라이트 작동과 리드 작동을 제어할 수 있는 컨트롤러 또는 버퍼 매니저(buffer manager)로 불릴 수 있다.
- [0063] 제3인터페이스(350)는, 제1CPU(330) 또는 제2CPU(331)의 제어에 따라, 복수의 메인 채널들(CHA, CHB, ..., CHC) 중에서 해당하는 메인 채널을 통해 메모리 클러스터(400)에 대한 데이터 처리 작동들(예컨대, 라이트 작동들 또는 리드 작동들, 등)을 제어할 수 있다.
- [0064] 실시 예들에 따라, 제3인터페이스(350)는 SATA 인터페이스, SATAe 인터페이스, SAS 인터페이스, PCIe 인터페이스, NVMe 인터페이스, AHCI 인터페이스, MMC 인터페이스, NAND-타입 플래시 메모리 인터페이스, 또는 NOR-타입 플래시 메모리 인터페이스로 구현될 수 있으나 이에 한정되는 것은 아니다.
- [0065] 예컨대, 제3인터페이스(350)는 에러 정정 코드(error correction code(ECC)) 엔진(미도시)을 포함할 수 있다. 상기 EEC 엔진은 메모리 클러스터(400)에 저장될 데이터 및/또는 메모리 클러스터(400)로부터 출력된 데이터에 포함된 에러를 정정할 수 있다. 실시 예에 따라, ECC 엔진은 컨트롤러(310) 내부의 어디라도 구현될 수 있다.
- [0066] 버퍼(360)는, 버퍼 컨트롤러(340)의 제어에 따라, 데이터(또는, 스왑 매핑 테이블)를 라이트하거나 데이터(또는 스왑 매핑 테이블)를 리드할 수 있다. 실시 예들에 따라, 휘발성 메모리 장치로 구현될 수 있는 버퍼(360)는 버퍼 메모리, RAM, SRAM, 또는 DRAM으로 구현될 수 있으나 이에 한정되는 것은 아니다.
- [0067] 버퍼(360)는 복수의 클러스터들(410, 430, ... 450)에 대한 논리 어드레스-물리 어드레스 변환을 위한 매핑 테이블 및/또는 스왑 매핑 테이블을 저장하는 제1영역과, 캐시(cache)의 기능을 수행할 수 있는 제2영역을 포함할 수 있으나 이에 한정되는 것은 아니다.
- [0068] 예컨대, CPU(330 및/또는 331)에 의해 실행되는 제1FTL 코드는 상기 제1영역에 저장된 매핑 테이블을 이용하여 논리 어드레스-물리 어드레스 변환을 수행하거나 상기 제1영역에 저장된 스왑 매핑 테이블을 이용하여 소스(source) 논리 어드레스들을 목적지(destination) 논리 어드레스들과 스왑(swap)할 수 있다. 예컨대, 매핑 테이블과 스왑 매핑 테이블은 하나의 매핑 테이블로 구현되거나 서로 다른 매핑 테이블로 구현될 수 있다.
- [0069] 실시 예에 따라, 컨트롤러(310)와 버퍼(360) 각각이 서로 다른 반도체 칩 (semiconductor chip)으로 구현될 때, 컨트롤러(310)와 버퍼(360)는 하나의 패키지, 예컨대, 패키지-온-패키지(package-on-package(PoP)), 멀티-칩 패키지(multi-chip package(MCP)) 또는 시스템-인 패키지(system-in package(SiP))로 구현될 수 있으나 상기 패키지의 종류가 이들에 한정되는 것은 아니다. 예컨대, 버퍼(360)를 포함하는 제1칩은 컨트롤러(310)를 포함하는 제2칩의 위(above)에 스택 볼들(stack balls)을 통해 적층될 수 있다.
- [0070] PMIC(370)는 컨트롤러(310), 버퍼(360), 및/또는 메모리 클러스터(400)로 공급되는 작동 전압들을 제어할 수 있다. 예컨대, 컨트롤러(310)로 공급되는 제1작동 전압, 버퍼(360)로 공급되는 제2작동 전압, 및 메모리 클러스터(400)로 공급되는 제3작동 전압은 서로 동일하거나 서로 다를 수 있다.
- [0071] 메모리 클러스터(400)는 복수의 클러스터들(410, 430, ... 450)을 포함할 수 있다. 제1클러스터(410)는 제1메인 채널(CHA)에 접속될 수 있고, 제2클러스터(430)는 제2메인 채널(CHB)에 접속될 수 있고, 제3클러스터(450)는 제3메인 채널(CHA)에 접속될 수 있다.
- [0072] 복수의 클러스터들(410, 430, ... 450) 각각의 구조는 도 2부터 도 8b를 참조하여 상세히 설명될 것이다. 복수의 클러스터들(410, 430, ... 450) 각각의 구조는 실질적으로 동일 또는 유사하게 구현될 수 있다.



- [0073] 본 명세서에서 메인 채널(main channel)은 컨트롤러(310), 즉, 제3인터페이스(350)와 해당하는 하나의 클러스터 사이에 존재하는 독립적인 데이터 패스(data path)를 의미할 수 있다. 상기 데이터 패스는 데이터 및/또는 제어 신호들을 전송할 수 있는 전송 라인들을 포함할 수 있다.
- [0074] 웨이(way)는 하나의 메인 채널을 공유하는 하나 또는 그 이상의 불휘발성 메모리 장치들의 그룹을 의미할 수 있다. 따라서, 하나의 메인 채널에는 복수의 웨이들이 접속될 수 있다. 상기 불휘발성 메모리 장치는 다이(die), 메모리 장치, 또는 반도체 패키지를 의미할 수 있다. 컨트롤러(310)는 A-채널\*B-웨이를 제어할 수 있다. 이때, A와 B 각각은 1 이상의 자연수를 의미할 수 있다.
- [0075] 도 2는 도 1에 도시된 데이터 저장 장치의 구체적인 블록도이고, 도 3은 도 2에 도시된 제1클러스터의 실시 예를 나타내는 블록도이다. 도 2에 도시된 데이터 저장 장치(300)는 스케일-아웃 구조를 갖는 데이터 저장 장치로 구현될 수 있다.
- [0076] 제1클러스터(410)는 적어도 하나의 스케일-아웃 장치와 적어도 하나의 불휘발성 메모리 장치, 예컨대 NAND-타입 플래시 메모리 장치(NAND)를 포함할 수 있다. 나머지 클러스터들(430~450) 각각은 적어도 하나의 스케일-아웃 장치와 적어도 하나의 불휘발성 메모리 장치, 예컨대 NAND-타입 플래시 메모리 장치(NAND)를 포함할 수 있다.
- [0077] 도 1부터 도 3을 참조하면, 각 클러스터(410~450)의 구조와 작동은 실질적으로 동일 또는 유사하므로, 제1클러스터(410)의 구조와 작동이 대표적으로 설명된다. 제1클러스터(410)는 복수의 스케일-아웃 저장 장치들(410-1~410-4)을 포함할 수 있다. 비록, 도 3에서는 4개의 스케일-아웃 저장 장치들(410-1~410-4)이 도시되어 있으나, 본 발명의 기술적 사상은 제1클러스터(410)에 포함되는 복수의 스케일-아웃 저장 장치들(410-1~410-4)의 개수에 한정되는 것은 아니다.
- [0078] 제1클러스터(410)는 스케일-아웃 장치들(410-1A, 410-2A, ..., 410-4A)과, 스케일-아웃 장치들(410-1A, 410-2A, ..., 410-4A) 각각에 접속된 NAND-타입 플래시 메모리 장치들(NAND)을 포함할 수 있다. 제1스케일-아웃 저장 장치(410-1)는 제1스케일-아웃 장치(410-1A)와 제1스케일-아웃 장치(410-1A)에 접속된 불휘발성 메모리 장치들(NAND)을 포함할 수 있다. 제2스케일-아웃 저장 장치(410-2)는 제2스케일-아웃 장치(410-2A)와 제2스케일-아웃 장치(410-2A)에 접속된 불휘발성 메모리 장치들(NAND)을 포함할 수 있다. 제4스케일-아웃 저장 장치(410-4)는 제4스케일-아웃 장치(410-4A)와 제4스케일-아웃 장치(410-4A)에 접속된 불휘발성 메모리 장치들(NAND)을 포함할 수 있다. 예컨대, 각 스케일-아웃 저장 장치(410-1), 410-2, ..., 410-4)는 멀티 칩 셋을 의미할 수 있다.
- [0079] 각 스케일-아웃 장치(410-1A, 410-2A, ..., 410-4A)는 각 스케일-아웃 컨트롤러(411-1, 411-2, ..., 411-4)와, 각 휘발성 메모리 장치(413-1, 413-2, ..., 413-4)를 포함할 수 있다. 각 스케일-아웃 컨트롤러(411-1, 411-2, ..., 411-4)의 구조는 실질적으로 동일 또는 유사하다.
- [0080] 제1스케일-아웃 컨트롤러(411-1)는 제1메인 채널(CHA), 채널들(CH11~CH1m; m은 2 이상의 자연수), 및 제1휘발성 메모리 장치(413-1)에 접속될 수 있다. 각 채널(CH11~CH1m)에는 대응되는 불휘발성 메모리 장치들(NAND)이 접속될 수 있다. 예컨대, 불휘발성 메모리 장치들(NAND) 각각은 NAND-타입 플래시 메모리 또는 매니지드 NAND 플래시 메모리 장치일 수 있으나 이에 한정되는 것은 아니다. 여기서, 각 채널은 제1스케일-아웃 컨트롤러(411-1)와 해당하는 불휘발성 메모리 장치들(NAND) 사이에 존재하는 독립적인 데이터 패스를 의미할 수 있다.
- [0081] 제1스케일-아웃 컨트롤러(411-1)는 각 채널(CH11~CH1m)을 통해 불휘발성 메모리 장치들(NAND) 각각의 작동을 제어할 수 있다. 예컨대, 컨트롤러(310)는, 제1메인 채널(CHA)을 통해, 제1스케일-아웃 컨트롤러(411-1)의 작동을 제어할 수 있다.
- [0082] 제1휘발성 메모리 장치(413-1)는 각 채널(CH11~CH1m)에 접속된 불휘발성 메모리 장치들(NAND)에 대한 논리 어드레스-물리 어드레스 매핑 정보(또는 매핑 테이블)를 저장할 수 있는 제1영역을 포함할 수 있다. 또한, 제1휘발성 메모리 장치(413-1)는 각 채널(CH11~CH1m)에 접속된 불휘발성 메모리 장치들(NAND) 중에서 적어도 하나에 라이트될 데이터 또는 각 채널(CH11~CH1m)에 접속된 불휘발성 메모리 장치들(NAND) 중에서 적어도 하나로부터 출력된 데이터를 임시로 저장할 수 있는 제2영역(또는 캐시 영역)을 더 포함할 수 있다.
- [0083] 예컨대, 도 7에 도시된 바와 같이 제1스케일-아웃 컨트롤러(411-1)에 포함된 CPU(423-1 또는 423-2)에 의해 실행되는 제2FTL 코드는 제1휘발성 메모리 장치(413-1)의 제1영역에 저장된 매핑 테이블을 이용하여 논리 어드레스-물리 어드레스 변환을 수행할 수 있다. 상기 제2FTL 코드는 제1클러스터(410)에 포함된 불휘발성 메모리 장치로부터 내부 메모리(425)로 로드되고 CPU(423-1 또는 423-2)에 의해 실행될 수 있다.

