



(19) 대한민국특허청(KR)  
 (12) 등록특허공보(B1)

(45) 공고일자 2019년05월17일  
 (11) 등록번호 10-1979715  
 (24) 등록일자 2019년05월13일

(51) 국제특허분류(Int. Cl.)  
*G06F 12/00* (2016.01) *G06F 3/06* (2006.01)

(21) 출원번호 10-2012-0109177

(22) 출원일자 2012년09월28일  
 심사청구일자 2017년07월06일

(65) 공개번호 10-2014-0042426

(43) 공개일자 2014년04월07일

(56) 선행기술조사문헌

US07257690 B1\*

US20080288713 A1\*

US20120011340 A1\*

\*는 심사관에 의하여 인용된 문헌

(73) 특허권자

삼성전자 주식회사

경기도 수원시 영통구 삼성로 129 (매탄동)

(72) 발명자

황주영

경기 수원시 영통구 청명로 132, 324동 1201호 (영통동, 청명마을3단지아파트)

이철

경기 화성시 동탄중앙로 213, 244동 2103호 (반송동, 시범한빛마을금호어울림아파트)

(뒷면에 계속)

(74) 대리인

특허법인가산

전체 청구항 수 : 총 10 항

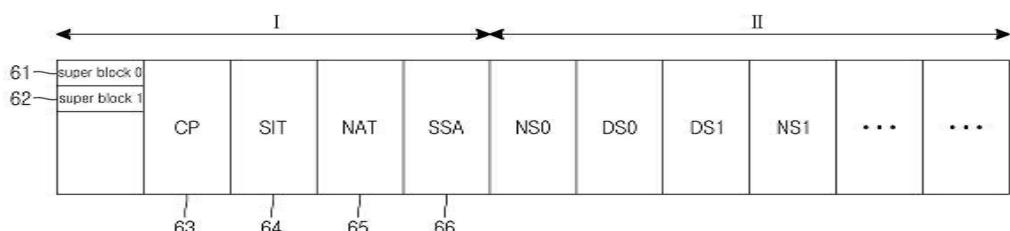
심사관 : 이상현

(54) 발명의 명칭 컴퓨팅 시스템 및 그 데이터 관리 방법

**(57) 요약**

컴퓨팅 시스템 및 그 데이터 관리 방법이 제공된다. 상기 컴퓨팅 시스템의 데이터 관리 방법은 스토리지 장치를 메타데이터 영역과 메인 영역으로 나누고, 상기 메인 영역에 다수의 데이터와, 상기 다수의 데이터와 연관된 다수의 노드(node)를 저장하고, 상기 메타데이터 영역에 노드 어드레스 테이블(node address table)를 저장하되, 상기 노드 어드레스 테이블은 상기 노드 각각에 대응되는 다수의 노드 식별자와, 상기 다수의 노드 식별자 각각에 대응되는 다수의 물리 어드레스를 포함한다.

**대 표 도** - 도5



(72) 발명자

김재극

경기 화성시 동탄반석로 277, 117동 1602호 (석우  
동, 예당마을우미린제일풍경채)

이창만

서울 강남구 논현로 209, 104동 802호 (도곡동, 경  
남아파트)

## 명세서

### 청구범위

#### 청구항 1

스토리지 장치를 메타데이터 영역과 메인 영역으로 나누고,

상기 메인 영역에 다수의 데이터와, 상기 다수의 데이터와 연관된 다수의 노드(node)를 저장하고, 상기 다수의 데이터는 제1 물리 어드레스에 대응하는 제1 블록 내 저장된 제1 데이터를 포함하고, 상기 다수의 노드는 상기 제1 데이터를 지시하는 제1 노드를 포함하되, 상기 제1 노드는 제2 물리 어드레스에 대응하는 제2 블록 내 저장되며,

상기 메타데이터 영역에 노드 어드레스 테이블(node address table)를 저장하되, 상기 노드 어드레스 테이블은 상기 노드 각각에 대응되는 다수의 노드 식별자를 포함하되, 상기 다수의 노드 식별자는 상기 제1 노드에 대응하는 제1 노드 식별자를 포함하고, 상기 다수의 노드 식별자 각각에 대응되는 다수의 물리 어드레스를 포함하되, 상기 다수의 물리 어드레스는 상기 제1 노드 식별자에 대응하는 상기 제2 물리 어드레스를 포함하며,

제1 데이터를 수정하여 제2 데이터를 생성하고, 상기 제2 데이터를 제3 물리 어드레스에 대응하는 비어있는 제3 블록 내에 저장하고,

제1 노드가 상기 제2 데이터를 지시하도록 수정하여, 상기 제1 노드를 제4 물리 어드레스에 대응되는 비어있는 제4 블록 내에 저장하고,

상기 노드 어드레스 테이블 내에서, 상기 제1 노드는 상기 제4 물리 어드레스에 대응하는 것으로 오버라이트(overwrite)하는 것을 포함하는 컴퓨팅 시스템의 데이터 관리 방법.

#### 청구항 2

제 1항에 있어서,

상기 메인 영역은 순차적 접근(sequential access) 방식으로 라이트되는 영역이고,

상기 메타데이터 영역은 랜덤 접근(random access) 방식으로 라이트되는 영역인 컴퓨팅 시스템의 데이터 관리 방법.

#### 청구항 3

제 1항에 있어서,

상기 메타데이터 영역은 상기 스토리지 장치의 앞부분에 저장되고,

상기 메인 영역은 상기 스토리지 장치의 뒷부분에 저장되는 컴퓨팅 시스템의 데이터 관리 방법.

#### 청구항 4

제 1항에 있어서,

상기 메인 영역은 다수의 세그먼트를 포함하고, 각 세그먼트는 다수의 페이지가 저장되고,

상기 메타데이터 영역은 세그먼트 정보 테이블(segment information table)를 더 포함하되, 상기 세그먼트 정보 테이블은 각 세그먼트의 유효한 페이지의 개수와, 다수의 페이지의 비트맵을 포함하는 컴퓨팅 시스템의 데이터 관리 방법.

#### 청구항 5

제 1항에 있어서,

상기 메인 영역은 다수의 세그먼트를 포함하고, 각 세그먼트는 다수의 블록으로 구분되고,

상기 메타데이터 영역은 세그먼트 요약 영역(segment summary area)을 더 포함하되, 상기 세그먼트 요약 영역은

각 세그먼트의 상기 다수의 블록이 속하는 노드의 정보를 포함하는 컴퓨팅 시스템의 데이터 관리 방법.

#### 청구항 6

제 1항에 있어서,

상기 메인 영역은 다수의 세그먼트를 포함하고, 각 세그먼트는 다수의 블록으로 구분되고,

상기 각 세그먼트의 적어도 하나의 블록에 세그먼트 요약 정보를 저장하되, 상기 세그먼트 요약 정보는 각 세그먼트의 상기 다수의 블록이 속하는 노드의 정보를 포함하는 컴퓨팅 시스템의 데이터 관리 방법.

#### 청구항 7

제 1항에 있어서,

상기 메인 영역은 다수의 세그먼트를 포함하고, 각 세그먼트는 다수의 블록으로 구분되고, 각 블록은 OOB(Out Of Band) 영역을 포함하고,

상기 OOB 영역에 세그먼트 요약 정보를 저장하되, 상기 세그먼트 요약 정보는 각 블록이 속하는 노드의 정보를 포함하는 컴퓨팅 시스템의 데이터 관리 방법.

#### 청구항 8

앞부분에 위치하는 메타데이터 영역과, 뒷부분에 위치하는 메인 영역을 포함하는 스토리지 장치;

상기 메인 영역에 다수의 데이터와, 상기 다수의 데이터와 연관된 다수의 노드(node)를 저장하되, 상기 다수의 데이터는 제1 물리 어드레스에 대응하는 제1 블록 내 저장된 제1 데이터를 포함하고, 상기 다수의 노드는 상기 제1 데이터를 지시하는 제1 노드를 포함하되, 상기 제1 노드는 제2 물리 어드레스에 대응하는 제2 블록 내 저장되며, 상기 메타데이터 영역에 노드 어드레스 테이블(node address table)를 저장하되, 상기 노드 어드레스 테이블은 상기 노드 각각에 대응되는 다수의 노드 식별자를 포함하되, 상기 다수의 노드 식별자는 상기 제1 노드에 대응하는 제1 노드 식별자를 포함하고 상기 다수의 노드 식별자 각각에 대응되는 다수의 물리 어드레스를 포함하는 호스트 장치를 포함하되,

상기 다수의 물리 어드레스는 상기 제1 노드 식별자에 대응하는 상기 제2 물리 어드레스를 포함하며, 제1 데이터를 수정하여 제2 데이터를 생성하고, 상기 제2 데이터를 제3 물리 어드레스에 대응하는 비어있는 제3 블록 내에 저장하고, 제1 노드가 상기 제2 데이터를 지시하도록 수정하여, 상기 제1 노드를 제4 물리 어드레스에 대응하는 비어있는 제4 블록 내에 저장하고, 상기 노드 어드레스 테이블 내에서 상기 제1 노드는 상기 제4 물리 어드레스에 대응하는 것으로 오버라이트(overwrite)하는 컴퓨팅 시스템.

