



(19) 대한민국특허청(KR)

(12) 등록특허공보(B1)

(45) 공고일자 2015년09월16일

(11) 등록번호 10-1553532

(24) 등록일자 2015년09월10일

- (51) 국제특허분류(Int. Cl.)  
*G06F 9/06* (2006.01) *G06F 12/00* (2006.01)  
*G06F 3/06* (2006.01)
- (21) 출원번호 10-2008-0035475
- (22) 출원일자 2008년04월17일  
 심사청구일자 2013년04월17일
- (65) 공개번호 10-2009-0109959
- (43) 공개일자 2009년10월21일
- (56) 선행기술조사문헌  
 KR1020060100684 A\*  
 JP08153014 A\*  
 KR1020070060301 A  
 US5873085 A  
 \*는 심사관에 의하여 인용된 문헌

- (73) 특허권자  
 삼성전자주식회사  
 경기도 수원시 영통구 삼성로 129 (매탄동)
- (72) 발명자  
 임근수  
 경기도 용인시 수지구 용구대로 2742, 동성1차아파트 103동 203호 (죽전동)
- (74) 대리인  
 특허법인 무한

전체 청구항 수 : 총 10 항

심사관 : 박승철

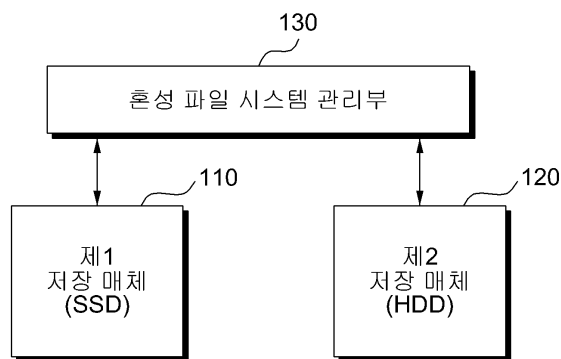
(54) 발명의 명칭 **스토리지 장치**

**(57) 요약**

스토리지 장치가 제공된다. 본 발명의 스토리지 장치는 제1 저장 매체, 상기 제1 저장 매체의 입출력 속도보다 느린 입출력 속도를 가지는 제2 저장 매체, 및 혼성 파일 시스템 관리부를 포함하고, 상기 혼성 파일 시스템 관리부는 상기 제1 저장 매체를 제어하는 제1 물리 파일 시스템 및 상기 제2 저장 매체를 제어하는 제2 물리 파일 시스템을 관리하며, 상기 제1 물리 파일 시스템 및 상기 제2 물리 파일 시스템으로부터 변환된 가상 파일 시스템을 호스트에 제공한다. 이를 통해 비휘발성 메모리에 최적화된 파일 시스템을 이용하는 스토리지 장치를 제공할 수 있다.

**대표도** - 도1

100



## 명세서

### 청구범위

#### 청구항 1

제1 저장 매체;

상기 제1 저장 매체의 입출력 속도보다 느린 입출력 속도를 가지는 제2 저장 매체; 및

상기 제1 저장 매체를 제어하는 제1 물리 파일 시스템 및 상기 제2 저장 매체를 제어하는 제2 물리 파일 시스템을 관리하며, 상기 제1 물리 파일 시스템 및 상기 제2 물리 파일 시스템으로부터 변환된 가상 파일 시스템을 가상 파일 시스템 관리자에 제공하는 혼성 파일 시스템 관리부

를 포함하고,

상기 혼성 파일 시스템 관리부는 상기 제1 물리 파일 시스템 상의 파일과 상기 제2 물리적 파일 시스템 상의 파일을 상기 가상 파일 시스템에서 이용될 수 있는 포맷으로 변환하고, 상기 가상 파일 시스템 상의 파일을 상기 제1 물리 파일 시스템 또는 상기 제2 물리 파일 시스템에서 이용될 수 있는 포맷으로 변환하는 컴퓨팅 시스템.

#### 청구항 2

제1항에 있어서,

상기 혼성 파일 시스템 관리부는

운영 체제로부터 요청된 파일의 접근 빈도에 기초하여 상기 요청된 파일을 상기 제1 저장 매체로부터 상기 제2 저장 매체로 이동하거나 상기 제2 저장 매체로부터 상기 제1 저장 매체로 이동할지 여부를 결정하는 컴퓨팅 시스템.

#### 청구항 3

제2항에 있어서,

상기 혼성 파일 시스템 관리부는

상기 요청된 파일이 상기 제2 저장 매체에 저장된 경우, 상기 요청된 파일의 접근 빈도에 기초하여 상기 요청된 파일의 상기 제1 저장 매체로의 이동에 따른 상대적 이득을 계산하고, 상기 요청된 파일의 상기 제1 저장 매체로의 이동에 따른 비용을 계산하고, 상기 계산된 상대적 이득이 상기 계산된 비용보다 크면 상기 요청된 파일을 상기 제1 저장 매체로 이동하는 컴퓨팅 시스템.