- [0084] 제2스케일-아웃 컨트롤러(411-2)는 제1메인 채널(CHA), 채널들(CH21~CH2m), 및 제2휘발성 메모리 장치(413-2)에 접속될 수 있다. 각 채널(CH21~CH2m)에는 대응되는 불휘발성 메모리 장치들(NAND)이 접속될 수 있다.
- [0085] 제2스케일-아웃 컨트롤러(411-2)는 각 채널(CH21~CH2m)을 통해 불휘발성 메모리 장치들(NAND) 각각의 작동을 제어할 수 있다. 예컨대, 컨트롤러(310)는, 메인 채널(CHA)을 통해, 제2스케일-아웃 컨트롤러(415-2)의 작동을 제어할 수 있다.
- [0086] 제2휘발성 메모리 장치(413-2)는 각 채널(CH21~CH2m)에 접속된 불휘발성 메모리 장치들(NAND)에 대한 논리 어드레스-물리 어드레스 매핑 정보(또는 매핑 테이블)를 저장할 수 있는 제1영역을 포함할 수 있다. 또한, 제2휘발성 메모리 장치(413-2)는 각 채널(CH21~CH2m)에 접속된 불휘발성 메모리 장치들(NAND) 중에서 적어도 하나에 라이트될 데이터 또는 각 채널(CH21~CH2m)에 접속된 불휘발성 메모리 장치들(NAND) 중에서 적어도 하나로부터 출력된 데이터를 임시로 저장할 수 있는 제2영역(또는 캐시 영역)을 더 포함할 수 있다.
- [0087] 도 7에 도시된 바와 같이 제1스케일-아웃 컨트롤러(411-1)의 구조와 제2스케일-아웃 컨트롤러(411-2)의 구조가 동일할 때, 제2스케일-아웃 컨트롤러(411-2)에 포함된 CPU(423-1 또는 423-2)에 의해 실행되는 제3FTL 코드는 제2휘발성 메모리 장치(413-2)의 제1영역에 저장된 매핑 테이블을 이용하여 논리 어드레스-물리 어드레스 변환을 수행할 수 있다. 상기 제3FTL 코드는 제1클러스터(410)에 포함된 불휘발성 메모리 장치로부터 내부 메모리(425)로 로드되고 CPU(423-1 또는 423-2)에 의해 실행될 수 있다.
- [0088] 제4스케일-아웃 컨트롤러(411-4)는 제1메인 채널(CHA), 채널들(CH41~CH4m), 및 제4휘발성 메모리 장치(413-4)에 접속될 수 있다. 각 채널(CH41~CH4m)에는 대응되는 불휘발성 메모리 장치들(NAND)이 접속될 수 있다.
- [0089] 제4스케일-아웃 컨트롤러(411-4)는 각 채널(CH41~CH4m)을 통해 불휘발성 메모리 장치들(NAND) 각각의 작동을 제어할 수 있다. 예컨대, 컨트롤러(310)는, 메인 채널(CHA)을 통해, 제4스케일-아웃 컨트롤러(411-4)의 작동을 제어할 수 있다.
- [0090] 제4휘발성 메모리 장치(413-4)는 각 채널(CH41~CH4m)에 접속된 대응되는 불휘발성 메모리 장치들(NAND)에 대한 논리 어드레스-물리 어드레스 매핑 정보(또는 매핑 테이블)를 저장할 수 있는 제1영역을 포함할 수 있다. 또한, 제4휘발성 메모리 장치(413-4)는 각 채널(CH41~CH4m)에 접속된 대응되는 불휘발성 메모리 장치들(NAND) 중에서 적어도 하나에 라이트될 데이터 또는 각 채널(CH41~CH4m)에 접속된 대응되는 불휘발성 메모리 장치들(NAND) 중에서 적어도 하나로부터 출력된 데이터를 임시로 저장할 수 있는 제2영역(또는 캐시 영역)을 더 포함할 수 있다.
- [0091] 도 7에 도시된 바와 같이 제1스케일-아웃 컨트롤러(411-1)의 구조와 제4스케일-아웃 컨트롤러(411-4)의 구조가 동일할 때, 제4스케일-아웃 컨트롤러(411-4)에 포함된 CPU(423-1 또는 423-2)에 의해 실행되는 제5FTL 코드는 제4휘발성 메모리 장치(413-4)의 제1영역에 저장된 매핑 테이블을 이용하여 논리 어드레스-물리 어드레스 변환을 수행할 수 있다. 상기 제5FTL 코드는 제1클러스터(410)에 포함된 불휘발성 메모리 장치로부터 내부 메모리(425)로 로드되고 CPU(423-1 또는 423-2)에 의해 실행될 수 있다.
- [0092] 각 휘발성 메모리 장치(413-1~413-4)는 버퍼 또는 버퍼 메모리를 의미할 수 있고, RAM, SRAM, 또는 DRAM으로 구현될 수 있으나 이에 한정되는 것은 아니다.
- [0093] 비록, 도 3에서는 각 스케일-아웃 저장 장치(410-1~410-4)에 동일한 개수의 채널들을 포함하는 것으로 도시되어 있으나, 실시 예들에 따라 각 스케일-아웃 저장 장치(410-1~410-4)에 포함된 채널들의 개수는 서로 다를 수 있다.
- [0094] 도 4는 도 3에 도시된 제1스케일-아웃 장치의 개략도이고, 도 7은 도 4에 도시된 제1스케일-아웃 컨트롤러의 블록도이다.
- [0095] 도 1부터 도 5, 및 도 7을 참조하면, 각 스케일-아웃 컨트롤러(411-1~411-4)의 구조와 작동은 실질적으로 동일 또는 유사하므로, 제1스케일-아웃 컨트롤러(411-1)의 구조와 작동이 대표적으로 설명될 것이다.
- [0096] 제1스케일-아웃 컨트롤러(411-1)는 컨트롤러(310), 제1휘발성 메모리 장치(413-1), 및 제1스케일-아웃 저장 장치(410-1)에 포함된 불휘발성 메모리 장치들(NAND) 사이에서 주거나 받는 명령 및/또는 데이터의 전송을 제어할 수 있다. 실시 예들에 따라, 제1스케일-아웃 컨트롤러(411-1)는 IC 또는 SoC로 구현될 수 있다.
- [0097] 제1스케일-아웃 컨트롤러(411-1)는 버스 구조(420), 제4인터페이스(421), 적어도 하나의 CPU(423-1 및/또는 423-2), 내부 메모리(425), 버퍼 컨트롤러(427), 및 제5인터페이스(429)를 포함할 수 있다. 버스 구조(420)는

AMBA, AHB, APB, AXI, ASB, ACE, 또는 이들의 결합으로 구현될 수 있으나 이에 한정되는 것은 아니다.

- [0098] 제4인터페이스(421)는 컨트롤러(310)로 전송될 응답 및/또는 데이터의 포맷을 변경하고, 변경된 포맷을 갖는 응답 및/또는 변경된 포맷을 갖는 데이터를 제1메인 채널(CHA)을 통해 컨트롤러(310)로 전송할 수 있다. 또한, 제4인터페이스(421)는 컨트롤러(310)로부터 전송된 명령 및/또는 데이터를 수신하고, 수신된 명령 및/또는 데이터의 포맷을 변경하고, 변경된 포맷을 갖는 명령 및/또는 변경된 포맷을 갖는 데이터를 적어도 하나의 CPU(423-1 및/또는 423-2) 및/또는 버퍼 컨트롤러(427)로 전송할 수 있다. 실시 예에 따라, 제4인터페이스(421)는 신호 및/또는 데이터를 송수신할 수 있는 송수신기를 포함할 수 있다.
- [0099] 제4인터페이스(421)의 구조와 작동은 제3인터페이스(350)의 구조와 작동에 적합하게 구현될 수 있다. 예컨대, 제4인터페이스(421)는 SATA 인터페이스, SATAe 인터페이스, SAS 인터페이스, PCIe 인터페이스, NVMe 인터페이스, AHCI 인터페이스, MMC 인터페이스, NAND-타입 플래시 메모리 인터페이스, 또는 NOR-타입 플래시 메모리 인터페이스로 구현될 수 있으나 이에 한정되는 것은 아니다.
- [0100] 실시 예에 따라, 도 4와 도 7을 참조하면, 제4인터페이스(421)는 하나의 채널(또는 하나의 메인 채널) 또는 하나의 웨이에 접속될 수 있다. 제1단자들(PT1)은 제1메인 채널(CHA)과 제4인터페이스(421)를 접속하기 위한 단자들이고, 제3단자들(PT3)은 제1휘발성 메모리 장치(413-1)와 버퍼 컨트롤러(427)를 접속하기 위한 단자들이고, 제2단자들(PT2)은 제5인터페이스(429)와 하나 또는 그 이상의 채널(CH11~CH1m)을 접속하기 위한 단자들이다. 상기 단자들은 핀들(pins) 또는 패드들(pads)을 의미할 수 있다.
- [0101] 하나 또는 그 이상의 CPU(423-1 및/또는 423-3)는, 버스 구조(420)를 통해, 제4인터페이스(421), 내부 메모리(425), 버퍼 컨트롤러(427), 및 제5인터페이스(429)를 제어할 수 있다. 각 CPU(423-1 및/또는 423-3)는 하나 또는 그 이상의 코어들을 포함할 수 있다.
- [0102] 예컨대, 제1CPU(423-1)는 제1메인 채널(CHA)과 제4인터페이스(421)를 통해 제3인터페이스(350)와 데이터를 주거나 받을 수 있고, 제2CPU(423-2)는 제5인터페이스(429)를 통해 각 채널(CH11~CH1m)에 접속된 불휘발성 메모리 장치들(NAND) 중에서 적어도 하나와 데이터를 주거나 받을 수 있다. 실시 예에 따라, 제1CPU(423-1)와 제2CPU(423-2)는 멀티-CPU를 구성할 수 있다. 실시 예에 따라, 제1CPU(423-1)는 제2CPU(423-2)를 제어할 수 있다.
- [0103] 내부 메모리(425)는 제1스케일-아웃 컨트롤러(415-1)의 작동에 필요한 데이터 또는 제1스케일-아웃 컨트롤러(415-1)에 의해 수행된 데이터 처리 작동(예컨대, 라이트 작동 또는 리드 작동)에 의해 생성된 데이터를 저장할 수 있다. 예컨대, 내부 메모리(425)는 CPU(423-1 및/또는 423-2)에 의해 실행될 수 있는 제2FTL(flash translation layer) 코드를 저장할 수 있다. 예컨대, 데이터 저장 장치(300)가 부팅될 때, 제2FTL 코드는 메모리 클러스터(400)에 포함된 불휘발성 메모리로부터 내부 메모리(425)로 로드될 수 있다. 실시 예들에 따라, 내부 메모리(425)는 RAM, DRAM, SRAM, 버퍼, 버퍼 메모리, 캐시, 또는 TCM으로 구현될 수 있으나 이에 한정되는 것은 아니다.
- [0104] 버퍼 컨트롤러(427)는, 제1CPU(423-1) 또는 제2CPU(423-2)의 제어에 따라, 제1휘발성 메모리 장치(413-1)에 데이터를 라이트하거나 제1휘발성 메모리 장치(413-1)로부터 데이터를 리드할 수 있다. 버퍼 컨트롤러(427)는 제1휘발성 메모리 장치(413-1)에 대한 라이트 작동과 리드 작동을 제어할 수 있는 컨트롤러 또는 버퍼 매니저로 불릴 수 있다.
- [0105] 제5인터페이스(429)는, 제1CPU(423-1) 또는 제2CPU(423-2)의 제어에 따라, 복수의 채널들(CH11~CH1m) 중에서 해당하는 채널을 통해 불휘발성 메모리 장치들(NAND)에 대한 데이터 처리 작동들을 제어할 수 있다.
- [0106] 도 4와 도 7을 참조하면, 복수의 채널들 및/또는 복수의 웨이들은 제5인터페이스(429)에 접속될 수 있다. 실시 예들에 따라, 제5인터페이스(429)는 SATA 인터페이스, SATAe 인터페이스, SAS 인터페이스, PCIe 인터페이스, NVMe 인터페이스, AHCI 인터페이스, MMC 인터페이스, NAND-타입 플래시 메모리 인터페이스, 또는 NOR-타입 플래시 메모리 인터페이스로 구현될 수 있으나 이에 한정되는 것은 아니다. 예컨대, 제5인터페이스(429)는 채널들(CH11~CH1m) 각각에 대응되는 메모리 컨트롤러들(429-1~429-m) 각각을 포함할 수 있다. 예컨대, 불휘발성 메모리 장치들(NAND) 각각이 NAND-타입 플래시 메모리일 때, 메모리 컨트롤러들(429-1~429-m) 각각은 플래시 메모리 컨트롤러일 수 있다.
- [0107] 제1휘발성 메모리 장치(413-1)는 제1스케일-아웃 저장 장치(410-1)에 포함된 불휘발성 메모리 장치들(NAND)에 대한 논리 어드레스-물리 어드레스 변환을 위한 매핑 테이블을 저장하는 제1영역과, 캐시(cache)의 기능을 수행할 수 있는 제2영역을 포함할 수 있으나 이에 한정되는 것은 아니다. 예컨대, CPU(423-1 및/또는 423-2)에 의해

실행되는 제2FTL 코드는 상기 제1영역에 저장된 매핑 테이블을 이용하여 논리 어드레스-물리 어드레스 변환을 수행할 수 있다.