#### 청구항 9

스토리지 장치를 관리하는 파일시스템(filesystem)을 포함하는 컴퓨팅 시스템에 있어서,

상기 파일시스템은 스토리지 장치를 메타데이터 영역과 메인 영역으로 나누고, 상기 메인 영역에 다수의 데이터와, 상기 다수의 데이터와 연관된 다수의 노드(node)를 저장하고, 상기 다수의 데이터는 제1 물리 어드레스에 대응하는 제1 블록 내 저장된 제1 데이터를 포함하고, 상기 다수의 노드는 상기 제1 데이터를 지시하는 제1 노드를 포함하되, 상기 제1 노드는 제2 물리 어드레스에 대응하는 제2 블록 내 저장되며,

상기 메타데이터 영역에 노드 어드레스 테이블(node address table)를 저장하되, 상기 노드 어드레스 테이블은 상기 노드 각각에 대응되는 다수의 노드 식별자를 포함하되, 상기 다수의 노드 식별자는 상기 제1 노드에 대응하는 제1 노드 식별자를 포함하고, 상기 다수의 노드 식별자 각각에 대응되는 다수의 물리 어드레스를 포함하되,

상기 다수의 물리 어드레스는 상기 제1 노드 식별자에 대응하는 상기 제2 물리 어드레스를 포함하며, 제1 데이터를 수정하여 제2 데이터를 생성하고, 상기 제2 데이터를 제3 물리 어드레스에 대응하는 비어있는 제3 블록 내에 저장하고, 제1 노드가 상기 제2 데이터를 지시하도록 수정하여, 상기 제1 노드를 제4 물리 어드레스에 비어있는 제4 블록 내에 저장하고, 상기 노드 어드레스 테이블 내에서, 상기 제1 노드는 상기 제4 물리 어드레스에 대응하는 것으로 오버라이트(overwrite)하는 컴퓨팅 시스템.

#### 청구항 10

메타데이터 영역;

상기 메타데이터 영역과 구분되는 메인 영역;

상기 메인 영역에 저장된 다수의 데이터;

상기 메인 영역에 저장되고, 상기 다수의 데이터와 연관된 다수의 노드(node)로서, 상기 다수의 데이터는 제1 물리 어드레스에 대응하는 제1 블록 내 저장된 제1 데이터를 포함하고, 상기 제1 데이터를 지시하는 제1 노드를 포함하되, 상기 제1 노드는 제2 물리 어드레스에 대응하는 제2 블록 내 저장되는 상기 다수의 노드; 및

상기 메타데이터 영역에 저장된 노드 어드레스 테이블(node address table)을 포함하되,

상기 노드 어드레스 테이블은 상기 노드 각각에 대응되는 다수의 노드 식별자를 포함하고, 상기 다수의 노드 식별자는 상기 제1 노드에 대응하는 제1 노드 식별자를 포함하고, 상기 다수의 노드 식별자 각각에 대응되는 다수의 물리 어드레스를 포함하는 노드 어드레스 테이블을 포함하되,

상기 다수의 물리 어드레스는 상기 제1 노드 식별자에 대응하는 상기 제2 물리 어드레스를 포함하며, 제1 데이터를 수정하여 제2 데이터를 생성하고, 상기 제2 데이터를 제3 물리 어드레스에 대응하는 비어있는 제3 블록 내에 저장하고, 제1 노드가 상기 제2 데이터를 지시하도록 수정하여, 상기 제1 노드를 제4 물리 어드레스에 대응되는 비어있는 제4 블록 내에 저장하고, 상기 노드 어드레스 테이블 내에서, 상기 제1 노드는 상기 제4 물리 어드레스에 대응하는 것으로 오버라이트(overwrite)하는 스토리지 장치.

## 발명의 설명

### 기술 분야

[0001]

본 발명은 컴퓨팅 시스템 및 그 데이터 관리 방법에 관한 것으로, 더 구체적으로 로그 구조화된 파일시스템(log structured filesystem)을 이용한 컴퓨팅 시스템 및 그 데이터 관리 방법에 관한 것이다.

### 배경 기술

[0002]

로그 구조화된 파일시스템은 하드 디스크 드라이브(Hard Disk Drive)를 사용하는 서버 스토리지 시스템에서 제안되었다. 하드 디스크 드라이브는 회전 모터를 사용하는 장치이므로, 검색 레이턴시(seek latency)와 회전 레이턴시(rotational latency)를 갖는다. 따라서, 로그 구조화된 파일시스템은 디스크 전체를 하나의 로그로 구성하여, 순차적인 라이트(sequential write)만을 한다. 즉, 로그 구조화된 파일시스템은, 파일을 수정할 때, 원래 위치의 데이터를 수정하지 않고, 로그의 마지막 부분에 수정된 데이터를 부가한다.

### 발명의 내용

#### 해결하려는 과제

[0003]

한편, 로그 구조화된 파일시스템은 수정된 데이터와 관련된 많은 메타데이터(metadata)의 수정도 해야 한다. 즉, 부가적인 라이트 동작이 필요하다. 이를 wandering tree 문제라고 한다.

[0004]

본 발명이 해결하려는 과제는, wandering tree 문제를 최소화한 컴퓨팅 시스템의 데이터 관리 방법을 제공하는 것이다.

[0005]

본 발명이 해결하려는 다른 과제는, wandering tree 문제를 최소화한 컴퓨팅 시스템을 제공하는 것이다.

[0006]

본 발명이 해결하려는 다른 과제는, wandering tree 문제를 최소화한 스토리지 장치를 제공하는 것이다.

[0007]

본 발명이 해결하려는 과제들은 이상에서 언급한 과제들로 제한되지 않으며, 언급되지 않은 또 다른 과제들은 아래의 기재로부터 당업자에게 명확하게 이해될 수 있을 것이다.

### 과제의 해결 수단

[0008]

상기 과제를 해결하기 위한 본 발명의 컴퓨팅 시스템의 데이터 관리 방법의 일 면(aspect)은 스토리지 장치를 메타데이터 영역과 메인 영역으로 나누고, 상기 메인 영역에 다수의 데이터와, 상기 다수의 데이터와 연관된 다수의 노드(node)를 저장하고, 상기 메타데이터 영역에 노드 어드레스 테이블(node address table)을 저장하되, 상기 노드 어드레스 테이블은 상기 노드 각각에 대응되는 다수의 노드 식별자와, 상기 다수의 노드 식별자 각각

에 대응되는 다수의 물리 어드레스를 포함한다.

- [0009] 상기 메인 영역은 순차적 접근(sequential access) 방식으로 라이트되는 영역이고, 상기 메타데이터 영역은 랜덤 접근(random access) 방식으로 라이트되는 영역일 수 있다.
- [0010] 상기 메타데이터 영역은 상기 스토리지 장치의 앞부분에 저장되고, 상기 메인 영역은 상기 스토리지 장치의 뒷부분에 저장될 수 있다. 상기 스토리지 장치는 SSD(Static Solid Disk)를 포함할 수 있다.
- [0011] 상기 메인 영역은 서로 분리된 데이터 세그먼트(data segment)과 노드 세그먼트(node segment)을 포함하고, 상기 다수의 데이터는 상기 데이터 세그먼트에 저장되고, 상기 다수의 노드는 상기 노드 세그먼트에 저장될 수 있다.
- [0012] 상기 메인 영역은 다수의 세그먼트를 포함하고, 각 세그먼트는 다수의 페이지가 저장되고, 상기 메타데이터 영역은 세그먼트 정보 테이블(segment information table)를 더 포함하되, 상기 세그먼트 정보 테이블은 각 세그먼트의 유효한 페이지의 개수와, 다수의 페이지의 비트맵을 포함할 수 있다.
- [0013] 상기 메인 영역은 다수의 세그먼트를 포함하고, 각 세그먼트는 다수의 블록으로 구분되고, 상기 메타데이터 영역은 세그먼트 요약 영역(segment summary area)을 더 포함하되, 상기 세그먼트 요약 영역은 각 세그먼트의 상기 다수의 블록이 속하는 노드의 정보를 포함할 수 있다.
- [0014] 상기 메인 영역은 다수의 세그먼트를 포함하고, 각 세그먼트는 다수의 블록으로 구분되고, 상기 각 세그먼트의 적어도 하나의 블록에 세그먼트 요약 정보를 저장하되, 상기 세그먼트 요약 정보는 각 세그먼트의 상기 다수의 블록이 속하는 노드의 정보를 포함할 수 있다.
- [0015] 상기 메인 영역은 다수의 세그먼트를 포함하고, 각 세그먼트는 다수의 블록으로 구분되고, 각 블록은 OOB(Out Of Band) 영역을 포함하고, 상기 OOB 영역에 세그먼트 요약 정보를 저장하되, 상기 세그먼트 요약 정보는 각 블록이 속하는 노드의 정보를 포함할 수 있다.
- [0016] 상기 다수의 노드는 아이노드(inode), 디렉트 노드(direct node), 인디렉트 노드(indirect node)를 포함하고, 상기 노드 어드레스 테이블은 상기 아이노드의 노드 식별자, 상기 디렉트 노드의 노드 식별자 및 상기 인디렉트 노드의 노드 식별자를 저장할 수 있다.
- [0017] 상기 다수의 데이터 중에서 제1 데이터는 제1 물리 어드레스에 대응되는 제1 블록 내에 저장되고, 제1 디렉트 노드는 상기 제1 데이터를 지시하고, 상기 제1 디렉트 노드는 제2 물리 어드레스에 대응되는 제2 블록 내에 저장되고, 상기 노드 어드레스 테이블 내에서, 상기 제1 디렉트 노드는 상기 제2 물리 어드레스에 대응되는 것으로 저장될 수 있다.
- [0018] 상기 제1 데이터를 수정하여 제2 데이터를 생성하고, 상기 제2 데이터를 상기 제1 물리 어드레스와 다른 제3 물리 어드레스에 대응되는 제3 블록 내에 저장하고, 상기 제1 디렉트 노드가 상기 제2 데이터를 지시하도록 수정하여, 상기 제2 물리 어드레스와 다른 제4 물리 어드레스에 대응되는 제4 블록 내에 저장하고, 상기 노드 어드레스 테이블 내에, 상기 제1 디렉트 노드는 상기 제4 물리 어드레스에 대응되는 것으로 오버라이트(overwrite)하는 것을 더 포함할 수 있다.
- [0019] 상기 과제를 해결하기 위한 본 발명의 컴퓨팅 시스템의 일 면(aspect)은 앞부분에 위치하는 메타데이터 영역과, 뒷부분에 위치하는 메인 영역을 포함하는 스토리지 장치; 상기 메인 영역에 다수의 데이터와, 상기 다수의 데이터와 연관된 다수의 노드(node)를 저장하고, 상기 메타데이터 영역에 노드 어드레스 테이블(node address table)을 저장하되, 상기 노드 어드레스 테이블은 상기 노드 각각에 대응되는 다수의 노드 식별자와, 상기 다수의 노드 식별자 각각에 대응되는 다수의 물리 어드레스를 포함하는 호스트 장치를 포함한다.
- [0020] 상기 메인 영역은 순차적 접근(sequential access) 방식으로 라이트되는 영역이고, 상기 메타데이터 영역은 랜덤 접근(random access) 방식으로 라이트되는 영역일 수 있다.
- [0021] 상기 메인 영역은 다수의 세그먼트를 포함하고, 각 세그먼트는 다수의 페이지가 저장되고, 상기 메타데이터 영역에 세그먼트 정보 테이블(segment information table)을 저장하되, 상기 세그먼트 정보 테이블은 각 세그먼트의 유효한 페이지의 개수와, 다수의 페이지의 비트맵을 포함할 수 있다.
- [0022] 상기 메인 영역은 다수의 세그먼트를 포함하고, 각 세그먼트는 다수의 블록으로 구분되고, 상기 메타데이터 영역에 세그먼트 요약 영역(segment summary area)을 저장하되, 상기 세그먼트 요약 영역은 각 세그먼트의 상기 다수의 블록이 속하는 노드의 정보를 포함할 수 있다.