#### 청구항 4

제3항에 있어서,

상기 혼성 파일 시스템 관리부는

상기 요청된 파일을 상기 제1 저장 매체로 이동하기 전에 상기 제1 저장 매체에 사용 가능한 공간이 남아 있는지 판정하고, 상기 제1 저장 매체에 사용 가능한 공간이 남아 있지 않으면 상기 제1 저장 매체에 저장된 파일 중 비용 대비 효과가 가장 낮은 파일을 선택하고, 상기 선택된 파일을 상기 제2 저장 매체로 이동하고, 상기 요청된 파일을 상기 제2 저장 매체로부터 상기 제1 저장 매체로 이동하는 컴퓨팅 시스템.

#### 청구항 5

제4항에 있어서,

제1 저장매체는 SSD이고,

제2 저장 매체는 HDD이고,

상기 혼성 파일 시스템 관리부는 상기 SSD를 관리하는 제1 물리 파일 시스템 및 상기 HDD를 관리하는 제2 물리 파일 시스템을 관리하고, 상기 제1 물리 파일 시스템 및 상기 제2 물리 파일 시스템으로부터 변환된 가상 파일

시스템을 가상 파일 시스템 관리자에 제공하는 것  
을 특징으로 하는 컴퓨팅 시스템.

**청구항 6**

제4항에 있어서,  
상기 혼성 파일 시스템 관리부는 요청된 파일이 제2 저장 매체에 저장되어 있으면, 제1 저장 매체로부터 호스트로의 데이터 전송 시간 및 제2 저장 매체로부터 호스트로의 데이터 전송 시간의 차이를 상기 요청된 파일의 접근 빈도에 곱하여 상대적 이득을 계산하는 것  
을 특징으로 하는 컴퓨팅 시스템.

**청구항 7**

제4항에 있어서,  
상기 혼성 파일 시스템 관리부는 요청된 파일이 제2 저장 매체에 저장되어 있으면, 일정한 주기 내에서 발생할 수 있는 요청한 파일에 대한 접근 시간의 단축 분인 상대적 이득을 계산하는 것  
을 특징으로 하는 컴퓨팅 시스템.

**청구항 8**

제4항에 있어서,  
상기 혼성 파일 시스템 관리부는 요청된 파일의 크기 및 제1 저장 매체에 대한 쓰기 속도에 기초하여 상기 비용을 계산하는 것  
을 특징으로 하는 컴퓨팅 시스템.

**청구항 9**

제4항에 있어서,  
상기 혼성 파일 시스템 관리부는 계산된 상대적 이득이 계산된 비용보다 크면 요청된 파일을 제1 저장 매체로 이동하는 것  
을 특징으로 하는 컴퓨팅 시스템.

**청구항 10**

제1항에 있어서,  
혼성 파일 시스템 관리부는 계산된 상대적 이득이 계산된 비용보다 작거나 같으면 요청된 파일을 제1 저장 매체로 이동하지 않고 호스트로 전송하는 것  
을 특징으로 하는 컴퓨팅 시스템.

**청구항 11**

삭제

**청구항 12**

삭제

**청구항 13**

삭제

**청구항 14**

삭제

청구항 15

삭제

청구항 16

삭제

청구항 17

삭제

청구항 18

삭제

청구항 19

삭제

청구항 20

삭제

청구항 21

삭제

청구항 22

삭제

**발명의 설명**

**발명의 상세한 설명**

**기술 분야**

[0001] 본 발명은 스토리지 장치에 관한 것으로서, 특히 비휘발성 메모리(non-volatile memory)에 기반한 스토리지 장치 및 스토리지 장치를 위한 파일 시스템에 관한 것이다.

**배경 기술**

[0002] 데이터를 저장하는 스토리지 장치로는 자기 디스크(magnetic disk), 반도체 메모리 등이 있을 수 있다. 스토리지 장치는 종류 별로 서로 다른 물리적 특성을 가지기 때문에 물리적 특성에 상응하는 관리 방법이 필요하다.

[0003] 종래의 스토리지 장치로는 자기 디스크가 널리 사용되어 왔다. 자기 디스크는 평균적으로 킬로바이트(kilobyte) 당 수 밀리초(millisecond)의 읽기/쓰기 시간을 특성으로 가진다. 또한, 자기 디스크는 데이터가 저장된 물리적 위치에 따라 암(arm)이 도달하는 시간이 다르기 때문에 읽기/쓰기 시간이 달라지는 특성을 가진다.

[0004] 호스트(host) 또는 프로세서(processor)는 스토리지 장치에 저장된 데이터를 읽거나 스토리지 장치에 데이터를 저장하기 위하여 스토리지 장치에 접근(access)한다. 호스트 또는 프로세서가 스토리지 장치로부터 데이터를 읽어 들이는 시간은 호스트 또는 프로세서의 입장에서는 상대적으로 긴 시간이므로 호스트 또는 프로세서는 스토리지 장치에 저장된 데이터의 일부를 저장하는 메모리 (자기 디스크보다 데이터 입출력 시간이 짧은)를 이용할 수 있다.

[0005] 호스트 또는 프로세서는 한번 읽은 데이터를 가까운 시간 내에 다시 한번 읽어 들이는 경향이 있다. 이러한 경향을 시간적 지역성(temporal locality)라고 하며, 종래의 스토리지 장치 및 관리 시스템(management system)은 시간적 지역성을 이용하여 호스트 또는 프로세서가 스토리지 장치로부터 읽은 데이터를 저장하는 버퍼 메모리(buffer memory)를 이용한다.