- [0108] 도 3을 참조하여 설명된 불휘발성 메모리 장치들(NAND) 중에서 적어도 하나가 제1스케일-아웃 장치(410-1A)로 대체되면, 메모리 클러스터(400)에 대한 용량의 확장성(scalability)은 증가할 수 있다.
- [0109] 도 5는 도 4에 도시된 제1스케일-아웃 장치의 반도체 패키지의 실시 예이다. 도 3부터 도 5를 참조하면, 제1스케일-아웃 장치(410-1A)는 반도체 패키지로 구현될 수 있다. 즉, 제1스케일-아웃 장치(410-1A)는 반도체 기판(substrate; 415)의 위(on)에 각각 부착(attached or mounted)된 제1스케일-아웃 컨트롤러(411-1)와 제1휘발성 메모리 장치(413-1)를 포함할 수 있다.
- [0110] 예컨대, 제1스케일-아웃 컨트롤러(411-1)는 플립-칩(flip-chip) 구조로 반도체 기판(415)에 부착될 수 있다. 또한, 제1스케일-아웃 컨트롤러(411-1)가 접착 물질(adhesive material)을 통해 반도체 기판(415)에 부착될 때, 제1스케일-아웃 컨트롤러(411-1)는 본딩 와이어들(bonding wires)을 통해 반도체 기판(415)과 전기적인 신호들을 주거나 받을 수 있다.
- [0111] 예컨대, 제1휘발성 메모리 장치(413-1)는 플립-칩 구조로 반도체 기판(415)에 부착될 수 있다. 또한, 제1휘발성 메모리 장치(413-1)가 접착 물질을 통해 반도체 기판(415)에 부착될 때, 제1휘발성 메모리 장치(413-1)는 본딩 와이어들을 통해 반도체 기판(415)과 전기적인 신호들을 주거나 받을 수 있다. 즉, 제1스케일-아웃 컨트롤러(411-1)는 반도체 기판(415)을 통해 제1휘발성 메모리 장치(413-1)와 명령 및/또는 데이터를 주거나 받을 수 있다. 각 스케일 아웃 장치(410-1A~410-4A)에 대한 반도체 패키지는 실질적으로 동일 또는 유사하다.
- [0112] 도 6은 도 1에 도시된 컨트롤러와 제1스케일-아웃 저장 장치를 포함하는 반도체 패키지의 실시 예이다. 도 6에 도시된 반도체 패키지(300-1A)는 ePOP (embedded package on package)로 구현될 수 있으나 반도체 패키지(300-1A)의 종류가 이에 한정되는 것은 아니다.
- [0113] 예컨대 반도체 패키지(300-1A)는 반도체 기판(415-1), 반도체 기판(415-1)의 위(over)에 배치된 제1패키지(PKG1), 및 제1패키지(PKG1)의 위(over)에 배치된 제2패키지(PKG2)를 포함할 수 있다. 제1패키지(PKG1)는 제1범프들(415-2)을 통해 반도체 기판(415-1)에 부착될 수 있고, 제1범프들(415-2)을 통해 반도체 기판(415-1)과 전기적인 신호들을 주거나 받을 수 있다. 제2패키지(PKG2)는 제2범프들(415-3)을 통해 제1패키지(PKG1)에 부착될 수 있고, 제2범프들(415-3)을 통해 반도체 기판(415-1)과 전기적인 신호들을 주거나 받을 수 있다. 제1패키지(PKG1)는 컨트롤러(310)를 포함할 수 있다. 제2패키지(PKG2)는 제1스케일-아웃 컨트롤러(411-1), 제1휘발성 메모리 장치(413-1), 및 적어도 하나의 NAND-타입 플래시 메모리(470)를 포함할 수 있다. 예컨대, 각 스케일-아웃 저장 장치(410-2, ..., 410-4)가 각 패키지로 구현되고, 반도체 기판(415-1)의 위(on)에 부착될 때, 컨트롤러(310)는 반도체 기판(415-1)을 통해 각 스케일-아웃 저장 장치(410-2, ..., 410-4)를 제어할 수 있다.
- [0114] 도 2, 도 3, 및 도 7을 다시 참조하면, 데이터 저장 장치(300)는 컨트롤러 (310)와 메모리 클러스터(400)를 포함하고, 메모리 클러스터(400)에 포함된 스케일-아웃 저장 장치들이 복수 개이고, 상기 스케일-아웃 저장 장치들 각각의 구조는 도 3에 도시된 제1스케일-아웃 저장 장치(410-1)의 구조와 실질적으로 동일하게 구현될 수 있다.
- [0115] 도 8a는 도 1에 도시된 데이터 처리 시스템의 작동을 설명하기 위한 블록도이다. 도 1, 도 2, 도 7, 및 도 8a를 참조하면, 데이터 처리 시스템(100A)은 호스트(200)와 데이터 저장 장치(300A)를 포함하고, 데이터 저장 장치(300A)는 컨트롤러(310), 버퍼(360), 및 메모리 클러스터(400)를 포함할 수 있다.
- [0116] 메모리 클러스터(400)는 제1클러스터(410)와 제2클러스터(430)를 포함하고, 제1클러스터(410)는 제1스케일-아웃 저장 장치(410-1)를 포함하고, 제2클러스터 (430)는 제2스케일-아웃 저장 장치(430-1)를 포함하고, 제1스케일-아웃 컨트롤러 (411-1)의 구조와 제2스케일-아웃 컨트롤러(413-1)의 구조는 실질적으로 동일하다고 가정한다.
- [0117] 비록, 도 8a에서는 제1스케일-아웃 저장 장치(410-1)를 포함하는 제1클러스터(410)와 제2스케일-아웃 저장 장치(430-1)를 포함하는 제2클러스터(430)가 예시적으로 도시되어 있으나, 본 발명의 기술적 사상은 각 클러스터(410과 430)에 포함된 스케일-아웃 저장 장치의 개수에 한정되는 것은 아니다. 따라서, 제1클러스터 (410)는 적어도 하나의 스케일-아웃 저장 장치를 포함할 수 있고, 제2클러스터 (430)는 적어도 하나의 스케일-아웃 저장 장치를 포함할 수 있다.
- [0118] 컨트롤러(310)의 CPU(330)는 수학적 1에 따라 논리 어드레스를 분배한다고 가정한다. 예컨대, 논리 어드레스는 논리 블록 어드레스(logical block address(LBA)) 또는 논리 페이지 넘버(logical page number(LPN))일 수 있



으나 이에 한정되는 것은 아니다.

- [0119] [수학식 1]
- [0120]  $GLPN=iLPN*X+NUM$
- [0121] 여기서, GLPN은 글로벌 논리 페이지 넘버(global logical page number)를 나타내고, iLPN은 i-번째 클러스터의 로컬(local) 논리 페이지 넘버를 나타내고, X는 데이터 저장 장치(300A)에 포함된 클러스터(또는 멀티 칩 셋)의 개수를 나타내고, NUM은 클러스터의 넘버를 나타낸다. 도 8a와 도 8b에서, X는 2개이고, 제1클러스터(410) 또는 제1스케일-아웃 저장 장치(410-1)의 NUM은 0이고, 제2클러스터 (430) 또는 제2스케일-아웃 저장 장치(430-1)의 NUM은 1이라고 가정한다.
- [0122] 또한, 테이블(361)은 호스트(200)로부터 출력된 논리 어드레스를 클러스터들 (410과 430) 중에서 어느 클러스터로 전송할지에 대한 전송 매핑 정보(이하, "제1매핑 정보"라 한다)를 포함하고 있다고 가정한다.
- [0123] 실시 예에 따라, 도 8a와 도 12를 참조하면, 호스트(200)의 CPU(220)가 글로벌 논리 어드레스(LA0=GLPN0)를 출력하면, 컨트롤러(310)의 CPU(330) 또는 CPU (330)에 의해 실행되는 제1FTL은, 수학식 1 및/또는 제1매핑 정보를 이용하여, 글로벌 논리 어드레스(LA0=GLPN0)에 해당하는 제1클러스터(410)로 글로벌 논리 어드레스(LA0=GLPN0)를 전송할 수 있다.
- [0124] 제1스케일-아웃 컨트롤러(411-1)의 CPU(423-1) 또는 CPU(423-1)에 의해 실행되는 제2FTL은, 휘발성 메모리(413-1)에 저장된 논리 어드레스-물리 어드레스 매핑 테이블(TB2)을 이용하여 글로벌 논리 어드레스(LA0=GLPN0)에 대응되는 로컬 논리 어드레스(1LPN0)를 물리 어드레스(1PPN0)로 변환할 수 있다. 예컨대, 로컬 논리 어드레스(1LPN0)는 논리 어드레스-물리 어드레스 매핑 테이블(TB2)의 엔트리(entry)를 의미할 수 있다.
- [0125] 실시 예에 따라, 도 8a와 도 12를 참조하면, 호스트(200)의 CPU(220)가 로컬 논리 어드레스(LA0=1LPN0)를 출력하면, 컨트롤러(310)의 CPU(330) 또는 CPU(330)에 의해 실행되는 제1FTL은, 수학식 1 및/또는 제1매핑 정보를 이용하여, 로컬 논리 어드레스(LA0=1LPN0)에 해당하는 제1클러스터(410)로 로컬 논리 어드레스 (LA0=1LPN0)를 전송할 수 있다.
- [0126] 제1스케일-아웃 컨트롤러(411-1)의 CPU(423-1) 또는 CPU(423-1)에 의해 실행되는 제2FTL은, 휘발성 메모리(413-1)에 저장된 논리 어드레스-물리 어드레스 매핑 테이블(TB2)을 이용하여 로컬 논리 어드레스(1LPN0)를 물리 어드레스(1PPN0)로 변환할 수 있다.
- [0127] 실시 예에 따라, 도 8a와 도 12를 참조하면, 호스트(200)의 CPU(220)가 글로벌 논리 어드레스(LA100=GLPN201)를 출력하면, 컨트롤러(310)의 CPU(330) 또는 CPU (330)에 의해 실행되는 제1FTL은, 수학식 1 및/또는 제1매핑 정보를 이용하여, 글로벌 논리 어드레스(LA100=GLPN201)에 해당하는 제2클러스터(430)로 글로벌 논리 어드레스 (LA100=GLPN201)를 전송할 수 있다.
- [0128] 제1스케일-아웃 컨트롤러(411-1)의 구조와 제2스케일-아웃 컨트롤러(431-1)가 실질적으로 동일할 때, 제2스케일-아웃 컨트롤러(431-1)의 CPU(423-1) 또는 CPU(423-1)에 의해 실행되는 제3FTL은, 휘발성 메모리(433-1)에 저장된 논리 어드레스-물리 어드레스 매핑 테이블(TB3)을 이용하여 글로벌 논리 어드레스 (LA100=GLPN201)에 대응되는 로컬 논리 어드레스(2LPN100)를 물리 어드레스 (2PPN100)로 변환할 수 있다. 예컨대, 로컬 논리 어드레스(2LPN100)는 논리 어드레스-물리 어드레스 매핑 테이블(TB3)의 엔트리를 의미할 수 있다.
- [0129] 실시 예에 따라, 도 8a와 도 12를 참조하면, 호스트(200)의 CPU(220)가 로컬 논리 어드레스(LA100=2LPN100)를 출력하면, 컨트롤러(310)의 CPU(330) 또는 CPU (330)에 의해 실행되는 제1FTL은, 수학식 1 및/또는 제1매핑 정보를 이용하여, 로컬 논리 어드레스(LA100=2LPN100)에 해당하는 제2클러스터(430)로 로컬 논리 어드레스 (LA100=2LPN100)를 전송할 수 있다.
- [0130] 제2스케일-아웃 컨트롤러(431-1)의 CPU(423-1) 또는 CPU(423-1)에 의해 실행되는 제3FTL은, 휘발성 메모리(433-1)에 저장된 논리 어드레스-물리 어드레스 매핑 테이블(TB3)을 이용하여 로컬 논리 어드레스(2LPN100)를 물리 어드레스(2PPN100)로 변환할 수 있다.
- [0131] 도 12에 도시된 바와 같이, 제1클러스터(410)에 대한 글로벌 논리 어드레스는 짝수로 할당되고, 제2클러스터 (430)에 대한 글로벌 논리 어드레스는 홀수로 할당될 수 있다.
- [0132] 도 9는 본 발명의 일 실시 예에 따른 스왑 매핑 테이블 업데이트 작동을 설명하기 위한 플로우이다. 도 1, 도 2, 도 8a, 및 도 9를 참조하면, 컨트롤러(310), 예컨대 CPU(330 및/또는 331)는 제1스케일-아웃 컨트롤러(즉,

제1코어; 411-1)로 제1요청(REQ1)을 전송할 수 있다(S101-1).

- [0133] 예컨대, 제1코어(411-1)는, 제1요청(REQ1)에 응답하여, 제1코어(411-1)에 의해 관리되는 적어도 하나의 NAND-타입 플래시 메모리 장치(NAND)의 메모리 셀 어레이에 포함된 블록들 각각의 웨어-아웃 카운트와 기준 웨어-아웃 카운트를 비교하고, 비교 결과에 따라 상기 기준 웨어-아웃 카운트보다 높은 카운트를 갖는 블록(들)을 서치하고, 서치 결과에 따라 제1리스트(LIST1)를 생성할 수 있다(S103-1). 제1코어(411-1)는 제1리스트(LIST1)를 컨트롤러(310)로 전송할 수 있다(S105-1). 제1리스트(LIST1)는, 제1코어(411-1)의 제어에 따라, 휘발성 메모리(413-1) 및/또는 제1코어(411-1)에 접속된 적어도 하나의 NAND-타입 플래시 메모리 장치(NAND)에 저장될 수 있다.
- [0134] 예컨대, 논리 어드레스(LA0=GLPN0, 또는 LA0=ILPN0)에 대응되는 제1저장 영역, 예컨대, 물리 어드레스(1PPN0)에 대응되는 페이지의 웨어-아웃 카운트가 기준 웨어-아웃 카운트보다 높을 때, 제1코어(411-1)는 제1저장 영역의 논리 어드레스(LA0=GLPN0, 또는 LA0=ILPN0)를 포함하는 제1리스트(LIST1)를 생성하고(S103-1), 제1리스트(LIST1)를 호스트(200)로 전송할 수 있다(S105-1). 상술한 바와 같이, 제1저장 영역에 대한 웨어-아웃 카운트는 이레이즈 카운트, 프로그램 카운트, 리드 카운트, 배드 블록 발생 횟수, 또는 리드 에러 발생 횟수일 수 있으나 이에 한정되는것은 아니다.
- [0135] 또한, 제1코어(411-1)는, 제1요청(REQ1)에 응답하여, 제1저장 영역이 핫 데이터와 콜드 데이터 중에서 어느 하나를 저장하는 영역인지를 판단하고, 판단 결과에 따라 제1저장 영역의 논리 어드레스(LA0=GLPN0, 또는 LA0=ILPN0)를 포함하는 제1리스트(LIST1)를 생성하고(S103-1), 제1리스트(LIST1)를 호스트(200)로 전송할 수 있다(S105-1).
- [0136] 컨트롤러(310), 예컨대 CPU(330 및/또는 331)는 제2스케일-아웃 컨트롤러(즉, 제2코어; 431-1)로 제2요청(REQ2)을 전송할 수 있다(S101-2).
- [0137] 예컨대, 제2코어(431-1)는, 제2요청(REQ2)에 응답하여, 제2코어(431-1)에 의해 관리되는 적어도 하나의 NAND-타입 플래시 메모리 장치(NAND)의 메모리 셀 어레이에 포함된 블록들 각각의 웨어-아웃 카운트와 기준 웨어-아웃 카운트를 비교하고, 비교 결과에 따라 상기 기준 웨어-아웃 카운트보다 낮은 카운트를 갖는 블록(들)을 서치하고, 서치 결과에 따라 제2리스트(LIST2)를 생성할 수 있다(S103-2). 제2코어(431-1)는 제2리스트(LIST2)를 컨트롤러(310)로 전송할 수 있다(S105-2).
- [0138] 제2리스트(LIST2)는, 제2코어(431-1)의 제어에 따라, 휘발성 메모리(433-1) 및/또는 제2코어(431-1)에 접속된 적어도 하나의 NAND-타입 플래시 메모리 장치 (NAND)에 저장될 수 있다.
- [0139] 예컨대, 논리 어드레스(LA100=GLPN201, 또는 LA1000=2LPN100)에 대응되는 제2저장 영역, 예컨대, 물리 어드레스(2PPN100)에 대응되는 페이지의 웨어-아웃 카운트가 기준 웨어-아웃 카운트보다 낮을 때, 제2코어(431-1)는 제2저장 영역의 논리 어드레스(LA100=GLPN201, 또는 LA100=2LPN100)를 포함하는 제2리스트(LIST2)를 생성하고(S103-2), 제2리스트(LIST2)를 호스트(200)로 전송할 수 있다(S105-2). 상술한 바와 같이, 제2저장 영역에 대한 웨어-아웃 카운트는 이레이즈 카운트, 프로그램 카운트, 리드 카운트, 배드 블록 발생 횟수, 또는 리드 에러 발생 횟수일 수 있으나 이에 한정되는 것은 아니다.
- [0140] 또한, 제2코어(431-1)는, 제2요청(REQ2)에 응답하여, 제2저장 영역이 핫 데이터와 콜드 데이터 중에서 다른 하나를 저장하는 영역인지를 판단하고, 판단 결과에 따라 제2저장 영역의 논리 어드레스(LA100=GLPN201, 또는 LA100=2LPN100)를 포함하는 제2리스트(LIST2)를 생성하고(S103-2), 제2리스트(LIST2)를 호스트(200)로 전송할 수 있다(S105-2).
- [0141] 단계들(S101-1과 S101-2, S103-1과 S103-2, 및 S105-1과 S105-2)은 동시에 또는 병렬로 수행될 수 있다. 예컨대, 각 단계(S101-1, S103-1, 및 S105-1)는 각 단계(S101-2, S103-2, 및 S105-2)보다 먼저 수행될 수도 있고 나중에 수행될 수 있다.
- [0142] 컨트롤러(310)의 CPU(330)는, 제1리스트(LIST1)와 제2리스트(LIST2)를 수신하고, 이들에 기초하여, 도 11에 도시된 바와 같은 스왑 매핑 테이블(TB1)을 생성하고, 스왑 매핑 테이블(TB1)을 휘발성 메모리(360)에 저장할 수 있다(S107). 예컨대, 스왑 매핑 테이블(TB1)은 테이블(TABLE1)의 적어도 일부일 수 있다.
- [0143] 도 11은 논리 어드레스 스왑 작동을 위한 스왑 매핑 테이블의 일 실시 예이다. 실시 예에 따라, 스왑 매핑 테이블(TB1)은 소스 글로벌 논리 어드레스 (GLPN0=0)와 목적지 글로벌 논리 어드레스(GLPN201=201)를 포함할 수 있다. 다른 실시 예에 따라, 스왑 매핑 테이블(TB1)은 소스 로컬 논리 어드레스(ILPN0)와 목적지 로컬 논리 어드레스(2LPN100)를 포함할 수 있다.