[0023] 상기 과제를 해결하기 위한 본 발명의 컴퓨팅 시스템의 다른 면은, 스토리지 장치를 관리하는 파일시스템(filesystem)을 포함하는 컴퓨팅 시스템에 있어서, 상기 파일시스템은 스토리지 장치를 메타데이터 영역과 메인 영역으로 나누고, 상기 메인 영역에 다수의 데이터와, 상기 다수의 데이터와 연관된 다수의 노드(node)를 저장하고, 상기 메타데이터 영역에 노드 어드레스 테이블(node address table)를 저장하되, 상기 노드 어드레스 테이블은 상기 노드 각각에 대응되는 다수의 노드 식별자와, 상기 다수의 노드 식별자 각각에 대응되는 다수의 물리 어드레스를 포함할 수 있다.

[0024] 상기 과제를 해결하기 위한 본 발명의 스토리지 장치의 일 면은, 메타데이터 영역; 상기 메타데이터 영역과 구분되는 메인 영역; 상기 메인 영역에 저장된 다수의 데이터; 상기 메인 영역에 저장되고, 상기 다수의 데이터와 연관된 다수의 노드(node); 및 상기 메타데이터 영역에 저장된 노드 어드레스 테이블(node address table)로서, 상기 노드 각각에 대응되는 다수의 노드 식별자와, 상기 다수의 노드 식별자 각각에 대응되는 다수의 물리 어드레스를 포함하는 노드 어드레스 테이블을 포함할 수 있다.

[0025] 본 발명의 기타 구체적인 사항들은 상세한 설명 및 도면들에 포함되어 있다.

### 도면의 간단한 설명

[0026] 도 1은 본 발명의 일 실시예에 따른 컴퓨팅 시스템을 설명하기 위한 블록도이다.

도 2는 도 1의 호스트를 설명하기 위한 블록도이다.

도 3은 도 1의 스토리지 장치를 설명하기 위한 블록도이다.

도 4는 도 1의 스토리지 장치에 저장되는 파일의 구조를 설명하기 위한 도면이다.

도 5는 도 1의 스토리지 장치를 설명하기 위한 블록도이다.

도 6은 노드 어드레스 테이블을 설명하기 위한 도면이다.

도 7 및 도 8은 본 발명의 일 실시예에 따른 컴퓨팅 시스템의 데이터 관리 방법을 설명하기 위한 개념도이다.

도 9는 본 발명의 다른 실시예에 따른 컴퓨팅 시스템의 스토리지 장치에 저장되는 파일의 구조를 설명하기 위한 블록도이다.

도 10은 본 발명의 또 다른 실시예에 따른 컴퓨팅 시스템의 스토리지 장치에 저장되는 파일의 구조를 설명하기 위한 블록도이다.

도 11은 본 발명의 또 다른 실시예에 따른 컴퓨팅 시스템의 스토리지 장치에 저장되는 파일의 구조를 설명하기 위한 블록도이다.

도 12는 본 발명의 몇몇 실시예에 따른 컴퓨팅 시스템의 구체적인 일 예를 설명하기 위한 블록도이다.

도 13 내지 도 15는 본 발명의 몇몇 실시예에 따른 컴퓨팅 시스템의 구체적인 다른 예를 설명하기 위한 블록도이다.

### 발명을 실시하기 위한 구체적인 내용

[0027] 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. 명세서 전체에 걸쳐 동일 참조 부호는 동일 구성 요소를 지칭한다.

[0028] 하나의 소자(elements)가 다른 소자와 "접속된(connected to)" 또는 "커플링된(coupled to)"이라고 지칭되는 것은, 다른 소자와 직접 연결 또는 커플링된 경우 또는 중간에 다른 소자를 개재한 경우를 모두 포함한다. 반면, 하나의 소자가 다른 소자와 "직접 접속된(directly connected to)" 또는 "직접 커플링된(directly coupled to)"으로 지칭되는 것은 중간에 다른 소자를 개재하지 않은 것을 나타낸다. 명세서 전체에 걸쳐 동일 참조 부호는 동일 구성 요소를 지칭한다.

[0029] 예컨대, 어느 하나의 구성요소가 다른 구성요소로 데이터 또는 신호를 "전송 또는 출력"하는 경우에는 상기 구성요소는 상기 다른 구성요소로 직접 상기 데이터 또는 신호를 "전송 또는 출력"할 수 있고, 적어도 하나의 또

다른 구성요소를 통하여 상기 데이터 또는 신호를 상기 다른 구성요소로 "전송 또는 출력"할 수 있음을 의미한다.

[0030] "및/또는"은 언급된 아이템들의 각각 및 하나 이상의 모든 조합을 포함한다.

[0031] 비록 제1, 제2 등이 다양한 소자, 구성요소 및/또는 섹션들을 서술하기 위해서 사용되나, 이들 소자, 구성요소 및/또는 섹션들은 이들 용어에 의해 제한되지 않음은 물론이다. 이들 용어들은 단지 하나의 소자, 구성요소 또는 섹션들을 다른 소자, 구성요소 또는 섹션들과 구별하기 위하여 사용하는 것이다. 따라서, 이하에서 언급되는 제1 소자, 제1 구성요소 또는 제1 섹션은 본 발명의 기술적 사상 내에서 제2 소자, 제2 구성요소 또는 제2 섹션일 수도 있음을 물론이다.

[0032] 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, 단계, 동작 및/또는 소자는 하나 이상의 다른 구성요소, 단계, 동작 및/또는 소자의 존재 또는 추가를 배제하지 않는다.

[0033] 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.

[0034] 도 1은 본 발명의 일 실시예에 따른 컴퓨팅 시스템을 설명하기 위한 블록도이다. 도 2는 도 1의 호스트를 설명하기 위한 블록도이다. 도 3 및 도 5는 도 1의 스토리지 장치를 설명하기 위한 블록도이다. 도 4는 도 1의 스토리지 장치에 저장되는 파일의 구조를 설명하기 위한 도면이다. 도 6은 노드 어드레스 테이블을 설명하기 위한 도면이다.

[0035] 우선, 도 1을 참조하면, 본 발명의 일 실시예에 따른 컴퓨팅 시스템(1)은 호스트(10)와 스토리지 장치(20)를 포함한다.

[0036] 호스트(10)와 스토리지 장치(20)는 특정한 프로토콜(protocol)을 이용하여 서로 통신한다. 예를 들어, USB(Universal Serial Bus) 프로토콜, MMC(multimedia card) 프로토콜, PCI(peripheral component interconnection) 프로토콜, PCI-E(PCI-express) 프로토콜, ATA(Advanced Technology Attachment) 프로토콜, Serial-ATA 프로토콜, Parallel-ATA 프로토콜, SCSI (small computer small interface) 프로토콜, ESDI(enhaned small disk interface) 프로토콜, 그리고 IDE(Integrated Drive Electronics) 프로토콜 등과 같은 다양한 인터페이스 프로토콜 중 적어도 하나를 통해 통신할 수 있으나, 이에 한정되는 것은 아니다.

[0037] 호스트(10)는 스토리지 장치(20)를 컨트롤한다. 예를 들어, 호스트(10)는 스토리지 장치(20)에 데이터를 라이트하거나, 스토리지 장치(20)로부터 데이터를 리드할 수 있다.

[0038] 도 2를 참조하면, 이러한 호스트(10)는 사용자 스페이스(user space)(11)와 커널 스페이스(kernel space)(13)를 포함한다.

[0039] 사용자 스페이스(11)는 사용자 어플리케이션(user application)(12)이 실행되는 영역이고, 커널 스페이스(13)는 커널 실행을 위해서 제한적으로 보장된(restrictively reserved) 영역이다. 사용자 스페이스(11)에서 커널 스페이스(13)를 접근하기 위해서, 시스템 콜(system call)이 이용될 수 있다.