**발명의 내용**

- [0006] 본 발명의 실시예들에 따른 스토리지 장치 또는 컴퓨팅 시스템은 비휘발성 메모리에 최적화된 파일 시스템을 이용함으로써 데이터 입력 및 출력 전송률을 높일 수 있다.
- [0007] 본 발명의 실시예들에 따른 스토리지 장치는 랜덤 쓰기 연산의 크기를 랜덤 쓰기 연산의 데이터 전송률을 기준치 이상으로 조정함으로써 랜덤 쓰기 연산의 데이터 전송률을 높일 수 있다.
- [0008] 본 발명의 실시예들에 따른 스토리지 장치는 랜덤 읽기 연산의 데이터 전송률이 상대적으로 높은 비휘발성 메모리에 최적화된 파일 시스템을 제공할 수 있다.
- [0009] 본 발명의 실시예들에 따른 스토리지 장치는 호스트와 무효 파일에 대한 정보를 공유함으로써 비휘발성 메모리의 가비지 콜렉션 연산으로 인한 부하를 줄일 수 있다.
- [0010] 본 발명의 실시예들에 따른 스토리지 장치는 호스트로부터 수신된 커맨드를 저장하고 스케줄링함으로써 보다 많은 데이터를 병렬 처리할 수 있다.
- [0011] 본 발명의 실시예들에 따른 스토리지 장치는 서로 다른 종류의 저장 매체들 각각에 최적화된 파일 시스템을 이용함으로써 데이터 전송률을 높일 수 있다. 본 발명의 실시예들에 따른 스토리지 장치는 접근 빈도에 기초하여 파일이 저장되는 저장 매체를 선택함으로써 데이터의 전반적인 전송 시간을 단축할 수 있다.
- [0012] 본 발명의 일 실시예에 따른 컴퓨팅 시스템은 제1 저장 매체, 제2 저장 매체 및 혼성 파일 시스템 관리부(Hybrid File System Manager)를 포함할 수 있다. 제2 저장 매체는 제1 저장 매체의 입출력 속도보다 느린 입출력 속도를 가질 수 있다. 혼성 파일 시스템 관리부는 제1 저장 매체를 제어하는 제1 물리 파일 시스템 및 상기 제2 저장 매체를 제어하는 제2 물리 파일 시스템을 관리할 수 있다. 혼성 파일 시스템 관리부는 상기 제1 물리 파일 시스템 및 상기 제2 물리 파일 시스템으로부터 변환된 가상 파일 시스템을 호스트에 제공할 수 있다.
- [0013] 본 발명의 다른 실시예에 따른 스토리지 장치는 제1 데이터 및 상기 제1 데이터의 메타 데이터를 저장하는 저장 매체 및 호스트로부터 상기 저장 매체로 전송되는 데이터를 순차적으로 저장하는 버퍼 메모리를 포함할 수 있다. 저장 매체는 상기 버퍼 메모리에 저장된 데이터의 크기가 기준치 이상이면 상기 제1 데이터 및 상기 제1 데이터의 메타 데이터의 위치의 다음 위치에 상기 버퍼 메모리에 저장된 데이터를 제2 데이터로서 저장할 수 있다. 저장 매체는 상기 제2 데이터의 메타 데이터를 상기 제2 데이터의 위치의 다음 위치에 저장할 수 있다.
- [0014] 본 발명의 또 다른 실시예에 따른 스토리지 장치는 비휘발성 메모리 및 컨트롤러를 포함할 수 있다. 비휘발성 메모리는 데이터 및 상기 데이터의 관리 정보를 포함하는 메타 데이터를 저장할 수 있다. 컨트롤러는 메타 데이터로부터 상기 비휘발성 메모리를 제어하고, 상기 메타 데이터의 업데이트가 필요하면 상기 비휘발성 메모리에 업데이트된 메타 데이터를 저장하고, 이전의 메타 데이터를 무효화할 수 있다. 컨트롤러는 호스트에 의해 무효화된 데이터의 정보를 상기 호스트로부터 수신하고, 상기 수신된 정보에 기초하여 상기 비휘발성 메모리의 데이터를 관리할 수 있다.
- [0015] 본 발명의 또 다른 실시예에 따른 스토리지 장치는 복수의 메모리 채널들, 커맨드 큐(command queue) 및 컨트롤러를 포함할 수 있다. 커맨드 큐는 상기 복수의 메모리 채널들에 대한 호스트로부터의 하나 이상의 커맨드를 순차적으로 수신하여 저장할 수 있다. 컨트롤러는 상기 커맨드 큐에 저장된 하나 이상의 커맨드 중 제1 커맨드를 선택하고, 상기 복수의 메모리 채널들 중 상기 선택된 제1 커맨드가 접근하는 제1 채널로 상기 선택된 제1 커맨드를 전송할 수 있다. 복수의 메모리 채널들 각각은 하나 이상의 비휘발성 메모리 유닛을 포함할 수 있다.