- [0144] 도 10은 본 발명의 다른 실시 예에 따른 스왑 매핑 테이블 업데이트 작동을 설명하기 위한 플로우이다.
- [0145] 도 9에서 각 코어(411-1과 431-1)는, 컨트롤러(310)로부터 출력된 각 요청 (REQ1과 REQ2)에 응답하여, 각 리스트(LIST1과 LIST2)를 생성하고, 각 리스트 (LIST1과 LIST2)를 호스트(200)로 전송할 수 있다. 그러나, 도 10에서는 컨트롤러 (310)로부터 요청이 없더라도, 각 코어(411-1과 431-1)는 주도적으로 각 리스트 (LIST1과 LIST2)를 생성하고, 각 리스트(LIST1과 LIST2)를 호스트(200)로 전송할 수 있다.
- [0146] 예컨대, 제1코어(411-1)는 제1코어(411-1)에 의해 관리되는 적어도 하나의 NAND-타입 플래시 메모리 장치(NAND)의 메모리 셀 어레이에 포함된 블록들 각각의 웨어-아웃 카운트와 기준 웨어-아웃 카운트를 비교하고, 비교 결과에 따라 상기 기준 웨어-아웃 카운트보다 높은 카운트를 갖는 블록(들)을 서치하고, 서치 결과에 따라 제1리스트(LIST1)를 생성할 수 있다(S103-1A). 제1코어(411-1)는 제1리스트 (LIST1)를 컨트롤러(310)로 전송할 수 있다(S105-1A). 제1리스트(LIST1)는, 제1코어(411-1)의 제어에 따라, 휘발성 메모리(413-1) 및/또는 제1코어(411-1)에 접속된 적어도 하나의 NAND-타입 플래시 메모리 장치(NAND)에 저장될 수 있다.
- [0147] 예컨대, 논리 어드레스(LA0=GLPN0, 또는 LA0=ILPN0)에 대응되는 제1저장 영역, 예컨대, 물리 어드레스(1PPN0)에 대응되는 페이지의 웨어-아웃 카운트가 기준 웨어-아웃 카운트보다 높을 때, 제1코어(411-1)는 제1저장 영역의 논리 어드레스 (LA0=GLPN0, 또는 LA0=ILPN0)를 포함하는 제1리스트(LIST1)를 생성하고(S103-1A), 제1리스트 (LIST1)를 호스트(200)로 전송할 수 있다(S105-1A).
- [0148] 또한, 제1코어(411-1)는 제1저장 영역이 핫 데이터와 콜드 데이터 중에서 어느 하나를 저장하는 영역인지를 판단하고, 판단 결과에 따라 제1저장 영역의 논리 어드레스(LA0=GLPN0, 또는 LA0=ILPN0)를 포함하는 제1리스트 (LIST1)를 생성하고 (S103-1A), 제1리스트(LIST1)를 호스트(200)로 전송할 수 있다(S105-1A).
- [0149] 예컨대, 제2코어(431-1)는 제2코어(431-1)에 의해 관리되는 적어도 하나의 NAND-타입 플래시 메모리 장치(NAND)의 메모리 셀 어레이에 포함된 블록들 각각의 웨어-아웃 카운트와 기준 웨어-아웃 카운트를 비교하고, 비교 결과에 따라 상기 기준 웨어-아웃 카운트보다 낮은 카운트를 갖는 블록(들)을 서치하고, 서치 결과에 따라 제2리스트(LIST2)를 생성할 수 있다(S103-2A). 제2코어(431-1)는 제2리스트 (LIST2)를 컨트롤러(310)로 전송할 수 있다(S105-2A).
- [0150] 제2리스트(LIST2)는, 제2코어(431-1)의 제어에 따라, 휘발성 메모리(433-1) 및/또는 제2코어(431-1)에 접속된 적어도 하나의 NAND-타입 플래시 메모리 장치 (NAND)에 저장될 수 있다.
- [0151] 예컨대, 논리 어드레스(LA100=GLPN201, 또는 LA1000=2LPN100)에 대응되는 제2저장 영역, 예컨대, 물리 어드레스 (2PPN100)에 대응되는 페이지의 웨어-아웃 카운트가 기준 웨어-아웃 카운트보다 낮을 때, 제2코어(431-1)는 제2저장 영역의 논리 어드레스(LA100=GLPN201, 또는 LA100=2LPN100)를 포함하는 제2리스트(LIST2)를 생성하고 (S103-2A), 제2리스트(LIST2)를 호스트(200)로 전송할 수 있다(S105-2A).
- [0152] 또한, 제2코어(431-1)는 제2저장 영역이 핫 데이터와 콜드 데이터 중에서 다른 하나를 저장하는 영역인지를 판단하고, 판단 결과에 따라 제2저장 영역의 논리 어드레스(LA100=GLPN201, 또는 LA100=2LPN100)를 포함하는 제2리스트(LIST2)를 생성하고(S103-2A), 제2리스트(LIST2)를 호스트(200)로 전송할 수 있다(S105-2A).
- [0153] 단계들(S103-1A과 S103-2A, 및 S105-1A과 S105-2A)은 동시에 또는 병렬로 수행될 수 있다. 예컨대, 각 단계 (S103-1A, 및 S105-1A)는 각 단계(S103-2A, 및 S105-2A)보다 먼저 수행될 수도 있고 나중에 수행될 수 있다.
- [0154] 도 8b는 논리 어드레스 스왑 작동을 수행하는 도 1에 도시된 데이터 처리 시스템의 작동을 설명하기 위한 블록도이다.
- [0155] 도 8b부터 도 12를 참조하면, 호스트(200)의 CPU(220)가 글로벌 논리 어드레스(LA0=GLPN0)를 출력하면, 컨트롤러(310)의 CPU(330) 또는 CPU (330)에 의해 실행되는 제1FTL은, 도 11에 도시된 스왑 매핑 테이블(TB1)을 이용하여, 글로벌 논리 어드레스(LA0=GLPN0=0)와 스왑된 글로벌 논리 어드레스(LA100=GLPN201=201)에 해당하는 제2클러스터(410)로 스왑된 글로벌 논리 어드레스(LA100=GLPN201=201)를 전송할 수 있다.
- [0156] 제2스케일-아웃 컨트롤러(431-1)의 CPU(423-1) 또는 CPU(423-1)에 의해 실행되는 제3FTL은, 휘발성 메모리 (433-1)에 저장된 논리 어드레스-물리 어드레스 매핑 테이블(TB3)을 이용하여 스왑된 글로벌 논리 어드레스 (LA100=GLPN201=201)에 대응되는 로컬 논리 어드레스(2LPN100)를 물리 어드레스(2PPN100)로 변환할 수 있다.
- [0157] 도 12는 논리 어드레스, 물리 어드레스, 및 글로벌 논리 어드레스의 관계를 설명하는 개념도이고, 도 13은 도 8b에 도시된 데이터 저장 장치의 데이터 스왑 작동을 설명하는 개념도이다.

- [0158] 도 8a부터 도 13을 참조하면, 제1메모리 맵(NAND1)은 제1코어(411-1)에 의해 관리되는 적어도 하나의 NAND-타입 플래시 메모리 장치(NAND)에 대한 전체 메모리 맵(total memory map)을 나타내고, 제2메모리 맵(NAND2)은 제2코어(431-1)에 의해 관리되는 적어도 하나의 NAND-타입 플래시 메모리 장치(NAND)에 대한 전체 메모리 맵을 나타낸다고 가정한다.
- [0159] 제1메모리 맵(NAND1)에서, 제1메모리 영역(10)은 제1오버프로비저닝 영역(12)과 제1사용자 데이터 영역(14)을 포함하고, 제1오버프로비저닝 영역(12)은 블록들(Block#0~Block#3)을 포함하고, 제1사용자 데이터 영역(14)은 블록들(Block#0~Block#3)을 포함하고, 블록들(Block#0~Block#3 및 Block#0~Block#3) 각각은 페이지들(Page#0~Page127)을 포함한다고 가정한다.
- [0160] 또한, 제2메모리 맵(NAND2)에서, 제2메모리 영역(20)은 제2오버프로비저닝 영역(22)과 제2사용자 데이터 영역(24)을 포함하고, 제2오버프로비저닝 영역(22)은 블록들(Block#0~Block#3)을 포함하고, 제2사용자 데이터 영역(24)은 블록들(Block#0~Block#3)을 포함하고, 블록들(Block#0~Block#3) 각각은 페이지들(OPage#0~OPage127)을 포함하고, 블록들(Block#0~Block#3) 각각은 페이지들(Page#0~Page127)을 포함한다고 가정한다.
- [0161] 도 12를 다시 참조하면, 매핑 테이블(TB2)의 제1엔트리들(PA101)은 제1메모리 영역(10)의 제1사용자 데이터 영역(14)에 대한 논리 어드레스-물리 어드레스 매핑을 위한 엔트리들이고, 매핑 테이블(TB2)의 제2엔트리들(PA103)은 제1메모리 영역(10)의 제1오버프로비저닝 영역(12)에 대한 논리 어드레스-물리 어드레스 매핑을 위한 엔트리들이라고 가정한다.
- [0162] 또한, 매핑 테이블(TB3)의 제1엔트리들(PA201)은 제2메모리 영역(20)의 제2사용자 데이터 영역(24)에 대한 논리 어드레스-물리 어드레스 매핑을 위한 엔트리들이고, 매핑 테이블(TB3)의 제2엔트리들(PA203)은 제2메모리 영역(20)의 제2오버프로비저닝 영역(22)에 대한 논리 어드레스-물리 어드레스 매핑을 위한 엔트리들이라고 가정한다. 논리 어드레스(LA0=GLPN0 또는 LA0=1LPN0)는 제1메모리 영역(10)의 페이지(16)에 대응되는 어드레스이고, 논리 어드레스(LA100=GLPN201 또는 LA100=2LPN100)는 제2메모리 영역(20)의 페이지(26)에 대응되는 어드레스라고 가정한다.
- [0163] 도 8b, 도 12, 및 도 13을 참조하면, 논리 어드레스(LA0=GLPN0 또는 LA0=1LPN0)가 프로그램(또는 라이트) 작동을 위한 논리 어드레스일 때, 컨트롤러(310)는 호스트(200)로부터 출력된 프로그램 데이터를 스왑된 논리 어드레스(LA100=GLPN201 또는 LA100=2LPN100)에 대응하는 물리 어드레스(2PPN100)에 의해 지정된 제2저장 영역, 예컨대, 제2메모리 영역(20)의 제2사용자 데이터 영역(24)의 제1블록(Block#0)의 제2페이지(Page#1; 26)에 프로그램할 수 있다.
- [0164] 즉, 컨트롤러(310)는 프로그램 데이터를 제1메모리 영역(10)의 제1사용자 데이터 영역(14)의 제1블록(Block#0)의 제1페이지(Page#0; 16) 대신에 제2메모리 영역(20)의 제2사용자 데이터 영역(24)의 제1블록(Block#0)의 제2페이지(Page#1; 26)에 프로그램할 수 있다.
- [0165] 경우(CASE1)에 도시된 바와 같이, 컨트롤러(310)는, 제1코어(411-1)과 제2코어(431-1)를 이용하여, 제1메모리 영역(10)의 제1사용자 데이터 영역(14)의 제1블록(Block#0)의 제1페이지(Page#0; 16)에 저장된 제1데이터와 제2메모리 영역(20)의 제2사용자 데이터 영역(24)의 제1블록(Block#0)의 제2페이지(Page#1; 26)에 저장된 제2데이터를 스왑할 수 있다. 따라서, 제2데이터는 제1페이지(Page#0; 16)에 저장되고, 제1데이터는 제2페이지(Page#1; 26)에 저장될 수 있다.
- [0166] 논리 어드레스(LA0=GLPN0 또는 LA0=1LPN0)가 제1데이터를 리드하기 위한 논리 어드레스일 때, 컨트롤러(310)는 스왑된 논리 어드레스(LA100=GLPN201 또는 LA100=2LPN100)를 제2코어(431-1)로 전송할 수 있다. 따라서, 제2코어(431-1)는 스왑된 논리 어드레스(LA100=GLPN201 또는 LA100=2LPN100)에 대응되는 제2페이지(Page#1; 26)에 저장된 제1데이터를 컨트롤러(310)로 전송하고, 컨트롤러(310)는 상기 제1데이터를 호스트(200)로 전송할 수 있다.
- [0167] 즉, 본 발명의 실시 예에 따른 데이터 저장 장치(300 또는 300A)는, 도 11의 스왑 매핑 테이블(TB1)을 이용하여, 웨어 레벨링이 필요한 데이터에 대한 논리 어드레스(예컨대, LA0)를 다른 논리 어드레스(예컨대, LA100)로 변경하거나 스왑할 수 있다. 따라서, 데이터 저장 장치(300 또는 300A)는 클러스터들(410과 430) 사이에서 또는 스케일-아웃 저장 장치들(410-1과 430-1) 사이에서 글로벌 웨어 레벨링을 지원할 수 있다.
- [0168] 도 8a와 도 8b에 도시된 바와 같이, 컨트롤러(310)는 제1스케일-아웃 저장 장치(410-1)의 웨어 레벨링 소스 블