[0040] 커널 스페이스(13)는 가상 파일시스템(14), 파일시스템(16), 장치 드라이버(18) 등을 포함할 수 있다. 파일시스템(16)은 하나 이상일 수 있다. 예를 들어, 파일시스템(16)은 ext2, ntfs, smbfs, proc 등일 수 있다. 또한, 본 발명의 일 실시예에 따른 컴퓨팅 시스템(1)은 F2FS 파일시스템을 포함할 수 있다. F2FS 파일시스템에 대해서는 도 3 내지 도 11을 이용하여 후술하도록 한다.

[0041] 가상 파일시스템(14)은 하나 이상의 파일시스템(16)이 서로 상호 동작할 수 있도록 한다. 서로 다른 미디어의 서로 다른 파일시스템(16)에 대해 리드/라이트 작업을 하기 위해서, 표준화된 시스템 콜을 사용할 수 있도록 한다. 예를 들어, open(), read(), write()와 같은 시스템 콜은, 파일시스템(16)의 종류에 관계없이 사용될 수 있다. 즉, 가상 파일시스템(14)은 사용자 스페이스(11)와 파일시스템(16) 사이에 존재하는 추상화 계층이다.

[0042] 장치 드라이버(18)는 하드웨어와 사용자 어플리케이션(또는 운영체제) 사이의 인터페이스를 담당한다. 장치 드라이버(18)는 하드웨어가 특정 운영체제 하에서 정상적으로 동작하기 위해 필요한 프로그램이다.

- [0043] 이하에서, F2FS 파일시스템이 스토리지 장치(20)를 어떻게 제어하는지를 설명한다.
- [0044] 스토리지 장치(20)는 SSD(Static Solid Disk), HDD(Hard Disk Drive), eMMC와 같은 각종 카드 스토리지, 데이터 서버 등일 수 있으나, 이에 한정되지 않는다.
- [0045] 스토리지 장치(20)는 도 3에 도시된 것과 같이, 구성될 수 있다. 세그먼트(SEGMENT)(53)는 다수의 블록(BLK)(51)을 포함하고, 섹션(SECTION)(55)은 다수의 세그먼트(53)를 포함하고, 존(ZONE)(57)은 다수의 섹션(55)을 포함할 수 있다. 예를 들어, 블록(51)은 4Kbyte이고, 세그먼트(53)는 512개의 블록(51)을 포함하여 2M byte일 수 있다. 이러한 구성은, 스토리지 장치(20)의 포맷(format) 시점에서 결정될 수 있으나, 이에 한정되는 것은 아니다. 섹션(55)과 존(57)의 사이즈는 포맷 시점에서 수정될 수도 있다. F2FS 파일시스템은 모든 데이터를 4Kbyte의 페이지 단위로 리드/라이트할 수 있다. 즉, 블록(51)에 하나의 페이지가 저장되고, 세그먼트(53)에 다수의 페이지가 저장될 수 있다.
- [0046] 한편, 스토리지 장치(20)에 저장되는 파일은 도 4에 도시된 것과 같은, 인덱싱 구조(indexing structure)를 가질 수 있다. 하나의 파일은 다수의 데이터와, 다수의 데이터와 연관된 다수의 노드를 포함할 수 있다. 데이터 블록(70)은 데이터를 저장하는 부분이고, 노드 블록(80, 81~88, 91~95)은 노드를 저장하는 부분이다.
- [0047] 노드 블록(80, 81~88, 91~95)은 디렉트 노드 블록(direct node block)(81~88), 인다이렉트 노드 블록(indirect node block)(91~95), 아이노드 블록(inode block)(80)을 포함할 수 있다.
- [0048] 디렉트 노드 블록(81~88)은 데이터 블록(70)을 직접 가리키는 데이터 포인터(data pointer)를 포함한다.
- [0049] 인다이렉트 노드 블록(91~95)은 데이터 블록(70)이 아닌, 다른 노드 블록(83~88)(즉, 하위의 노드 블록)을 가리키는 포인터를 포함한다. 인다이렉트 노드 블록(91~95)은 예를 들어, 제1 인다이렉트 노드 블록(91~94), 제2 인다이렉트 노드 블록(95) 등을 포함할 수 있다. 제1 인다이렉트 노드 블록(91~94)은 디렉트 노드 블록(83~88)을 가리키는 제1 노드 포인터를 포함한다. 제2 인다이렉트 노드 블록(95)은 제1 인다이렉트 노드 블록(93, 94)을 가리키는 제2 노드 포인터를 포함한다.
- [0050] 아이노드 블록(80)은 데이터 포인터, 디렉트 노드 블록(81, 82)를 가리키는 제1 노드 포인터, 제1 인다이렉트 노드 블록(91, 92)을 가리키는 제2 노드 포인터, 제2 인다이렉트 노드 블록(95)를 가리키는 제3 노드 포인터 중 적어도 하나를 포함할 수 있다. 하나의 파일은 예를 들어, 최대 3Tbyte일 수 있고, 이러한 대용량의 파일을 다음과 같은 인덱스 구조를 가질 수 있다. 예를 들어, 아이노드 블록(80) 내의 데이터 포인터는 994개이고, 994개의 데이터 포인터 각각은 994개의 데이터 블록(70) 각각을 가리킬 수 있다. 제1 노드 포인터는 2개이고, 2개의 제1 노드 포인터 각각은 2개의 디렉트 노드 블록(81, 82)를 가리킬 수 있다. 제2 노드 포인터는 2개이고, 2개의 제2 노드 포인터 각각은 2개의 제1 인다이렉트 노드 블록(91, 92)을 가리킬 수 있다. 제3 노드 포인터는 1개이고, 제2 인다이렉트 노드 블록(95)를 가리킬 수 있다.
- [0051] 또한, 파일별로 아이노드 메타데이터를 포함하는 아이노드 페이지가 존재한다.
- [0052] 한편, 도 5에서와 같이, 본 발명의 일 실시예에 따른 컴퓨팅 시스템(1)에서, 스토리지 장치(20)는 메타데이터 영역(I)과 메인 영역(II)으로 나누어질 수 있다. 파일시스템(16)은 포맷(format)할 때, 스토리지 장치(20)를 메타데이터 영역(I)과 메인 영역(II)으로 나눌 수 있으나, 이에 한정되는 것은 아니다. 메타데이터 영역(I)은 메타데이터가 저장되는 공간이다. 메타데이터는 시스템 전체로 관리되는 각종 정보를 지칭하는 것으로, 예를 들어, 현재 할당된 파일 수, 유효한 페이지 수, 위치 등의 정보를 나타낼 수 있다. 메인 영역(II)은 실제 사용자가 사용하고 있는 각종 디렉토리 정보, 데이터, 파일 정보 등을 저장하는 공간이다.
- [0053] 또한, 메타데이터 영역(I)은 스토리지 장치(20)의 앞부분에 저장되고, 메인 영역(II)은 스토리지 장치(20)의 뒷부분에 저장될 수 있다. 여기서, 앞부분은 뒷부분보다 물리 어드레스(physical address)를 기준으로 앞에 있음을 의미한다.
- [0054] 구체적으로, 메타데이터 영역(I)은 슈퍼블록(61, 62), 체크 포인트 영역(CheckPoint area, CP)(63), 세그먼트 정보 테이블(Segment Information Table, SIT)(64), 노드 어드레스 테이블(Node Address Table, NAT)(65), 세그먼트 요약 영역(Segment Summary Area, SSA)(66) 등을 포함할 수 있다.
- [0055] 먼저, 슈퍼블록(61, 62)에는, 파일시스템(16)의 디폴트 정보가 저장된다. 예를 들어, 블록(51)의 크기, 블록(51)의 개수, 파일시스템(16)의 상태 플러그(clean, stable, active, logging, unknown) 등이 저장될 수 있다. 도시된 것과 같이, 슈퍼블록(61, 62)은 2개일 수 있고, 각각에는 동일한 내용이 저장될 수 있다. 따라서, 둘 중

어느 하나에 문제가 발생하더라도, 다른 하나를 이용할 수 있다.

[0056] 체크 포인트 영역(63)은 체크 포인트를 저장한다. 체크 포인트는 논리적인 중단점으로서, 이러한 중단점까지의 상태가 완전하게 보존된다. 컴퓨팅 시스템의 동작 중에 사고(예를 들어, 셧다운(shutdown))가 발생하면, 파일시스템(16)은 보존된 체크 포인트를 이용하여 데이터를 복구할 수 있다. 이러한 체크 포인트의 생성 시점은, 예를 들어, 주기적으로 생성, Umount 시점, System shutdown 시점 등일 수 있으나, 이에 한정되는 것은 아니다.

[0057] 노드 어드레스 테이블(65)은 도 6에 도시된 것과 같이, 노드 각각에 대응되는 다수의 노드 식별자(NODE ID)와, 다수의 노드 식별자 각각에 대응되는 다수의 물리 어드레스를 포함할 수 있다. 예를 들어, 노드 식별자 N0에 대응되는 노드 블록은 물리 어드레스 a에 대응되고, 노드 식별자 N1에 대응되는 노드 블록은 물리 어드레스 b에 대응되고, 노드 식별자 N2에 대응되는 노드 블록은 물리 어드레스 c에 대응될 수 있다. 모든 노드(아이노드, 디렉트 노드, 인다이렉트 노드 등)는 각각 고유의 노드 식별자를 갖는다. 다르게 설명하면, 모든 노드(아이노드, 디렉트 노드, 인다이렉트 노드 등)는 노드 어드레스 테이블(65)로부터 고유의 노드 식별자를 할당 받을 수 있다. 노드 어드레스 테이블(65)은 아이노드의 노드 식별자, 디렉트 노드의 노드 식별자 및 인다이렉트 노드의 노드 식별자 등을 저장할 수 있다. 각 노드 식별자에 대응되는 각 물리 어드레스는 업데이트(update)될 수 있다.