**발명의 실시를 위한 구체적인 내용**

- [0016] 이하에서, 본 발명에 따른 바람직한 실시예를 첨부된 도면을 참조하여 상세하게 설명한다. 그러나, 본 발명이 실시예들에 의해 제한되거나 한정되는 것은 아니다. 각 도면에 제시된 동일한 참조 부호는 동일한 부재를 나타낸다.
- [0017] 스토리지 장치로 널리 이용되는 자기 디스크의 종류로는 플로피 디스크 드라이브(Floppy Disk Drive) 또는 하드 디스크 드라이브(Hard Disk Drive, HDD) 등이 있다. HDD를 이용하면 매우 큰 용량의 스토리지 장치를 구현할 수 있기 때문에 널리 이용되고 있다. HDD는 회전하는 자기 디스크에 암(arm)이 접근하여 자기 디스크의 트랙으로부터 데이터를 읽거나 트랙에 데이터를 저장한다.

- [0018] 스토리지 장치를 포함하는 컴퓨팅 시스템의 초기 부팅 과정에서 HDD의 자기 디스크가 회전하는 시간 및 반도체 메모리에 비해 상대적으로 긴 데이터 읽기 시간으로 인하여 HDD를 이용하는 컴퓨팅 시스템의 부팅 시간이 길 수 있다.
- [0019] HDD는 자기 디스크의 회전 속도가 일정한 레벨(예를 들어 6000 rpm - rpm은 rotation per minute)에 도달해야 데이터의 안정적인 입출력이 가능하다. 따라서 부팅 시 자기 디스크의 회전 속도가 일정한 레벨에 도달하기까지 자기 디스크의 회전을 가속하는 시간이 필요하다.
- [0020] 또한, 부팅 시 HDD에 저장된 데이터를 메인 메모리로 로드(load)하는 시간이 소요된다. HDD는 반도체 메모리에 비해 상대적으로 긴 데이터 읽기 시간을 가지므로 부팅 시 HDD에 저장된 데이터를 로드하는 시간이 길 수 있다.
- [0021] 메인 메모리가 휘발성(volatile) 메모리인 경우 데이터의 일관성을 유지하기 위하여 빈번하게 데이터를 HDD에 저장할 필요가 있다. HDD는 자기 디스크의 회전을 이용하므로 데이터를 입력 및/또는 출력하는 데 상대적으로 긴 시간이 소요된다. HDD의 암 끝부분에 위치한 헤드(head)가 원하는 트랙 또는 섹터의 위치로 이동하는 과정을 통해서 HDD의 데이터에 대한 접근이 가능하므로, 랜덤한(random) 데이터의 입력 및/또는 출력 - random I/O라 불린다 - 시에 성능이 크게 저하될 수 있다.
- [0022] 메인 메모리의 크기는 스토리지 장치의 크기보다 작으므로 어플리케이션이 요구하는 메모리 공간이 메인 메모리의 크기보다 큰 경우가 빈번하게 발생할 수 있다. 메인 메모리 및 스토리지 장치 간에 필요한 데이터를 가져오고(fetch), 가져오는 데이터와 위치를 스왑하는 데이터를 방출(evict)하는 과정을 페이지 스와핑이라 부르기도 한다. 페이지 스와핑 속도는 메인 메모리 및 스토리지 장치 중 더 느린 장치의 동작 속도에 의해 결정되므로 일반적으로 스토리지 장치의 느린 속도가 페이지 스와핑 속도를 결정할 수 있다.
- [0023] 휘발성 메모리인 메인 메모리에 저장된 데이터가 업데이트되면 스토리지 장치에 저장된 데이터 - 스토리지 장치에 저장된 데이터는 메인 메모리에 저장된 데이터에 대응한다 - 도 업데이트되어야 한다. 메인 메모리에 저장된 데이터가 업데이트될 때마다 스토리지 장치에 저장된 데이터도 업데이트되는 기법을 라이트-스루 기법(write-through scheme)이라 한다. 한편, 메인 메모리에 저장된 데이터가 업데이트되면 업데이트되었음을 표시하고, 일정 시간 주기마다 또는 업데이트의 표시가 일정 수 이상이면 스토리지 장치에 저장된 데이터가 업데이트되는 기법을 라이트-백 기법(write-back scheme)이라 한다. 라이트-백 기법은 라이트-비하인드 기법(write-behind scheme)이라 불리기도 한다. 메인 메모리에 저장된 데이터가 업데이트된 경우 메인 메모리에 저장된 데이터가 dirty하다고 표현하기도 한다.
- [0024] 메모리에 저장된 데이터 중 업데이트되는 데이터는 일반적으로 높은 시간적 지역성(temporal locality)을 가지므로 가까운 시간 내에 빈번하게 다시 업데이트될 가능성이 높다. 이 경우 동일한 위치에 저장된 데이터가 반복적으로 스토리지 장치에 다시 저장될 가능성이 높으므로 스토리지 장치인 HDD의 느린 동작 속도가 전체 컴퓨팅 시스템의 동작 속도를 저하시킬 수 있다.
- [0025] HDD는 디스크의 회전 및 암의 이동 등 기계적으로 동작하는 과정에서 상대적으로 많은 양의 전력을 소모한다. HDD가 사용되는 빈도는 전체 어플리케이션의 일부이지만 상대적으로 많은 양의 전력을 소모하므로, HDD가 사용되지 않는 동안에 HDD의 전원을 차단하는 기법이 고려되기도 하였으나 HDD의 전원을 차단하였다가 다시 동작 가능한 상태로 전환하는 데 소모되는 전력 양이 크기 때문에 HDD의 디스크의 회전 속도를 조정하는 dynamic power management 기법이 사용되고 있다.
- [0026] 메인 메모리에 널리 이용되는 DRAM(Dynamic Random Access Memory)은 일정 시간 주기마다 메모리 셀에 저장된 데이터를 refresh하는 동작으로 인해 소모 전력이 상대적으로 높다. 본 발명의 실시예들에 따르면 휘발성 메모리의 소모 전력보다 작은 전력을 소모하는 비휘발성 메모리를 이용함으로써 컴퓨팅 시스템의 소모 전력을 줄일 수 있다.
- [0027] 비휘발성 메모리를 이용하여 대용량 스토리지 장치(storage device)를 구현한 제품으로 SSD(Solid State Disk)가 이용되고 있다. SSD는 복수의 채널들을 이용하여 병렬적으로 데이터를 전송할 수 있고, 채널들 각각은 독립적인 연산을 수행할 수 있다. 하나의 채널은 복수의 메모리 뱅크들을 포함할 수 있고, SSD는 복수의 메모리 뱅크들을 이용하여 데이터의 스트루풋(throughput)을 높일 수 있다.
- [0028] SSD는 HDD처럼 기계적인 동작을 수행하지 않으므로 HDD가 소모하는 전력보다 낮은 전력을 소모할 수 있고, HDD의 동작 시 발생하는 소음보다 작은 소음을 발생시킬 수 있다.
- [0029] HDD는 암(arm)을 움직여 데이터에 접근하므로 랜덤 읽기(random read) 또는 랜덤 쓰기(random write) 연산을