록(예컨대, 웨어-카운트가 높은 블록) 또는 상기 웨어 레벨링 소스 블록의 저장 영역(예컨대, 16)에 대한 논리 어드레스(예컨대, LA0)와 제2스케일-아웃 저장 장치(430-1)의 웨어 레벨링 목적지 블록(예컨대, 웨어-카운트가 낮은 블록) 또는 상기 웨어 레벨링 목적지 블록의 저장 영역(예컨대, 26)에 대한 논리 어드레스(예컨대, LA100)를 스왑할 수 있다. 따라서, 데이터 저장 장치(300 또는 300A)는 클러스터들(410과 430) 또는 스케일-아웃 저장 장치들(410-1과 430-1)에 대한 웨어 레벨링(즉, 글로벌 웨어 레벨링)을 수행할 수 있다.

[0169] 도 8b, 도 12, 및 도 13에 도시된 경우2(CASE2)는 제1메모리 영역(10)의 제1사용자 데이터 영역(14)의 제1저장 영역(16)의 논리 어드레스(LA0=GLPN0 또는 LA0=1LPN0)를 제2메모리 영역(20)의 제2오버프로비저닝 영역(20)의 제3저장 영역(36)의 논리 어드레스(LA101=GLPN203 또는 LA101=2LPN101)로 매핑, 변경, 또는 스왑하는 방법을 설명하기 위한 개념도이다.

[0170] 컨트롤러(310)는 제1스케일-아웃 저장 장치(410-1)의 웨어 레벨링 소스 블록(예컨대, 웨어-카운트가 높은 블록) 또는 상기 웨어 레벨링 소스 블록의 저장 영역(예컨대, 16)에 대한 논리 어드레스(예컨대, LA0)를 제2스케일-아웃 저장 장치(430-1)의 히든(hidden) 목적지 블록(예컨대, 웨어-카운트가 낮은 블록) 또는 상기 히든 목적지 블록의 저장 영역(예컨대, 36)에 대한 논리 어드레스(예컨대, LA101)로 매핑할 수 있다. 예컨대, 저장 영역(예컨대, 36)은 제2오버프로비저닝 영역(20)에 존재할 수 있다.

[0171] 도 14는 논리 어드레스 스왑 작동을 위한 스왑 매핑 테이블의 다른 실시 예이다. 도 9, 도 10, 및 도 14를 참조하면, 제1코어(411-1)는, 제1요청(REQ1)에 응답하여 또는 스스로, 제1코어(411-1)에 의해 관리되는 적어도 하나의 NAND-타입 플래시 메모리 장치(NAND)의 메모리 셀 어레이에 포함된 블록들 각각의 웨어-아웃 카운트와 기준 웨어-아웃 카운트를 비교하고, 비교 결과에 따라 상기 기준 웨어-아웃 카운트보다 높은 카운트를 갖는 블록(들)을 서치하고, 서치 결과에 따라 제1리스트(LIST1)를 생성할 수 있다(S103-1 또는 S103-1A). 제1코어(411-1)는 제1리스트(LIST1)를 컨트롤러(310)로 전송할 수 있다(S105-1 또는 S105-1A). 제1리스트(LIST1)는, 제1코어(411-1)의 제어에 따라, 휘발성 메모리(413-1) 및/또는 제1코어(411-1)에 접속된 적어도 하나의 NAND-타입 플래시 메모리 장치(NAND)에 저장될 수 있다.

[0172] 제2코어(431-1)는, 제2요청(REQ2)에 응답하여 또는 스스로, 제2코어(431-1)에 의해 관리되는 적어도 하나의 NAND-타입 플래시 메모리 장치(NAND)의 제2오버프로비저닝 영역(22)에 포함된 블록들(OBlock#0~OBlock#3) 각각의 웨어-아웃 카운트와 기준 웨어-아웃 카운트를 비교하고, 비교 결과에 따라 상기 기준 웨어-아웃 카운트보다 낮은 카운트를 갖는 블록(들)을 서치하고, 서치 결과에 따라 제2리스트(LIST2)를 생성할 수 있다(S103-2 또는 S103-2A). 제2코어(431-1)는 제2리스트(LIST2)를 컨트롤러(310)로 전송할 수 있다(S105-2 또는 S105-2A). 제2리스트(LIST2)는, 제2코어(431-1)의 제어에 따라, 휘발성 메모리(433-1) 및/또는 제2코어(431-1)에 접속된 적어도 하나의 NAND-타입 플래시 메모리 장치(NAND)에 저장될 수 있다.

[0173] 컨트롤러(310)의 CPU(330)는, 제1리스트(LIST1)와 제2리스트(LIST2)를 수신하고, 이들에 기초하여, 도 14에 도시된 바와 같은 매핑 테이블(TB4)을 생성하고, 매핑 테이블(TB4)을 휘발성 메모리(360)에 저장할 수 있다(S107). 예컨대, 매핑 테이블(TB4)은 테이블(TABLE1)의 적어도 일부일 수 있다.

[0174] 논리 어드레스(LA0=GLPN0 또는 LA0=1LPN0)가 프로그램(또는 라이트) 작동을 위한 논리 어드레스일 때, 경우 2(CASE2)에 도시된 바와 같이, 컨트롤러(310)는 호스트(200)로부터 출력된 프로그램 데이터를 도 14에 도시된 논리 어드레스(LA101=GLPN203=203 또는 LA101=2LPN101)에 대응하는 물리 어드레스(2PPN101)에 의해 지정된 제3저장 영역, 예컨대, 제2메모리 영역(20)의 제2프로비저닝 영역(22)의 제1블록(Block#0)의 제1페이지(OPage#0; 36)에 프로그램할 수 있다.

[0175] 즉, 컨트롤러(310)는 프로그램 데이터를 제1메모리 영역(10)의 제1저장 영역(Page#0; 16) 대신에 제2메모리 영역(20)의 제3저장 영역(OPage#0; 36)에 프로그램할 수 있다. 이에 따라, 제2메모리 영역(20)의 유효 페이지(valid page)가 증가한다.

[0176] 데이터 저장 장치(300 또는 300A)에 포함된 클러스터들(410과 430) 사이에서 또는 스케일-아웃 저장 장치들(410-1과 430-1)에서 성능의 불균형이 발생할 경우, 컨트롤러(310)는, 상기 성능의 불균형을 제어 또는 제거하기 위해, 각 오버프로비저닝 영역(12와 22)의 크기를 제어할 수 있다.

[0177] 예컨대, 제1스케일-아웃 저장 장치(410-1)의 성능이 제2스케일-아웃 저장 장치(430-1)의 성능보다 상대적으로 낮아지면, 컨트롤러(310)는, 코어들(411-1과 431-1)을 이용하여, 제1메모리 영역(10)의 제1저장 영역(16)에 저장된 콜드 데이터(cold data)를 제2메모리 영역(20)의 제3저장 영역(36)으로 옮길 수 있다. 이에 따라, 제1스케일-아웃 저장 장치(410-1)의 제1메모리 영역(10)의 제1오버프로비저닝 영역(12)이 증가하고, 제1스케일-아웃

저장 장치(410-1)의 성능은 증가할 수 있다.

- [0178] 다른 예로서, 제1스케일-아웃 저장 장치(410-1)의 성능이 제2스케일-아웃 저장 장치(430-1)의 성능보다 상대적으로 낮아지면, 컨트롤러(310)는, 코어들(411-1과 431-1)을 이용하여, 제1메모리 영역(10)의 제1저장 영역(16)에 저장된 핫 데이터(hot data)를 제2메모리 영역(20)의 제3저장 영역(36)으로 옮길 수 있다. 이에 따라, 제1스케일-아웃 저장 장치(410-1)의 워크로드(workload)는 제2스케일-아웃 저장 장치(430-1)로 분산될 수 있으므로, 제1스케일-아웃 저장 장치(410-1)의 성능은 증가할 수 있다.
- [0179] 제2스케일-아웃 저장 장치(430-1)의 성능이 제1스케일-아웃 저장 장치(410-1)의 성능보다 상대적으로 낮아지면, 컨트롤러(310)는, 코어들(411-1과 431-1)을 이용하여, 제2메모리 영역(20)의 제2사용자 데이터 영역(24)에 저장된 콜드 데이터 또는 핫 데이터를 제1메모리 영역(10)의 제1오버프로니저닝 영역(12)으로 옮길 수 있다.
- [0180] 핫 데이터는 호스트(200) 또는 컨트롤러(310)에 의해 상대적으로 자주 (frequently) 액세스되는 데이터이고, 콜드 데이터는 호스트(200) 또는 컨트롤러 (310)에 의해 상대적으로 덜 자주(less frequently) 액세스되는 데이터이다.
- [0181] 상술한 바와 같이, 제1스케일-아웃 저장 장치(410-1)의 수명과 성능 중에서 적어도 하나가 제2스케일-아웃 저장 장치(430-1)의 수명과 성능 중에서 적어도 하나와 다를 때, 컨트롤러(310)는, 도 11과 도 14를 참조하여 설명한 매핑 테이블을 이용하여, 핫 데이터를 저장하는 저장 영역의 논리 어드레스와 콜드 데이터를 저장하는 저장 영역의 논리 어드레스를 스왑하거나 상기 핫 데이터와 상기 콜드 데이터를 스왑할 수 있다. 이러한 방법들을 통해, 스케일-아웃 저장 장치들(410-1과 430-1) 사이의 성능 차이 및/또는 수명 차이는 제거될 수 있다.
- [0182] 도 15는 스왑 매핑 테이블이 스케일-아웃 저장 장치에 포함될 때, 데이터 저장 장치의 작동을 설명하는 개념도이다.
- [0183] 도 11과 도 14를 참조하여 설명된 매핑 테이블(TB1 또는 TB4)은 도 8a와 도 8b에 도시된 바와 같이 컨트롤러(310)에 의해 사용되는 버퍼(360), 예컨대 DRAM에 저장될 수 있다. 각 매핑 테이블(TB1 또는 TB4)은 매핑 테이블(TABLE1)의 일부 또는 전부일 수 있다.
- [0184] 그러나, 도 11과 도 14를 참조하여 설명된 매핑 테이블(TB1 또는 TB4)은 각 스케일-아웃 컨트롤러(411-1과 431-1)에 의해 관리되는 각 휘발성 메모리(413-1과 433-1)에 저장될 수 있다.
- [0185] 본 명세서에서 각 매핑 테이블(TB1~TB4)은 컨트롤러(310, 411-1, 또는 431-1) 외부의 휘발성 메모리(360, 413-1, 또는 433-1)에 저장되는 것으로 도시되어 있으나, 각 매핑 테이블(TB1~TB4)은 컨트롤러(310, 411-1, 또는 431-1) 내부의 휘발성 메모리(315 또는 425)에 저장될 수 있다.
- [0186] 도 15를 참조하면, 도 11과 도 14를 참조하여 설명된 매핑 테이블(TB1 또는 TB4, 집합적으로 "TABLE1")은 휘발성 메모리(413-1과 433-1)에 저장된다. 각 스케일-아웃 컨트롤러(또는 각 코어; 411-1과 431-1)는 도 9 또는 도 10을 참조하여 설명된 각 리스트들(LIST1과 LIST2)을 생성하고, 각 리스트들(LIST1과 LIST2)을 이용하여 매핑 테이블(TB1 또는 TB4, 집합적으로 "TABLE1")을 휘발성 메모리(413-1과 433-1)에 저장할 수 있다.
- [0187] 휘발성 메모리(413-1)는 도 12에 도시된 논리 어드레스-물리 어드레스 매핑 테이블(TABLE2=TB2)을 저장할 수 있고, 휘발성 메모리(433-1)는 도 12에 도시된 논리 어드레스-물리 어드레스 매핑 테이블(TABLE3=TB3)을 저장할 수 있다.
- [0188] 예컨대, 제1스케일-아웃 컨트롤러(411-1)는 제1리스트(LIST1)를 컨트롤러 (310)를 통해 또는 전용 라인을 통해 제2스케일-아웃 컨트롤러(431-1)로 전송할 수 있다. 또한, 제2스케일-아웃 컨트롤러(431-1)는 제2리스트(LIST2)를 컨트롤러 (310)를 통해 또는 상기 전용 라인을 통해 제1스케일-아웃 컨트롤러(411-1)로 전송할 수 있다.
- [0189] 도 8b부터 도 15를 참조하면, 논리 어드레스(LA0)가 프로그램(또는 라이트) 작동을 위한 논리 어드레스일 때, 컨트롤러(310)는 호스트(200)로부터 출력된 논리 어드레스(LA0)와 프로그램 데이터를 제1메인 채널(CHA)을 통해 제1스케일-아웃 컨트롤러(411-1)로 전송할 수 있다.
- [0190] 제1스케일-아웃 컨트롤러(411-1)는, 휘발성 메모리(413-1)에 저장된 매핑 테이블(TABLE1=TB1)을 이용하여, 논리 어드레스(LA0)를 스왑된 논리 어드레스 (LA201=GLPN201 또는 LA201=2LPN100)로 변경하고, 스왑된 논리 어드레스 (LA201=GLPN201 또는 LA201=2LPN100)를 컨트롤러(310)를 통해 또는 상기 전용 라인을 통해 제2스케일-아웃 컨트롤러(431-1)로 출력할 수 있다. 이때, 제1스케일-아웃 컨트롤러(411-1)는 호스트(200)로부터 출력된 프로그램 데이터를 컨트롤러(310)를 통해 또는 상기 전용 라인을 통해 제2스케일-아웃 컨트롤러(431-1)로 출력할 수