[0058] 세그먼트 정보 테이블(64)은 각 세그먼트의 유효한 페이지의 개수와, 다수의 페이지의 비트맵을 포함한다. 비트맵은 각 페이지가 유효한지 여부를 0 또는 1로 표시한 것을 의미한다. 세그먼트 정보 테이블(64)은 클리닝(cleaning) 작업(또는 가비지 컬렉션(garbage collection))에서 사용될 수 있다. 특히 비트맵은 클리닝 작업을 수행할 때, 불필요한 리드 요청을 줄일 수 있고, 어댑티브 데이터 로깅(adaptive data logging) 시 블록 할당할 때 이용될 수 있다.

[0059] 세그먼트 요약 영역(66)은 메인 영역의 각 세그먼트의 요약 정보를 모아둔 영역이다. 구체적으로, 세그먼트 요약 영역(66)은 메인 영역의 각 세그먼트의 다수의 블록이 속하는 노드의 정보를 기술한다. 세그먼트 요약 영역(66)은 클리닝(cleaning) 작업(또는 가비지 컬렉션(garbage collection))에서 사용될 수 있다. 구체적으로 설명하면, 노드 블록(80, 81~88, 91~95)은 데이터 블록(70) 또는 하위의 노드 블록(예를 들어, 디렉트 노드 블록 등)의 위치를 확인하기 위해, 노드 식별자 목록 또는 어드레스를 가지고 있다. 이와는 반대로, 세그먼트 요약 영역(66)은 데이터 블록(70) 또는 하위의 노드 블록(80, 81~88, 91~95)이 상위의 노드 블록(80, 81~88, 91~95)의 위치를 확인할 수 있는 인덱스(index)를 제공한다. 세그먼트 요약 영역(66)은 다수의 세그먼트 요약 블록을 포함한다. 하나의 세그먼트 요약 블록은, 메인 영역에 위치한 하나의 세그먼트에 대한 정보를 가지고 있다. 또한, 세그먼트 요약 블록은 다수의 요약 정보로 구성되어 있으며, 하나의 요약 정보는 하나의 데이터 블록 또는 하나의 노드 블록에 대응된다.

[0060] 또한, 메인 영역(II)은 서로 분리된 데이터 세그먼트(data segment)(DS0, DS1)과 노드 세그먼트(node segment)(NS0, NS1)을 포함할 수 있다. 다수의 데이터는 데이터 세그먼트(DS0, DS1)에 저장되고, 다수의 노드는 노드 세그먼트(NS0, NS1)에 저장될 수 있다. 데이터와 노드가 분리되는 영역이 서로 다르면, 효율적으로 세그먼트를 관리할 수 있고, 데이터를 리드할 때 보다 효과적으로 빠른 시간 내에 리드할 수 있다.

[0061] 한편, 메인 영역(II)은 순차적 접근(sequential access) 방식으로 라이트되는 영역이고, 메타데이터 영역(I)은 랜덤 접근(random access) 방식으로 라이트되는 영역일 수 있다. 메인 영역(II)은 스토리지 장치(20)의 뒷부분에 저장될 수 있고, 메타데이터 영역(I)은 스토리지 장치(20)의 앞부분에 저장된다.

[0062] 스토리지 장치(20)가 예를 들어, SSD(Static Solid Disk)인 경우, SSD 내부에 버퍼(buffer)가 있을 수 있다. 버퍼는 예를 들어, 리드/라이트 속도가 빠른 SLC(Single Layer Cell) 메모리일 수 있다. 따라서, 이러한 버퍼는 한정된 공간의 랜덤 접근 방식의 라이트 속도를 빠르게 할 수 있다. 따라서, 이러한 버퍼를 활용해서, 메타데이터 영역(I)을 스토리지 장치(20)의 앞부분에 위치시킴으로써 성능 저하를 막을 수 있다.

[0063] 도면에서는, 메타데이터 영역(I)은 슈퍼블록(61, 62), 체크 포인트 영역(62), 세그먼트 정보 테이블(64), 노드 어드레스 테이블(65), 세그먼트 요약 영역(66) 순서로 되어 있으나, 이에 한정되는 것은 아니다. 예를 들어, 세그먼트 정보 테이블(64)과 노드 어드레스 테이블(65)의 위치가 바뀌어도 무방하고, 노드 어드레스 테이블(65)과 세그먼트 요약 영역(66)의 위치가 바뀌어도 무방하다.

[0064] 이하에서, 도 7 및 도 8을 이용하여, 본 발명의 일 실시예에 따른 컴퓨팅 시스템의 데이터 관리 방법을 설명하도록 한다. 도 7 및 도 8은 본 발명의 일 실시예에 따른 컴퓨팅 시스템의 데이터 관리 방법을 설명하기 위한 개념도이다.

- [0065] 도 7을 참조하면, 파일시스템(16)은 스토리지 장치를 메타데이터 영역(I)과 메인 영역(II)으로 나눈다. 전술한 것과 같이, 포맷 시점에서 메타데이터 영역(I)과 메인 영역(II)로 나누어질 수 있다.
- [0066] 파일시스템(16)은 도 4를 이용하여 설명한 것과 같이, 하나의 파일을 다수의 데이터와, 다수의 데이터와 관련된 다수의 노드(예를 들어, 아이노드, 디렉트 노드, 인다이렉트 노드 등)로 구성하여, 스토리지 장치(20) 내에 저장할 수 있다. 이 때, 모든 노드는, 노드 어드레스 테이블(65)로부터 노드 식별자(NODE ID)를 할당받는다. 예를 들어, N0~N5를 할당받았다고 가정하자. N0~N5에 대응되는 노드 블록은 각각 물리 어드레스 a, b, … d에 대응될 수 있다. 도 7에 도시된 해칭된 부분은, 메인 영역(II) 내에서 다수의 데이터, 다수의 노드가 라이트된 부분을 의미한다.
- [0067] 예를 들어, 디렉트 노드 N5는 DATA10을 가리킬 수 있다. 디렉트 노드 N5는 물리 어드레스 d에 해당하는 노드 블록에 저장되어 있다. 즉, 노드 어드레스 테이블(65) 내에서, 디렉트 노드 N5는 물리 어드레스 d에 대응되는 것으로 저장되어 있다.
- [0068] 도 8을 참조하여, 파일 내의 일부 데이터 DATA10을 DATA10a로 수정하는 경우를 설명하면 다음과 같다.
- [0069] 메인 영역(II)은 순차적 접근(sequential access) 방식으로 라이트되는 영역이다. 따라서, 수정된 데이터 DATA10a는 비어있는 데이터 블록 내에 새롭게 저장된다. 또한, 디렉트 노드 N5는, 수정된 데이터 DATA10a가 저장된 데이터 블록을 가리키도록 수정하여, 비어있는 노드 블록 내에 새롭게 저장된다.
- [0070] 메타데이터 영역은 랜덤 접근(random access) 방식으로 라이트되는 영역이다. 따라서, 디렉트 노드 N5는 물리 어드레스 f에 대응되는 것으로, 오버라이트(overwrite)한다.
- [0071] 파일 내의 일부 데이터를 수정하는 것을 정리하면, 다수의 데이터 중에서 제1 데이터는 제1 물리 어드레스에 대응되는 제1 블록 내에 저장되고, 제1 디렉트 노드는 제1 데이터를 지시하고 제1 디렉트 노드는 제2 물리 어드레스에 대응되는 제2 블록 내에 저장된다. 노드 어드레스 테이블 내에서, 제1 디렉트 노드는 제2 물리 어드레스에 대응되는 것으로 저장되어 있다. 여기서, 제1 데이터를 수정하여 제2 데이터를 생성한다. 제2 데이터를 제1 물리 어드레스와 다른 제3 물리 어드레스에 대응되는 제3 블록 내에 라이트한다. 제1 디렉트 노드가 제2 데이터를 지시하도록 수정하여, 제2 물리 어드레스와 다른 제4 물리 어드레스에 대응되는 제4 블록 내에 라이트 한다. 또한, 노드 어드레스 테이블 내에, 제1 디렉트 노드는 제4 물리 어드레스에 대응되는 것으로 오버라이트(overwrite)한다.
- [0072] 로그 구조화된 파일시스템에서, 노드 어드레스 테이블(65)을 이용하면, 파일의 일부 데이터를 수정할 때, 수정해야 하는 데이터 및 노드의 양을 최소화할 수 있다. 즉, 수정된 데이터와, 수정된 데이터를 직접 가리키는 디렉트 노드만 순차적 접근 방식으로 라이트하고, 디렉트 노드를 가리키는 인다이렉트 노드 또는 아이노드 등은 수정할 필요가 없다. 노드 어드레스 테이블(65) 내에서, 디렉트 노드에 대응되는 물리 어드레스를 수정하였기 때문이다.
- [0073] 도 9는 본 발명의 다른 실시예에 따른 컴퓨팅 시스템의 스토리지 장치에 저장되는 파일의 구조를 설명하기 위한 블록도이다. 이하에서 설명의 편의를 위해서, 도 1 내지 도 8을 이용하여 설명한 내용과 다른 점을 위주로 설명한다.
- [0074] 도 9를 참조하면, 본 발명의 다른 실시예에 따른 컴퓨팅 시스템의 스토리지 장치에서, 메인 영역(II)은 서로 분리된 다수의 세그먼트(S1~Sn, 단, n은 자연수)를 포함할 수 있다. 각 세그먼트(S1~Sn)에는, 데이터와 노드 구분 없이 저장될 수 있다.
- [0075] 반면, 본 발명의 일 실시예에 따른 컴퓨팅 시스템에서, 스토리지 장치는 서로 분리된 데이터 세그먼트(DS0, DS1), 노드 세그먼트(NS0, NS1)를 포함한다. 다수의 데이터는 데이터 세그먼트(DS0, DS1)에 저장되고, 다수의 노드는 노드 세그먼트(NS0, NS1)에 저장될 수 있다.
- [0076] 도 10은 본 발명의 또 다른 실시예에 따른 컴퓨팅 시스템의 스토리지 장치에 저장되는 파일의 구조를 설명하기 위한 블록도이다. 이하에서 설명의 편의를 위해서, 도 1 내지 도 8을 이용하여 설명한 내용과 다른 점을 위주로 설명한다.
- [0077] 도 10을 참조하면, 본 발명의 또 다른 실시예에 따른 컴퓨팅 시스템의 스토리지 장치에서, 메타데이터 영역(I)은 세그먼트 요약 영역(도 5의 66 참조)을 포함하지 않는다. 즉, 메타데이터 영역(I)은 슈퍼블록(61, 62), 체크포인트 영역(62), 세그먼트 정보 테이블(64), 노드 어드레스 테이블(65)을 포함한다.