수행할 때 상대적으로 긴 시간이 필요함은 앞에서 설명한 바와 같다. SSD는 HDD처럼 기계적인 동작을 수행하지 않으므로 순차적 읽기(sequential read) 연산을 수행할 때와 랜덤 읽기 연산을 수행할 때 소요되는 시간의 차이가 HDD에 비해 매우 작다.

- [0030] 도 1은 본 발명의 일 실시예에 따른 컴퓨팅 시스템(100)을 도시하는 도면이다.
- [0031] 도 1을 참조하면, 컴퓨팅 시스템(100)은 제1 저장 매체(110), 제2 저장 매체(120) 및 혼성 파일 시스템 관리부(Hybrid File System Manager)(130)를 포함한다.
- [0032] 제2 저장 매체(120)는 제1 저장 매체(110)의 입출력 속도보다 느린 입출력 속도를 가진다.
- [0033] 혼성 파일 시스템 관리부(130)는 제1 저장 매체(110)를 제어하는 제1 물리 파일 시스템 및 제2 저장 매체(120)를 제어하는 제2 물리 파일 시스템을 관리한다. 혼성 파일 시스템 관리부(130)는 제1 물리 파일 시스템 및 제2 물리 파일 시스템으로부터 변환된 가상 파일 시스템을 가상 파일 시스템 관리자(Virtual File System Manager)에 제공한다.
- [0034] 제1 저장 매체(110)는 제2 저장 매체(120)와 다른 종류의 저장 매체일 수 있다. 예를 들어 제1 저장 매체(110)는 SSD일 수 있고, 제2 저장 매체(120)는 HDD일 수 있다.
- [0035] 호스트 또는 프로세서는 혼성 파일 시스템 관리부(130)를 포함할 수 있다. 가상 파일 시스템 관리자는 호스트 또는 프로세서에 포함될 수 있으며 가상 파일 시스템을 관리할 수 있다.
- [0036] 도 2는 도 1의 컴퓨팅 시스템(100)의 동작의 일 예를 도시하는 도면이다.
- [0037] 도 2를 참조하면, 혼성 파일 시스템 관리부(130)는 제1 저장 매체(110)를 제어하는 제1 물리 파일 시스템(210)을 관리하고, 제2 저장 매체(120)를 제어하는 제2 물리 파일 시스템(220)을 관리한다. 제1 저장 매체(110)가 SSD이면 제1 물리 파일 시스템(210)은 SSD에 최적화된 파일 시스템일 수 있고, 제2 저장 매체(120)가 HDD이면 제2 물리 파일 시스템(220)은 HDD에 최적화된 파일 시스템일 수 있다.
- [0038] 혼성 파일 시스템 관리부(130)는 제1 물리 파일 시스템(210) 상의 파일을 가상 파일 시스템(230)에서 이용될 수 있는 포맷으로 변환할 수 있고, 가상 파일 시스템(230) 상의 파일을 제1 물리 파일 시스템(210)에서 이용될 수 있는 포맷으로 변환할 수 있다. 혼성 파일 시스템 관리부(130)는 제2 물리 파일 시스템(220) 상의 파일을 가상 파일 시스템(230)에서 이용될 수 있는 포맷으로 변환할 수 있고, 가상 파일 시스템(230) 상의 파일을 제2 물리 파일 시스템(220)에서 이용될 수 있는 포맷으로 변환할 수 있다.
- [0039] 혼성 파일 시스템 관리부(130)는 가상 파일 시스템(230) 상의 파일을 상기 파일의 주소에 따라 제1 물리 파일 시스템(210) 또는 제2 물리 파일 시스템(220)에서 이용될 수 있는 포맷으로 변환할 수 있다.
- [0040] 혼성 파일 시스템 관리부(130)는 하위 계층(layer)에 존재하는 저장 매체가 제1 저장 매체(110)인지 제2 저장 매체(120)인지에 무관하게 운영 체제가 인식할 수 있는 가상 파일 시스템(230)을 운영 체제에 제공할 수 있다.
- [0041] 혼성 파일 시스템 관리부(130)는 운영 체제로부터 요청된 파일의 접근 빈도에 기초하여 요청된 파일을 제1 물리 파일 시스템(210) 및 제2 물리 파일 시스템(220) 간에 migration할 수 있다. 혼성 파일 시스템 관리부(130)는 요청된 파일의 접근 빈도에 기초하여 요청된 파일을 제1 저장 매체(110)로부터 제2 저장 매체(120)로 이동할 지 여부를 결정할 수 있고, 요청된 파일을 제2 저장 매체(120)로부터 제1 저장 매체(110)로 이동할 지 여부를 결정할 수 있다.
- [0042] 도 3은 도 1의 혼성 파일 시스템 관리부(130)의 동작의 일 예를 도시하는 동작 흐름도이다.
- [0043] 도 3을 참조하면, 혼성 파일 시스템 관리부(130)는 운영 체제로부터 요청된 파일이 제1 저장 매체(110)에 저장되어 있는지 판정한다(S310).
- [0044] 혼성 파일 시스템 관리부(130)는 요청된 파일이 제2 저장 매체(120)에 저장되어 있으면 요청된 파일의 접근 빈도에 기초하여 요청된 파일의 제1 저장 매체(110)로의 이동에 따른 상대적 이득을 계산한다(S320). 상대적 이득의 예로는 혼성 파일 시스템 관리부(130)는 제1 저장 매체(110)로부터 호스트로의 데이터 전송 시간 및 제2 저장 매체(110)로부터 호스트로의 데이터 전송 시간의 차이를 요청된 파일의 접근 빈도에 곱하여 상대적 이득을 계산할 수 있다. 접근 빈도는 일정한 주기 내의 요청된 파일의 접근 횟수일 수 있다. 계산된 상대적 이득은 상기 일정한 주기 내에서 발생할 수 있는 요청된 파일에 대한 접근 시간(access time)의 단축 분일 수 있다.
- [0045] 혼성 파일 시스템 관리부(130)는 요청된 파일이 제2 저장 매체(120)에 저장되어 있으면 요청된 파일의 제1 저장