있다.

- [0191] 제2스케일-아웃 컨트롤러(431-1)는, 휘발성 메모리(433-1)에 저장된 매핑 테이블(TB3)을 참조하여, 논리 어드레스(LA201=GLPN201 또는 LA201=2LPN100)를 물리 어드레스(2PPN100)로 변환하고, 제1스케일-아웃 컨트롤러(411-1)로부터 전송된 프로그램 데이터를 물리 어드레스(2PPN100)에 해당하는 제2저장 영역(26)에 프로그램할 수 있다.
- [0192] 논리 어드레스(LA0)가 리드 작동을 위한 논리 어드레스일 때, 컨트롤러(310)는 호스트(200)로부터 출력된 논리 어드레스(LA0)를 제1메인 채널(CHA)을 통해 제1스케일-아웃 컨트롤러(411-1)로 전송할 수 있다.
- [0193] 제1스케일-아웃 컨트롤러(411-1)는, 휘발성 메모리(413-1)에 저장된 매핑 테이블(TABLE1=TB1)을 이용하여, 논리 어드레스(LA0)에 대응되는 스왑된 논리 어드레스(LA201=GLPN201 또는 LA201=2LPN100)를 컨트롤러(310)를 통해 또는 상기 전용 라인을 통해 제2스케일-아웃 컨트롤러(431-1)로 출력할 수 있다.
- [0194] 제2스케일-아웃 컨트롤러(431-1)는, 휘발성 메모리(433-1)에 저장된 매핑 테이블(TB3)을 참조하여, 논리 어드레스(LA201=GLPN201 또는 LA201=2LPN100)를 물리 어드레스(2PPN100)로 변환하고, 물리 어드레스(2PPN100)에 해당하는 제2저장 영역 (26)으로부터 데이터를 리드하고, 리드된 데이터를 제2메인 채널(CHB)을 통해 컨트롤러(310)로 전송한다. 컨트롤러(310)는 제2저장 영역(26)으로부터 출력된 데이터를 호스트(200)로 전송할 수 있다.
- [0195] 논리 어드레스(LA0)가 프로그램(또는 라이트) 작동을 위한 논리 어드레스일 때, 컨트롤러(310)는 호스트(200)로부터 출력된 논리 어드레스(LA0)와 프로그램 데이터를 제1메인 채널(CHA)을 통해 제1스케일-아웃 컨트롤러(411-1)로 전송할 수 있다.
- [0196] 제1스케일-아웃 컨트롤러(411-1)는, 휘발성 메모리(413-1)에 저장된 매핑 테이블(TABLE1=TB4)을 이용하여, 논리 어드레스(LA0)를 매핑된 논리 어드레스 (LA203=GLPN203 또는 LA203=2LPN101)로 변경하고, 논리 어드레스 (LA203=GLPN203 또는 LA203=2LPN101)를 컨트롤러(310)를 통해 또는 상기 전용 라인을 통해 제2스케일-아웃 컨트롤러(431-1)로 출력할 수 있다. 이때, 제1스케일-아웃 컨트롤러(411-1)는 호스트(200)로부터 출력된 프로그램 데이터를 컨트롤러(310)를 통해 또는 상기 전용 라인을 통해 제2스케일-아웃 컨트롤러(431-1)로 출력할 수 있다.
- [0197] 제2스케일-아웃 컨트롤러(431-1)는, 휘발성 메모리(433-1)에 저장된 매핑 테이블(TB3)을 참조하여, 논리 어드레스(LA203=GLPN203 또는 LA203=2LPN101)를 물리 어드레스(2PPN101)로 변환하고, 제1스케일-아웃 컨트롤러(411-1)로부터 전송된 프로그램 데이터를 물리 어드레스(2PPN101)에 해당하는 제3저장 영역(36)에 프로그램할 수 있다.
- [0198] 논리 어드레스(LA0)가 리드 작동을 위한 논리 어드레스일 때, 컨트롤러(310)는 호스트(200)로부터 출력된 논리 어드레스(LA0)를 제1메인 채널(CHA)을 통해 제1스케일-아웃 컨트롤러(411-1)로 전송할 수 있다.
- [0199] 제1스케일-아웃 컨트롤러(411-1)는, 휘발성 메모리(413-1)에 저장된 매핑 테이블(TABLE1=TB4)을 이용하여, 논리 어드레스(LA0)에 대응되는 매핑된 논리 어드레스(LA203=GLPN203 또는 LA203=2LPN101)를 컨트롤러(310)를 통해 또는 상기 전용 라인을 통해 제2스케일-아웃 컨트롤러(431-1)로 출력할 수 있다.
- [0200] 제2스케일-아웃 컨트롤러(431-1)는, 휘발성 메모리(433-1)에 저장된 매핑 테이블(TB3)을 참조하여, 논리 어드레스(LA203=GLPN203 또는 LA203=2LPN101)를 물리 어드레스(2PPN101)로 변환하고, 물리 어드레스(2PPN101)에 해당하는 제3저장 영역 (36)으로부터 데이터를 리드하고, 리드된 데이터를 제2메인 채널(CHB)을 통해 컨트롤러(310)로 전송한다. 컨트롤러(310)는 제3저장 영역(36)으로부터 출력된 데이터를 호스트(200)로 전송할 수 있다.
- [0201] 도 16은 본 발명의 다른 실시 예에 따른 데이터 처리 시스템을 나타내는 블록도이다. 도 1부터 도 16을 참조하면, 데이터 처리 시스템(100B)은 호스트(200)와 DRAM이 없는(DRAM-less) 데이터 저장 장치(300B)를 포함할 수 있다. 예컨대, 데이터 처리 시스템(100B)은 태블릿 PC 또는 eMMC로 구현될 수 있다. DRAM이 없는 데이터 저장 장치(300B)는 컨트롤러(310-1)와 메모리 클러스터(400)를 포함할 수 있다.
- [0202] 버퍼 컨트롤러(340-1)는 외부 버퍼에 대한 데이터 처리를 수행하지 않는다. 라이트 작동 시 또는 리드 작동 시, 버퍼 컨트롤러(340-1)는 내부 버퍼(341)를 이용하여 데이터를 제2인터페이스(320) 또는 제3인터페이스(350)로 전송할 수 있다. 버퍼 컨트롤러(340-1)의 작동을 제외하면, 도 16의 컨트롤러(310-1)의 구조와 작동은 도 1의 컨트롤러(310)의 구조와 작동과 실질적으로 동일 또는 유사하므로, 컨트롤러(310-1)에 대한 상세한 설명은 생략한다.

- [0203] 도 17은 본 발명의 또 다른 실시 예에 따른 데이터 처리 시스템을 나타내는 블록도이다. 도 1부터 도 17을 참조하면, 데이터 처리 시스템(500)은 데이터베이스 (520), 데이터베이스 서버(530), 제2네트워크(540), 및 복수의 클라이언트 컴퓨터들(550과 551)을 포함할 수 있다. 데이터 센터, 인터넷 데이터 센터, 또는 클라우드 데이터 센터(510)는 데이터베이스(520)와 데이터베이스 서버(530)를 포함할 수 있다.
- [0204] 데이터베이스(520)는 복수의 데이터 저장 장치들(300, 300A, 또는 300B, 집합적으로 "300")을 포함할 수 있다. 복수의 데이터 저장 장치들(300)은 랙(rack)에 설치될 수 있다. 각 데이터 저장 장치(300)의 구조와 작동은 도 1부터 도 16을 참조하여 설명된 데이터 저장 장치(300)의 구조와 작동과 실질적으로 동일 또는 유사하다.
- [0205] 데이터베이스 서버(530)는 복수의 데이터 저장 장치들(300) 각각의 작동을 제어할 수 있다. 예컨대, 데이터베이스 서버(530)는 도 1에 도시된 호스트(200)의 기능을 수행할 수 있다. 데이터베이스 서버(530)는 제1네트워크(535), 예컨대 랜 (Local Area Network(LAN))을 통해 제2무선 네트워크(540), 예컨대, 인터넷 또는 Wi-Fi에 접속될 수 있다. 복수의 클라이언트 컴퓨터들(550과 551) 각각은 제2네트워크(540)를 통해 데이터베이스 서버(530)에 접속할 수 있다.
- [0206] 도 18은 도 8a 또는 도 8b에 도시된 데이터 저장 장치의 작동을 설명하는 플로우 차트이다. 도 8a부터 도 18을 참조하면, 컨트롤러(310)는 호스트(200)로부터 논리 어드레스를 수신할 수 있다(S210). 논리 어드레스가 라이트 작동을 위한 논리 어드레스일 때(S220의 YES), 컨트롤러(310)는 호스트(200)로부터 데이터(DATA)를 수신할 수 있다(S230).
- [0207] 도 8a에 도시된 바와 같이, 각 논리 어드레스(LA0와 LA100)가 스왑되지 않았을 때(S240의 NO), 각 스케일-아웃 컨트롤러(411-1과 431-1)는 각 논리 어드레스 (LA0과 LA100)에 해당하는 각 영역에 각 데이터(DATA)를 프로그램 (또는 라이트)할 수 있다(S242).
- [0208] 그러나, 도 8b에 도시된 바와 같이, 논리 어드레스들이 스왑됐을 때(S240의 YES), 스케일-아웃 컨트롤러(431-1)는, 매핑 테이블(도 11의 TB1 또는 도 14의 TB4)을 이용하여 스왑된 논리 어드레스에 해당하는 영역(예컨대, 26 또는 36)에 데이터(DATA)를 프로그램할 수 있다(S244).
- [0209] 호스트(200)로부터 출력된 논리 어드레스가 리드 작동을 위한 논리 어드레스일 때(S220의 NO), 컨트롤러(310)는 논리 어드레스들의 스왑 여부를 판단할 수 있다(S250).
- [0210] 도 8a에 도시된 바와 같이, 각 논리 어드레스(LA0와 LA100)가 스왑되지 않았을 때(S250의 NO), 각 스케일-아웃 컨트롤러(411-1과 431-1)는 각 논리 어드레스 (LA0과 LA100)에 해당하는 각 영역으로부터 데이터를 리드하고, 상기 데이터를 컨트롤러(310)로 출력할 수 있다(S252). 컨트롤러(310)는 상기 데이터를 호스트(200)로 출력할 수 있다.
- [0211] 그러나, 도 8b에 도시된 바와 같이, 논리 어드레스들이 스왑됐을 때(S250의 YES), 스케일-아웃 컨트롤러(431-1)는, 매핑 테이블(도 11의 TB1 또는 도 14의 TB4)을 이용하여 스왑된 논리 어드레스에 해당하는 영역(예컨대, 26 또는 36)으로부터 데이터를 리드하고, 상기 데이터를 컨트롤러(310)로 출력할 수 있다(S254). 컨트롤러(310)는 상기 데이터를 호스트(200)로 출력할 수 있다.
- [0212] 본 발명은 도면에 도시된 실시 예를 참고로 설명되었으나 이는 예시적인 것에 불과하며, 본 기술 분야의 통상의 지식을 가진 자라면 이로부터 다양한 변형 및 균등한 타 실시 예가 가능하다는 점을 이해할 것이다. 따라서, 본 발명의 진정한 기술적 보호 범위는 첨부된 등록청구범위의 기술적 사상에 의해 정해져야 할 것이다.