- [0078] 세그먼트 요약 정보는 메인 영역(II)내에 저장될 수 있다. 구체적으로, 메인 영역(II)은 다수의 세그먼트(S0~Sn)를 포함하고, 각 세그먼트(S0~Sn)는 다수의 블록으로 구분된다. 각 세그먼트(S0~Sn)의 적어도 하나의 블록(SS0~SSn)에 세그먼트 요약 정보를 저장할 수 있다.
- [0079] 도 11은 본 발명의 또 다른 실시예에 따른 컴퓨팅 시스템의 스토리지 장치에 저장되는 파일의 구조를 설명하기 위한 블록도이다. 이하에서 설명의 편의를 위해서, 도 1 내지 도 8을 이용하여 설명한 내용과 다른 점을 위주로 설명한다.
- [0080] 도 11을 참조하면, 본 발명의 또 다른 실시예에 따른 컴퓨팅 시스템의 스토리지 장치에서, 메타데이터 영역(I)은 세그먼트 요약 영역(도 5의 66 참조)을 포함하지 않는다. 즉, 메타데이터 영역(I)은 슈퍼블록(61, 62), 체크 포인트 영역(62), 세그먼트 정보 테이블(64), 노드 어드레스 테이블(65)을 포함한다.
- [0081] 세그먼트 요약 정보는 메인 영역(II)내에 저장될 수 있다. 메인 영역은 다수의 세그먼트(53)를 포함하고, 각 세그먼트(53)는 다수의 블록(BLK0~BLKm)으로 구분되고, 각 블록(BLK0~BLKm)은 OOB(Out Of Band)(OOB1~OOBm, 단, m은 자연수) 영역을 포함할 수 있다. OOB 영역(OOB1~OOBm)에 세그먼트 요약 정보를 저장할 수 있다.
- [0082] 이하에서, 본 발명의 몇몇 실시예에 따른 컴퓨팅 시스템이 적용될 수 있는 구체적인 시스템을 설명한다. 이하에서 설명되는 시스템은 예시적인 것에 불과하고, 이에 한정되는 것은 아니다.
- [0083] 도 12는 본 발명의 몇몇 실시예에 따른 컴퓨팅 시스템의 구체적인 일 예를 설명하기 위한 블록도이다.
- [0084] 도 12를 참조하면, 호스트 서버(300)는 네트워크(20)를 통해서 다수의 데이터 베이스 서버(330, 340, 350, 360)와 접속되어 있다. 호스트 서버(300) 내에, 데이터 베이스 서버(330, 340, 350, 360)의 데이터를 관리하기 위한 파일시스템(316)을 설치될 수 있다. 파일시스템(316)은 도 1 내지 도 11을 이용하여 설명한 F2FS 파일시스템일 수 있다.
- [0085] 도 13 내지 도 15는 본 발명의 몇몇 실시예에 따른 컴퓨팅 시스템의 구체적인 다른 예를 설명하기 위한 블록도이다.
- [0086] 우선, 도 13을 참조하면, 스토리지 장치(100)(도 1의 20에 대응됨)는 비휘발성 메모리 장치(1100) 및 컨트롤러(1200)를 포함할 수 있다.
- [0087] 여기서, 비휘발성 메모리 장치(1100)에는, 전술한 슈퍼블록(61, 62), 체크 포인트 영역(63), 세그먼트 정보 테이블(64), 노드 어드레스 테이블(65) 등이 저장되어 있을 수 있다.
- [0088] 컨트롤러(1200)는 호스트 및 비휘발성 메모리 장치(1100)에 연결된다. 호스트(Host)로부터의 요청에 응답하여, 컨트롤러(1200)는 비휘발성 메모리 장치(1100)를 액세스하도록 구성된다. 예를 들면, 컨트롤러(1200)는 비휘발성 메모리 장치(1100)의 리드, 라이트, 이레이즈, 그리고 백그라운드(background) 동작을 제어하도록 구성된다. 컨트롤러(1200)는 비휘발성 메모리 장치(1100) 및 호스트(Host) 사이에 인터페이스를 제공하도록 구성된다. 컨트롤러(1200)는 비휘발성 메모리 장치(1100)를 제어하기 위한 펌웨어(firmware)를 구동하도록 구성된다.
- [0089] 예시적으로, 컨트롤러(1200)는 램(RAM, Random Access Memory), 프로세싱 유닛(processsing unit), 호스트 인터페이스(host interface), 그리고 메모리 인터페이스(memory interface)와 같은 잘 알려진 구성 요소들을 더 포함한다. 램(RAM)은 프로세싱 유닛의 동작 메모리, 비휘발성 메모리 장치(1100) 및 호스트(Host) 사이의 캐시 메모리, 그리고 비휘발성 메모리 장치(1100) 및 호스트(Host) 사이의 버퍼 메모리 중 적어도 하나로서 이용된다. 프로세싱 유닛은 컨트롤러(1200)의 제반 동작을 제어한다.
- [0090] 컨트롤러(1200) 및 비휘발성 메모리 장치(1100)는 하나의 반도체 장치로 집적될 수 있다. 예시적으로, 컨트롤러(1200) 및 비휘발성 메모리 장치(1100)는 하나의 반도체 장치로 집적되어, 메모리 카드를 구성할 수 있다. 예를 들면, 컨트롤러(1200) 및 비휘발성 메모리 장치(1100)는 하나의 반도체 장치로 집적되어 PC 카드(PCMCIA, personal computer memory card international association), 컴팩트 플래시 카드(CF), 스마트 미디어 카드(SM, SMC), 메모리 스틱, 멀티미디어 카드(MMC, RS-MMC, MMCmicro), SD 카드(SD, miniSD, microSD, SDHC), 유니버설 플래시 기억장치(UFS) 등과 같은 메모리 카드를 구성할 것이다.
- [0091] 컨트롤러(1200) 및 비휘발성 메모리 장치(1100)는 하나의 반도체 장치로 집적되어 SSD(Solid State Drive)를 구성할 수 있다. SSD는 반도체 메모리에 데이터를 저장하도록 구성되는 저장 장치를 포함한다. 시스템(1000)이 반도체 드라이브(SSD)로 이용되는 경우, 시스템(1000)에 연결된 호스트(Host)의 동작 속도는 획기적으로 개선될 수 있다.

[0092] 다른 예로서, 시스템(1000)은 컴퓨터, UMPC (Ultra Mobile PC), 워크스테이션, 넷북(net-book), PDA (Personal Digital Assistants), 포터블(portable) 컴퓨터, 웹 태블렛(web tablet), 무선 전화기(wireless phone), 모바일 폰(mobile phone), 스마트폰(smart phone), e-북(e-book), PMP(portable multimedia player), 휴대용 게임기, 네비게이션(navigation) 장치, 블랙박스(black box), 디지털 카메라(digital camera), 3차원 수상기(3-dimensional television), 디지털 음성 녹음기(digital audio recorder), 디지털 음성 재생기(digital audio player), 디지털 영상 녹화기(digital picture recorder), 디지털 영상 재생기(digital picture player), 디지털 동영상 녹화기(digital video recorder), 디지털 동영상 재생기(digital video player), 정보를 무선 환경에서 송수신할 수 있는 장치, 홈 네트워크를 구성하는 다양한 전자 장치들 중 하나, 컴퓨터 네트워크를 구성하는 다양한 전자 장치들 중 하나, 텔레매틱스 네트워크를 구성하는 다양한 전자 장치들 중 하나, RFID 장치, 또는 컴퓨팅 시스템을 구성하는 다양한 구성 요소들 중 하나 등과 같은 전자 장치의 다양한 구성 요소들 중 하나로 제공된다.

[0093] 예시적으로, 비휘발성 메모리 장치(1100) 또는 시스템(1000)은 다양한 형태들의 패키지로 실장될 수 있다. 예를 들면, 비휘발성 메모리 장치(1100) 또는 시스템(1000)은 PoP(Package on Package), Ball grid arrays(BGAs), Chip scale packages(CSPs), Plastic Leaded Chip Carrier(PLCC), Plastic Dual In Line Package(PDIP), Die in Waffle Pack, Die in Wafer Form, Chip On Board(COB), Ceramic Dual In Line Package(CERDIP), Plastic Metric Quad Flat Pack(MQFP), Thin Quad Flatpack(TQFP), Small Outline(SOIC), Shrink Small Outline Package(SSOP), Thin Small Outline(TSOP), Thin Quad Flatpack(TQFP), System In Package(SIP), Multi Chip Package(MCP), Wafer-level Fabricated Package(WFP), Wafer-Level Processed Stack Package(WSP) 등과 같은 방식으로 패키지화되어 실장될 수 있다.