매체(110)로의 이동에 따른 비용을 계산한다(S330). 혼성 파일 시스템 관리부(130)는 요청된 파일의 크기 및 제1 저장 매체(120)에 대한 쓰기 속도에 기초하여 상기 비용을 계산할 수 있다.

- [0046] 혼성 파일 시스템 관리부(130)는 계산된 상대적 이득이 계산된 비용보다 큰지 여부를 판정한다(S340).
- [0047] 혼성 파일 시스템 관리부(130)는 계산된 상대적 이득이 계산된 비용보다 크면 요청된 파일을 제2 저장 매체(120)로부터 제1 저장 매체(110)로 이동한다(S350).
- [0048] 혼성 파일 시스템 관리부(130)는 요청된 파일을 호스트로 전송한다(S360).
- [0049] 혼성 파일 시스템 관리부(130)는 요청된 파일이 제1 저장 매체(110)에 저장되어 있으면 요청된 파일을 호스트로 전송한다(S360).
- [0050] 혼성 파일 시스템 관리부(130)는 계산된 상대적 이득이 계산된 비용보다 크지 않으면 요청된 파일을 제1 저장 매체(110)로 이동하지 않고 호스트로 전송한다(S360).
- [0051] 도 4는 도 3의 단계(S350)의 일 예를 상세히 도시하는 동작 흐름도이다.
- [0052] 도 4를 참조하면, 혼성 파일 시스템 관리부(130)는 제1 저장 매체(110)에 사용 가능한 공간이 남아 있는지 여부를 판정한다(S410). 혼성 파일 시스템 관리부(130)는 사용 가능한 공간 대 전체 공간(제1 저장 매체(110))의 비율이 기준치 이상인지 여부를 판정할 수도 있다. 상기 비율이 기준치 이상이면 혼성 파일 시스템 관리부(130)는 제1 저장 매체(110)에 사용 가능한 공간이 남아 있는 것으로 간주할 수도 있다.
- [0053] 혼성 파일 시스템 관리부(130)는 제1 저장 매체(110)에 사용 가능한 공간이 남아 있지 않으면 제1 저장 매체(110)에 저장된 파일 중 비용 대비 효과가 가장 낮은 데이터를 선택한다(S420). 혼성 파일 시스템 관리부(130)는 파일 각각의 크기 및 파일 각각의 접근 빈도에 기초하여 파일 각각의 비용 대비 효과를 계산할 수 있다.
- [0054] 혼성 파일 시스템 관리부(130)는 선택된 파일을 제2 저장 매체(120)로 이동한다(S430).
- [0055] 혼성 파일 시스템 관리부(130)는 요청된 파일을 제1 저장 매체(110)로 이동한다(S440).
- [0056] 혼성 파일 시스템 관리부(130)는 제1 저장 매체(110)에 사용 가능한 공간이 남아 있으면 요청된 파일을 제1 저장 매체(110)로 이동한다(S440).
- [0057] 도 4의 각 단계는 호스트로부터 제1 저장 매체(110)에 파일을 쓰는 연산이 수행될 때에도 적용될 수 있다.
- [0058] 도 5는 본 발명의 다른 실시예에 따른 스토리지 장치(500)를 도시하는 도면이다.
- [0059] 도 5를 참조하면, 스토리지 장치(500)는 버퍼 메모리(510) 및 저장 매체(520)를 포함한다.
- [0060] 저장 매체(520)는 제1 데이터 및 제1 데이터의 메타 데이터를 저장한다. 제1 데이터의 메타 데이터는 상기 제1 데이터에 대한 관리 정보를 포함한다.
- [0061] 버퍼 메모리(510)는 호스트로부터 저장 매체(520)로 전송되는 데이터를 순차적으로 저장한다.
- [0062] 저장 매체(520)는 버퍼 메모리(510)에 저장된 데이터의 크기가 기준치 이상이면 제1 데이터 및 제1 데이터의 메타 데이터의 위치의 다음 위치에 버퍼 메모리(510)에 저장된 데이터를 제2 데이터로서 저장한다. 저장 매체(520)는 제2 데이터를 저장한 후 제2 데이터의 메타 데이터를 제2 데이터의 다음 위치에 저장한다.
- [0063] 저장 매체(520)는 일정한 크기(= 상기 기준치) 단위로 데이터 및 메타 데이터를 저장하는 파일 시스템을 제공할 수 있다.
- [0064] 파일 시스템은 호스트에 의해 관리되며, 스토리지 장치(500)는 호스트의 제어에 따라 데이터 및 메타 데이터를 저장할 수 있다.
- [0065] 저장 매체(520)에 의해 제공되는 파일 시스템 하에서는 랜덤 읽기(random read) 연산이 빈번하게 발생할 수 있다. 상기 파일 시스템은 랜덤 읽기 연산의 성능 저하가 크지 않은 SSD에 적합한 파일 시스템일 수 있다. 상기 랜덤 읽기 연산의 성능은 스루풋(throughput), 데이터 접근 속도(data access speed) 등의 척도를 이용하여 측정될 수 있다.
- [0066] 저장 매체(520)에 의해 제공되는 파일 시스템 하에서는 일정한 크기 단위로 데이터 및 메타 데이터가 저장된다. SSD에 대한 랜덤 쓰기(random write) 연산의 성능은 순차적 쓰기(sequential write) 연산의 성능보다 낮지만 랜덤 쓰기 연산의 데이터 크기를 크게 할수록 랜덤 쓰기 연산의 성능이 증가할 수 있다.