## 부호의 설명

- [0213] 100: 데이터 처리 시스템
- 200: 호스트
- 300: 데이터 저장 장치
- 310: 컨트롤러
- 400: 메모리 클러스터
- 410, 430, 및 450: 클러스터
- 410-1, 410-2, 410-4: 스케일-아웃 저장 장치, 멀티 칩 셋

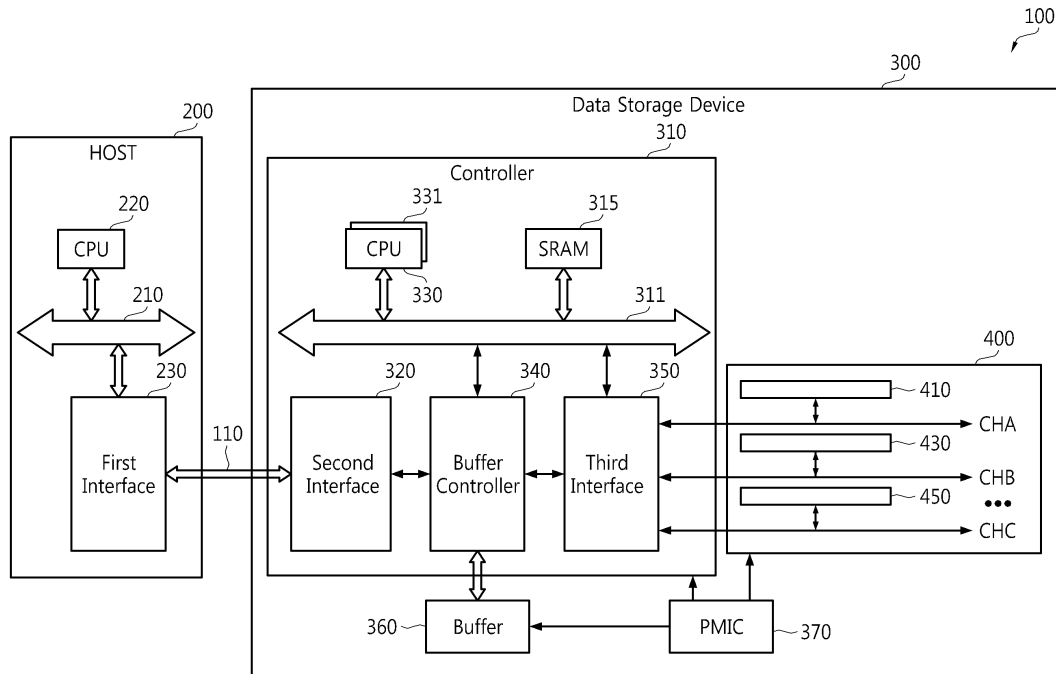


410-1A, 410-2A, 410-4A: 스케일-아웃 장치

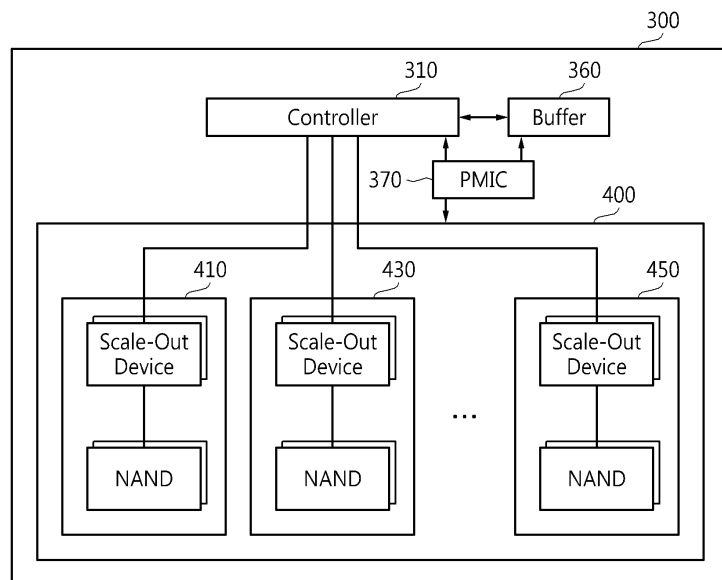
411-1, 411-2, 411-4, 431-1: 스케일-아웃 컨트롤러, 코어