[0094] 이어서, 도 14을 참조하면, 시스템(2000)은 비휘발성 메모리 장치(2100) 및 컨트롤러(2200)를 포함한다. 비휘발성 메모리 장치(2100)는 복수의 비휘발성 메모리 칩들을 포함한다. 복수의 비휘발성 메모리 칩들은 복수의 그룹들로 분할된다. 복수의 비휘발성 메모리 칩들의 각 그룹은 하나의 공통 채널을 통해 컨트롤러(2200)와 통신하도록 구성된다. 예를 들어, 복수의 비휘발성 메모리 칩들은 제 1 내지 제 k 채널들(CH1~CHk)을 통해 컨트롤러(2200)와 통신하는 것으로 도시되어 있다.

[0095] 도 14에서, 하나의 채널에 복수의 비휘발성 메모리 칩들이 연결되는 것으로 설명되었다. 그러나, 하나의 채널에 하나의 비휘발성 메모리 칩이 연결되도록 시스템(2000)이 변형될 수 있음이 이해될 것이다.

[0096] 이어서, 도 15을 참조하면, 시스템(3000)은 중앙 처리 장치(3100), 램(3200, RAM, Random Access Memory), 사용자 인터페이스(3300), 전원(3400), 그리고 도 14의 시스템(2000)을 포함한다.

[0097] 시스템(2000)은 시스템 버스(3500)를 통해, 중앙처리장치(3100), 램(3200), 사용자 인터페이스(3300), 그리고 전원(3400)에 전기적으로 연결된다. 사용자 인터페이스(3300)를 통해 제공되거나, 중앙 처리 장치(3100)에 의해서 처리된 데이터는 시스템(2000)에 저장된다.

[0098] 도 15에서, 비휘발성 메모리 장치(2100)는 컨트롤러(2200)를 통해 시스템 버스(3500)에 연결되는 것으로 도시되어 있다. 그러나, 비휘발성 메모리 장치(2100)는 시스템 버스(3500)에 직접 연결되도록 구성될 수 있다.

[0099] 이상 첨부된 도면을 참조하여 본 발명의 실시예를 설명하였지만, 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자는 본 발명이 그 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.

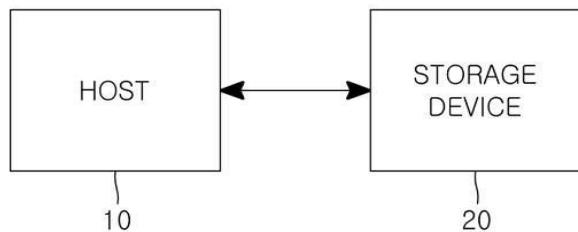
### 부호의 설명

10: 호스트	20: 스토리지 장치
61, 62: 슈퍼블록	63: 체크 포인트 블록
64: 세그먼트 정보 테이블	65: 노드 어드레스 테이블
66: 세그먼트 요약 영역	