- [0067] 저장 매체(520)에 의해 제공되는 파일 시스템 하에서는 상기 기준치를 조정함으로써 데이터 및 메타 데이터가 저장되는 단위 크기를 조절할 수 있다. 저장 매체(520)는 상기 기준치를 조정함으로써 랜덤 쓰기 연산의 성능을 증가시킬 수 있다.
- [0068] 저장 매체(520)가 비휘발성 메모리인 경우 저장 매체(520) 내의 사용 가능한 공간을 확보하기 위해 가비지 콜렉션(garbage collection) 연산이 수행될 수 있다. 이 때, 가비지 콜렉션 연산은 비휘발성 메모리 블록의 k배 단위로 수행될 수 있다. 상기 파일 시스템 하에서 데이터 및 메타 데이터가 저장되는 단위 크기가 증가할수록 랜덤 쓰기 연산의 성능이 증가하지만 상기 단위 크기가 가비지 콜렉션 연산이 수행되는 단위보다 큰 경우 성능의 증가 폭이 감소하는 경향이 있을 수 있다.
- [0069] 저장 매체(520)는 상기 기준치를 저장 매체(520)의 가비지 콜렉션 단위와 같도록 설정할 수 있다. 또는 저장 매체(520)는 상기 기준치를 저장 매체(520)의 가비지 콜렉션 단위보다 크도록 설정할 수 있다.
- [0070] 호스트가 저장 매체(520)의 파일 시스템을 효과적으로 이용하려면 호스트는 저장 매체(520)의 가비지 콜렉션 단위를 알 필요가 있다.
- [0071] 저장 매체(520)는 가비지 콜렉션 단위를 호스트에 통지할 수 있다.
- [0072] 또는 호스트는 저장 매체(520)에 복수의 랜덤 쓰기 연산을 요청할 수 있다. 저장 매체(520)는 요청된 복수의 랜덤 쓰기 연산을 수행함으로써 호스트가 가비지 콜렉션 단위를 인식하도록 할 수 있다. 상기 복수의 랜덤 쓰기 연산 각각은 서로 다른 크기의 데이터를 쓰는 연산이고, 호스트는 복수의 랜덤 쓰기 연산 각각에 대한 응답 시간에 기초하여 저장 매체(520)의 가비지 콜렉션 단위를 인식할 수 있다.
- [0073] 실시예에 따라서는, 버퍼 메모리(510)는 비휘발성 메모리일 수도 있다. 버퍼 메모리(510)는 바이트 단위로 접근 가능(accessible)한 비휘발성 메모리일 수 있다. 버퍼 메모리(510)는 NVRAM(Non-volatile RAM)일 수 있다.
- [0074] 실시예에 따라서는 버퍼 메모리(510)는 호스트 내부에 위치할 수도 있고, 저장 매체(520)와 동일한 패키지 내부에 위치할 수도 있다.
- [0075] 실시예에 따라서는 버퍼 메모리(510)는 호스트와 관련된 칩셋(chipset) 또는 메인보드(mainboard) 상에 위치할 수도 있다.
- [0076] 도 6은 도 5의 스토리지 장치(500)의 동작의 일 예를 도시하는 도면이다.
- [0077] 도 6을 참조하면, 버퍼 메모리(510)는 호스트로부터 전송된 데이터를 순차적으로 저장한다. 버퍼 메모리(510)에 저장된 데이터(610)가 기준치 이상이면 스토리지 장치(520)는 데이터(610)를 제2 데이터(640)로서 저장할 수 있다.
- [0078] 저장 매체(520)는 제1 데이터(620) 및 제1 메타 데이터(630)를 저장한다. 제1 메타 데이터(630)는 제1 데이터(620)에 대한 관리 정보를 포함하는 메타 데이터이다.