도면

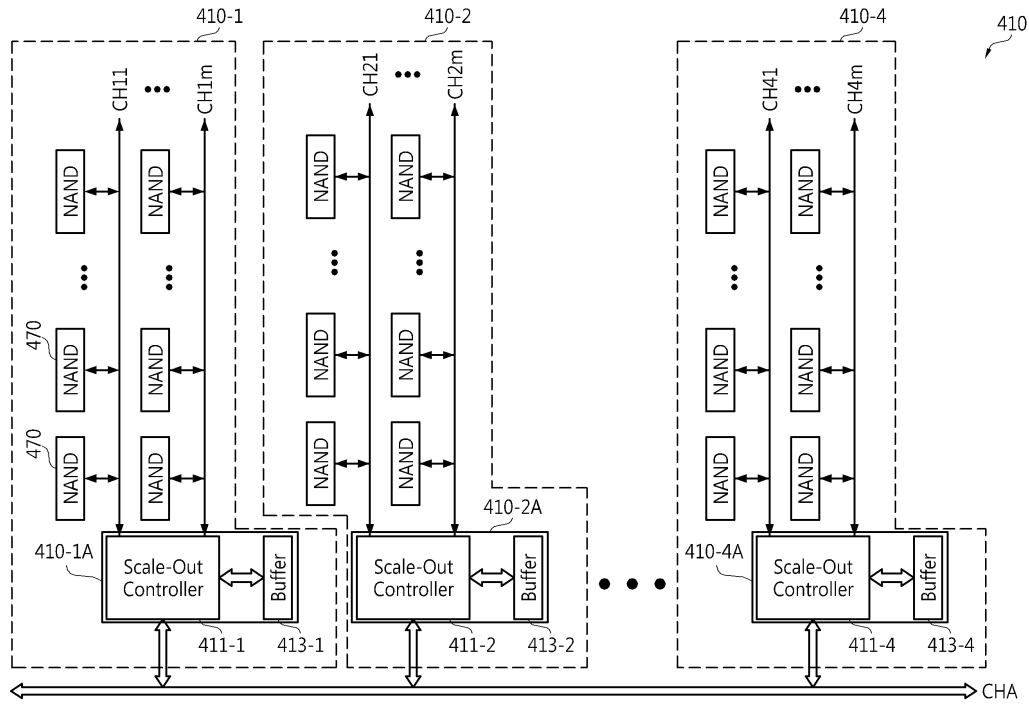
도면1



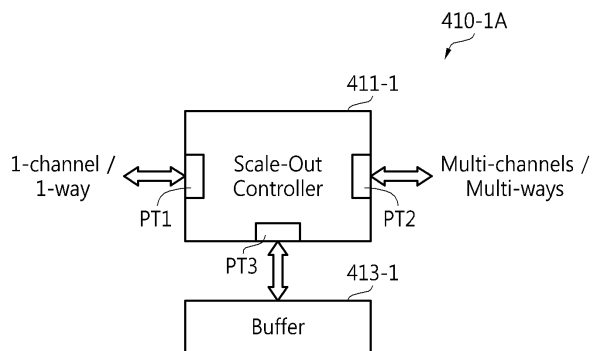
도면2



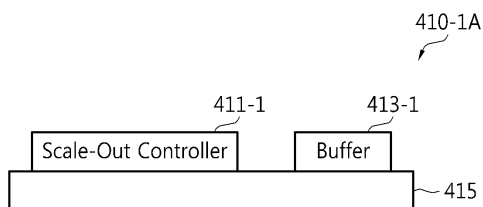
도면3



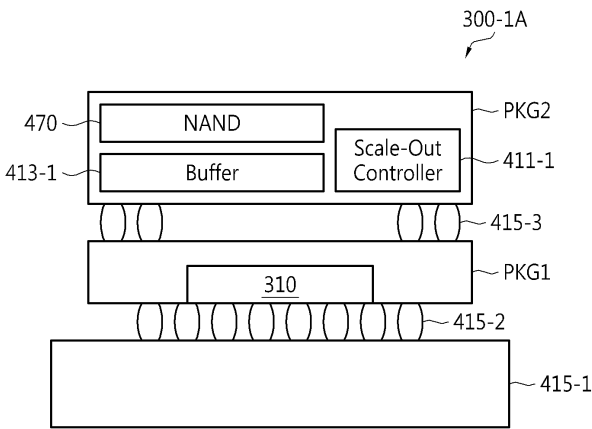
도면4



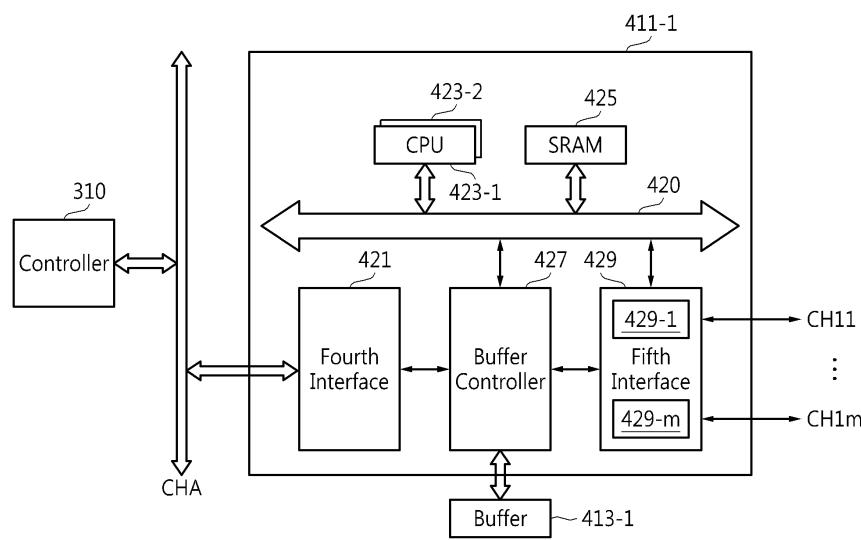
도면5



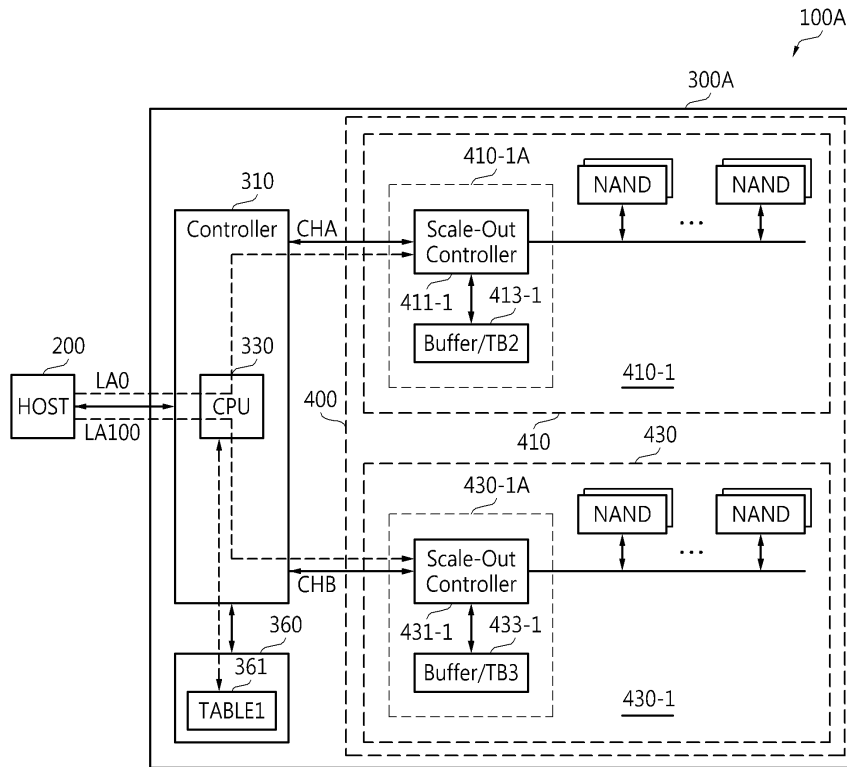
도면6



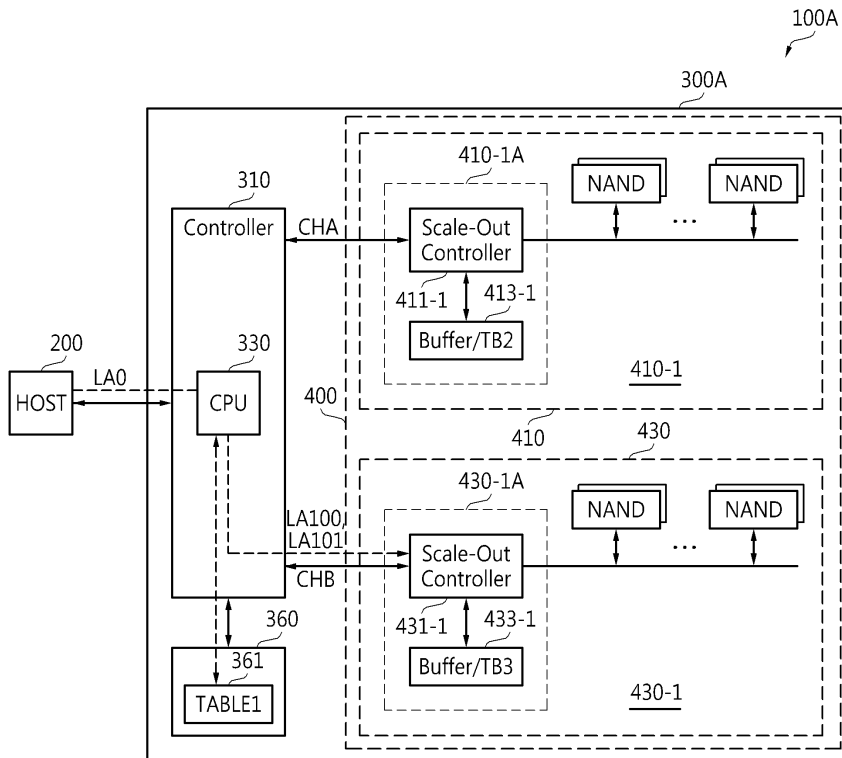
도면7



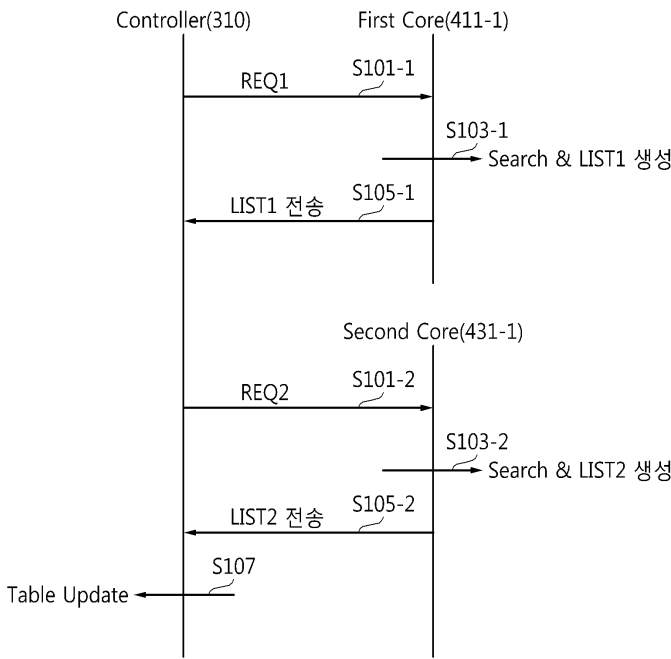
도면8a



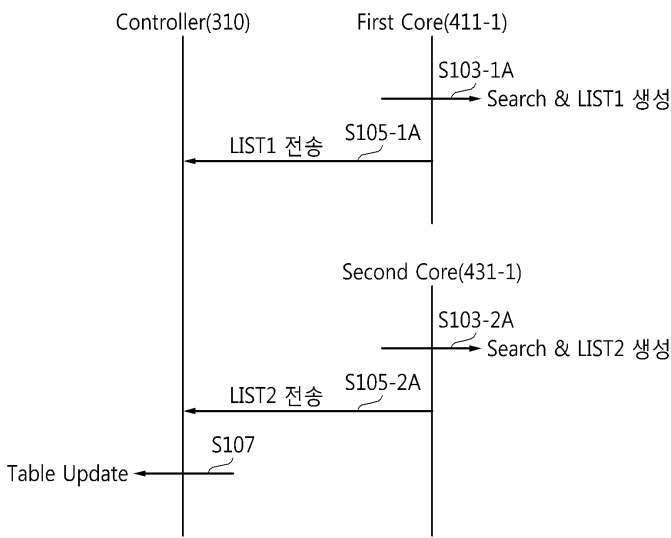
도면8b



도면9



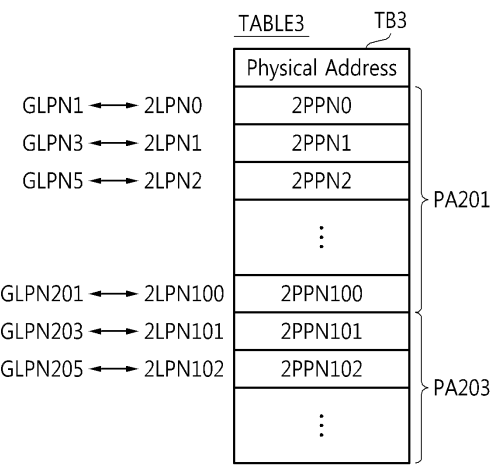
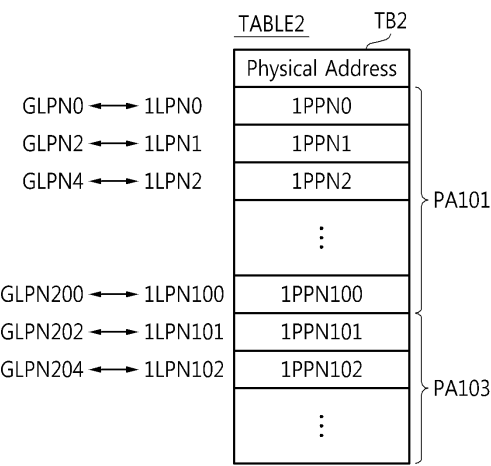
도면10



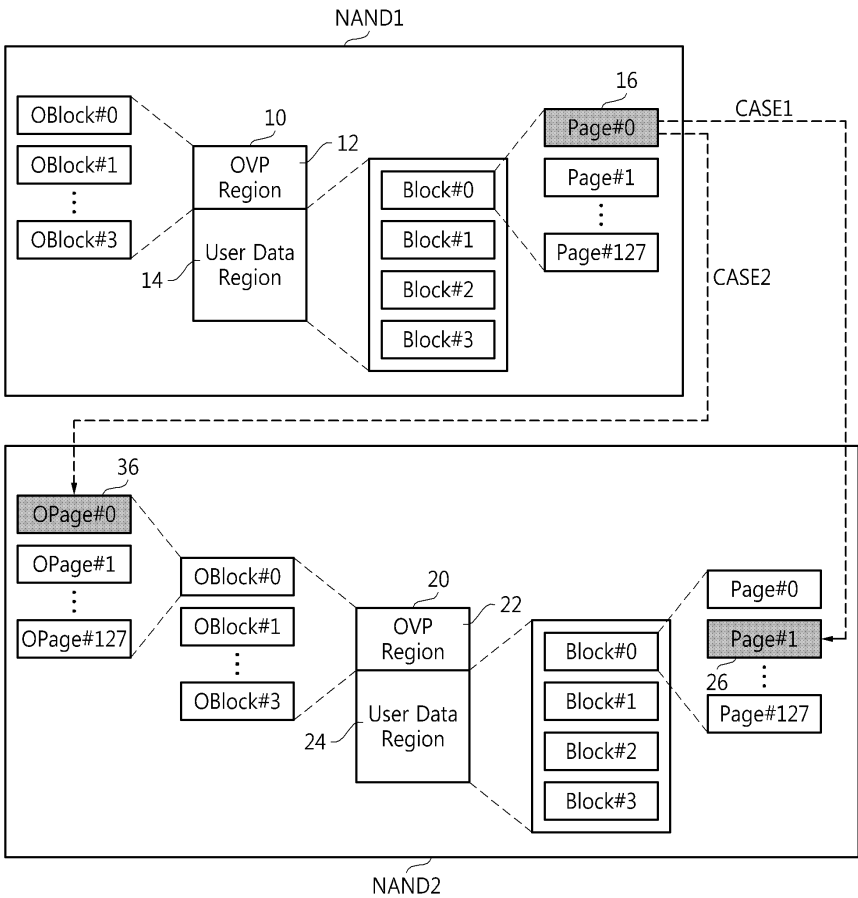
도면11

TABLE1		TB1	
Source Global/Local Logical Address		Destination Global/Local Logical Address	
0/1LPN0		201/2LPN100	
⋮		⋮	

도면12



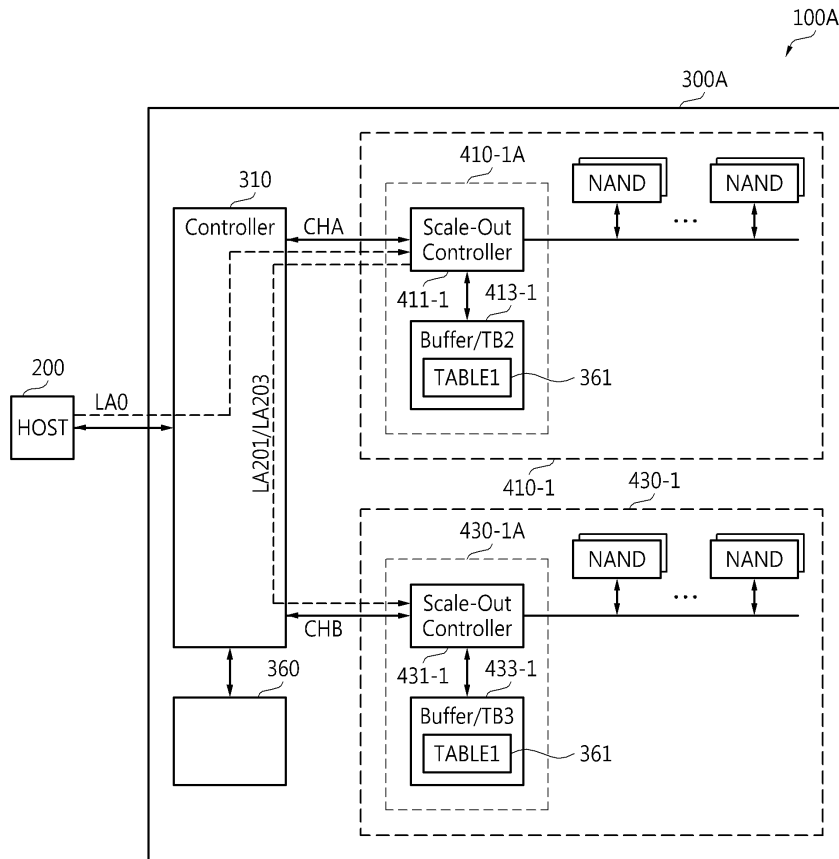
도면13



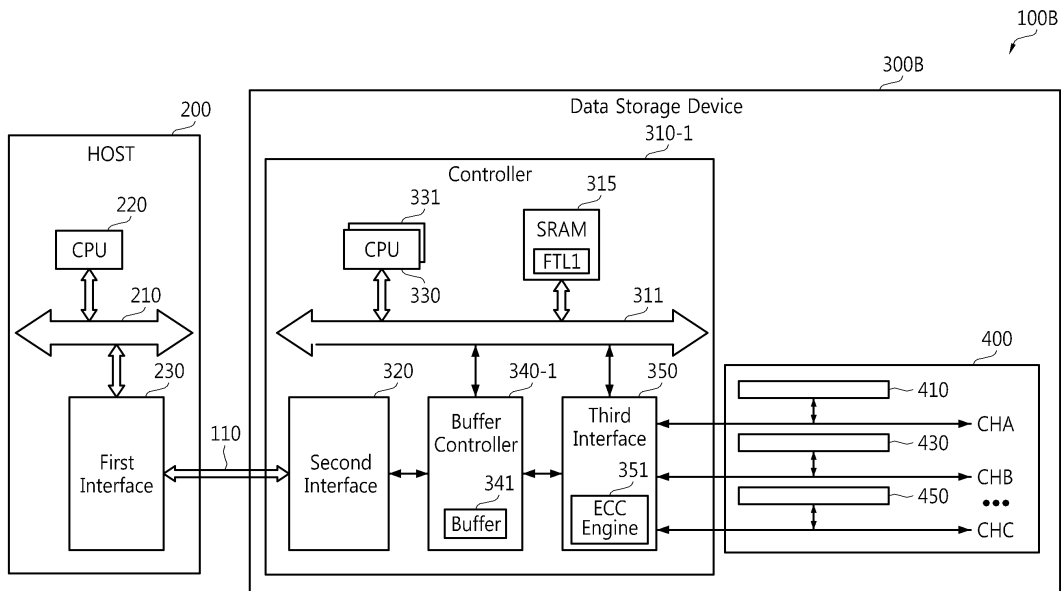
도면14

TABLE1		TB4	
Source Global/Local Logical Address		Destination Global/Local Logical Address	
0/1LPN0		203/2LPN101	
⋮		⋮	

도면15

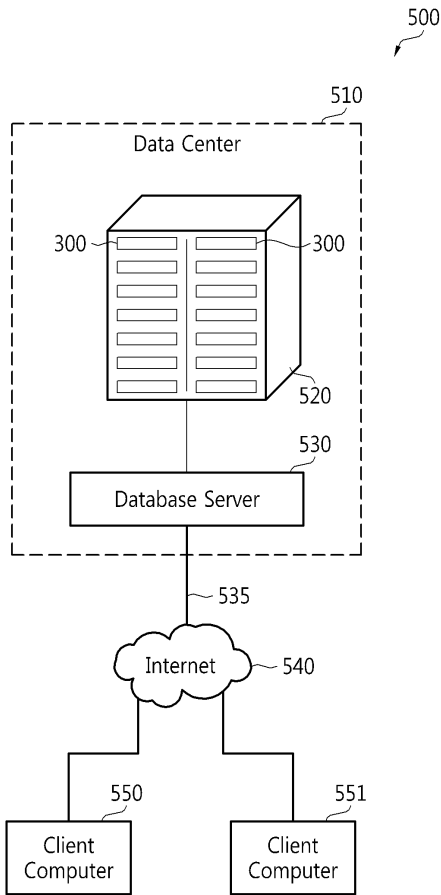


도면16





도면17



도면18