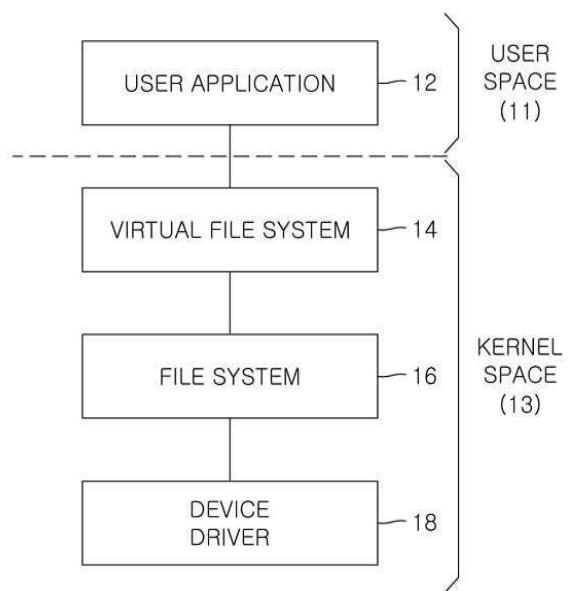
도면

도면1

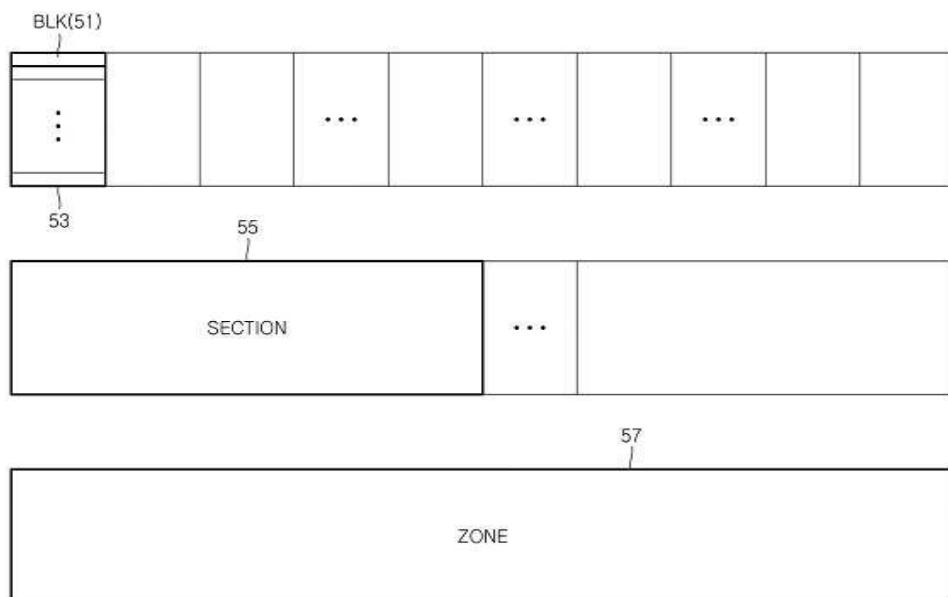
1



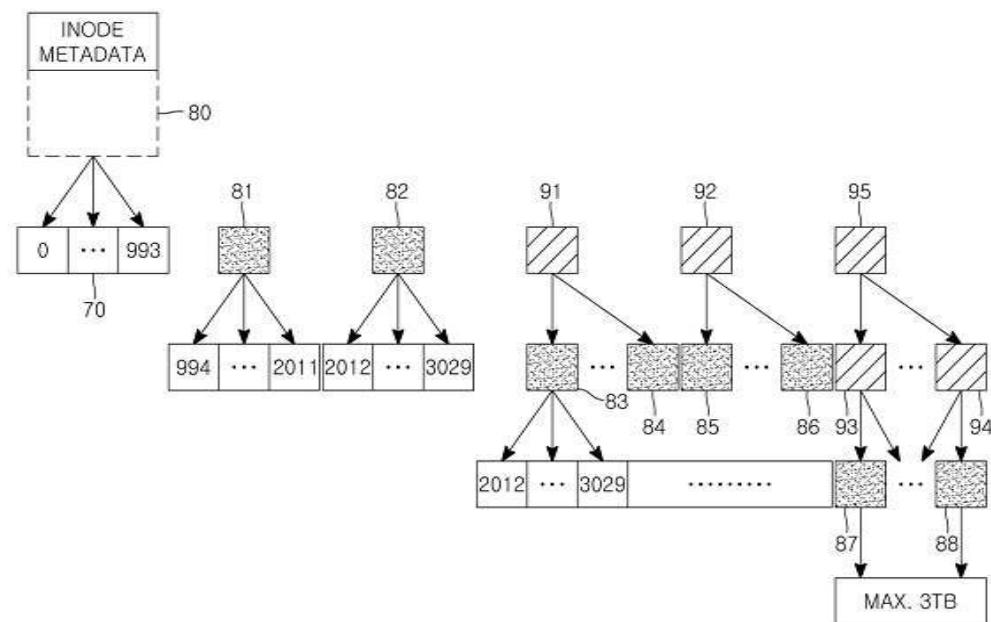
도면2



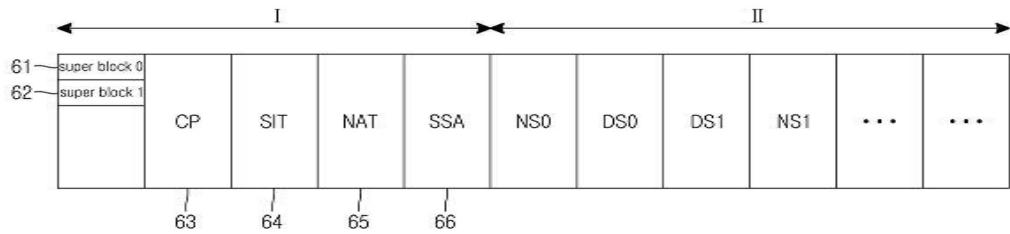
## 도면3



## 도면4



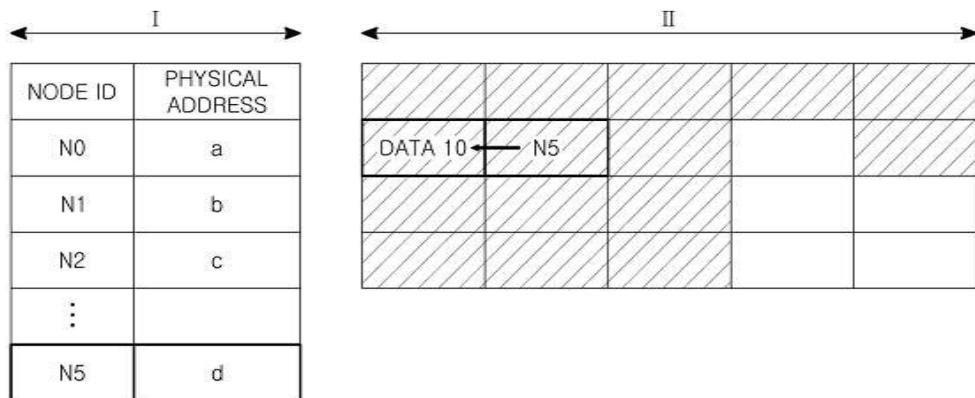
## 도면5



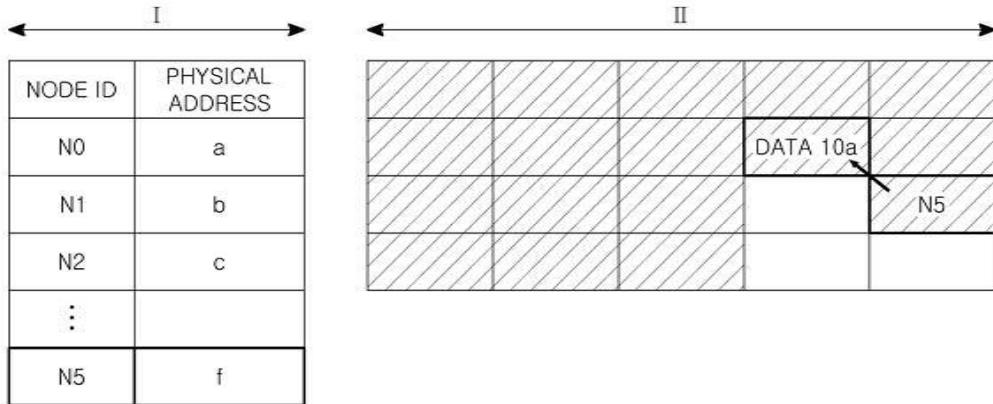
## 도면6

NODE ID	PHYSICAL ADDRESS
N0	a
N1	b
N2	c
:	

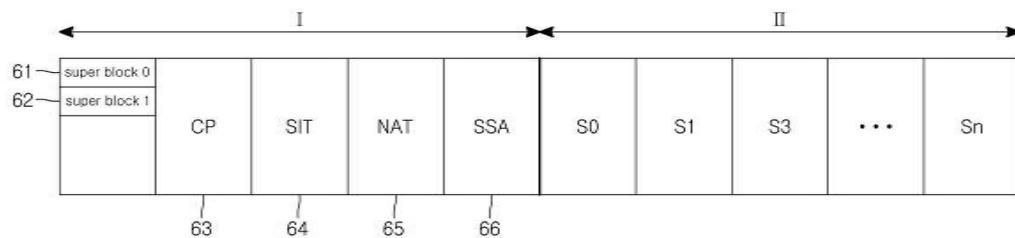
## 도면7



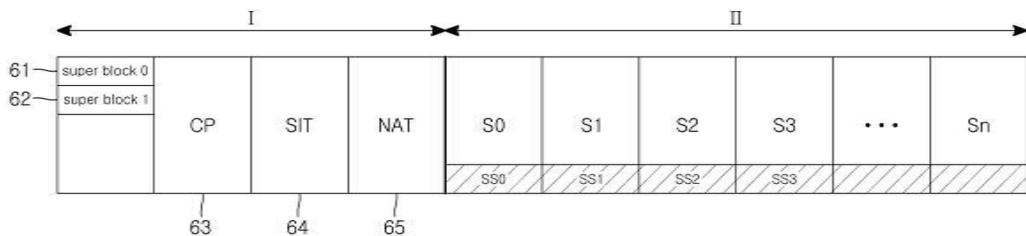
## 도면8



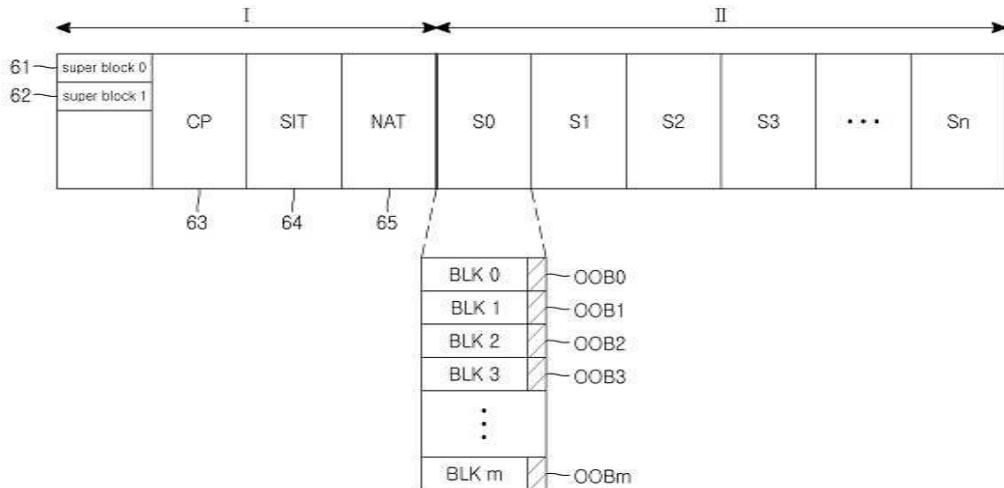
## 도면9



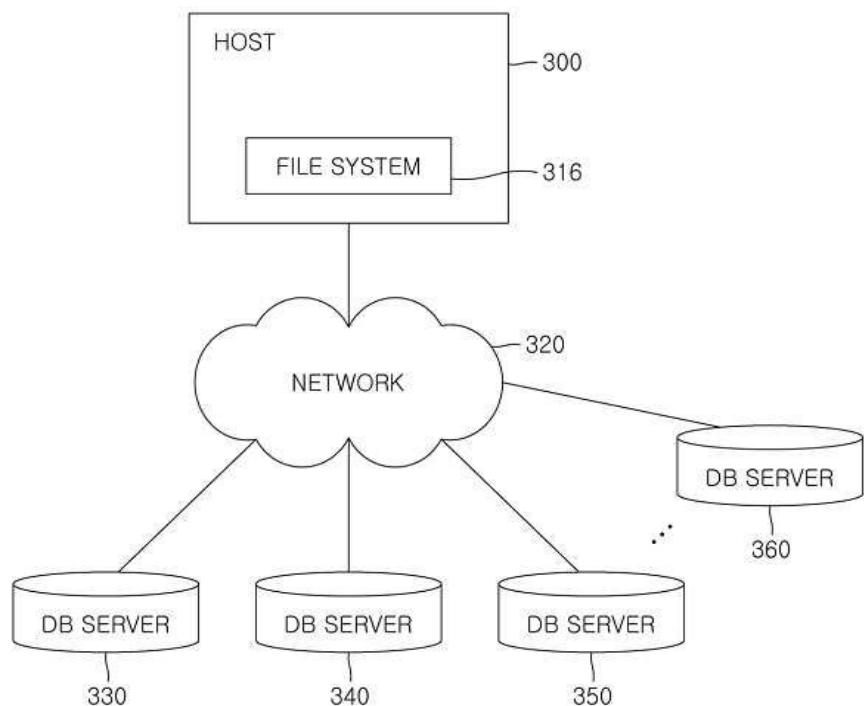
## 도면10



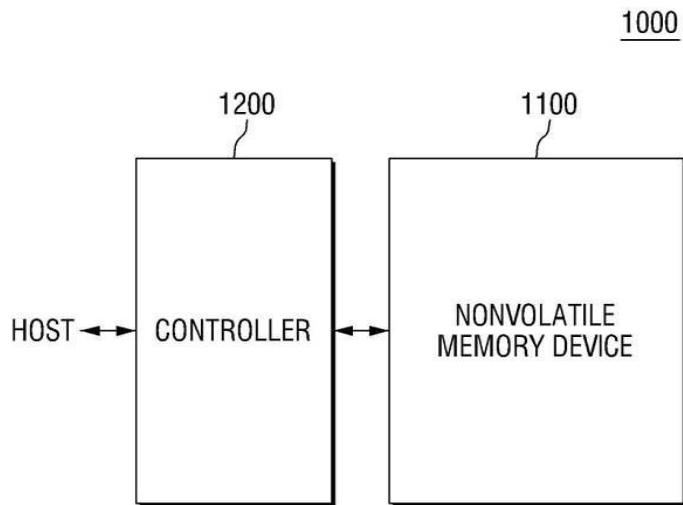
도면11



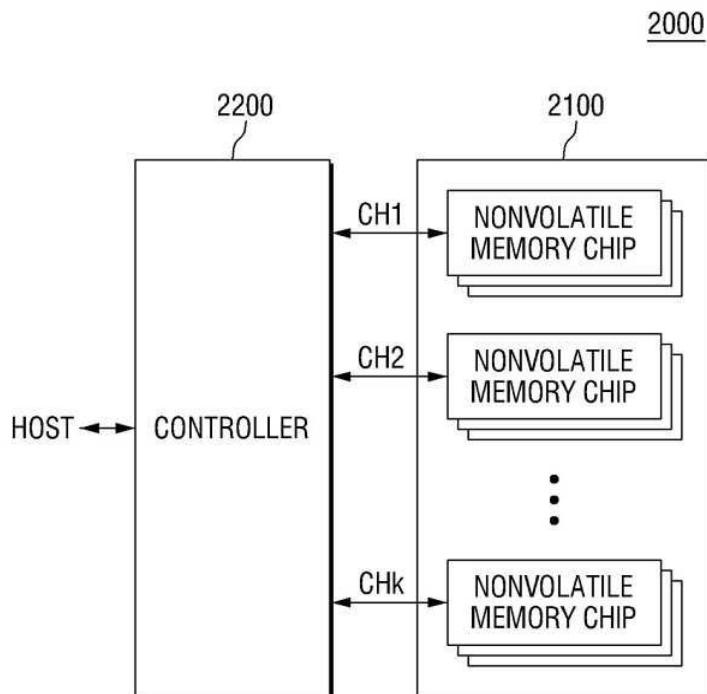
도면12



도면13



도면14



도면15