- [0079] 저장 매체(520)는 제1 메타 데이터(630)의 다음 위치에 제2 데이터(640)를 저장한다. 저장 매체(520)는 제2 데이터(640)를 저장한 후 제2 메타 데이터(650)를 제2 데이터(640)의 다음 위치에 저장한다. 제2 메타 데이터(650)는 제2 데이터(640)에 대한 관리 정보를 포함하는 메타 데이터이다.
- [0080] 스토리지 장치(500)에 의해 제공되는 파일 시스템은 도 2의 제1 물리 파일 시스템(210)으로 이용될 수 있다. 제1 저장 매체(110)가 SSD이면 스토리지 장치(500)에 의해 제공되는 파일 시스템은 SSD에 최적화된 제1 물리 파일 시스템(210)으로 이용될 수 있다.
- [0081] 도 7은 본 발명의 또 다른 실시예에 따른 스토리지 장치(700)를 도시하는 도면이다.
- [0082] 도 7을 참조하면, 스토리지 장치(700)는 컨트롤러(710) 및 비휘발성 메모리(720)를 포함한다.
- [0083] 비휘발성 메모리(720)는 데이터 및 상기 데이터의 관리 정보를 포함하는 메타 데이터를 저장한다.
- [0084] 컨트롤러(710)는 메타 데이터를 이용하여 비휘발성 메모리(720)를 제어하고, 메타 데이터의 업데이트가 필요하면 비휘발성 메모리(720)에 업데이트된 메타 데이터를 저장하고, 이전의 메타 데이터를 무효화(invalidate)한다.
- [0085] 컨트롤러(710)는 호스트에 의해 무효화된 데이터의 정보를 호스트로부터 수신하고, 상기 수신된 정보에 기초하여 비휘발성 메모리(720)의 데이터를 관리한다.

- [0086] 컨트롤러(710)는 호스트에 의해 수행되는 프로세스가 종료되면 상기 프로세스에 동적으로 할당된 비휘발성 메모리(720) 상의 저장 공간에 대한 무효화 정보를 호스트로부터 수신할 수 있다. 컨트롤러(710)는 수신된 무효화 정보에 기초하여 상기 프로세스에 동적으로 할당된 비휘발성 메모리(720) 상의 저장 공간을 무효화할 수 있다.
- [0087] 스토리지 장치(700)는 비휘발성 메모리(720)에 저장된 데이터에 대한 유효/무효 정보를 호스트와 공유함으로써 비휘발성 메모리(720)를 용이하게 관리할 수 있다. 스토리지 장치(700)는 호스트로부터 수신된 무효 정보에 기초하여 비휘발성 메모리(720)에 대한 가비지 콜렉션 연산의 수행 시간을 단축할 수 있다. 스토리지 장치(700)는 호스트에 의하여 무효화된 데이터를 가비지 콜렉션 연산 수행 동안 카피하지 않음으로써 가비지 콜렉션 연산의 수행 부하(load)를 줄일 수 있다.
- [0088] 도 8은 도 7의 비휘발성 메모리(720)의 한 블록에 저장된 데이터의 일 예를 도시하는 도면이다.
- [0089] 도 8을 참조하면, 비휘발성 메모리(720)의 블록은 6개의 페이지들(810 내지 860)을 포함한다.
- [0090] 호스트로부터 데이터 쓰기 커맨드를 수신하면 컨트롤러(710)는 제1 페이지(810), 제2 페이지(820), 제3 페이지(830) 및 제4 페이지(840)에 데이터를 저장한다. 컨트롤러(710)는 제5 페이지(850)에 상기 제1 페이지(810), 제2 페이지(820), 제3 페이지(830) 및 제4 페이지(840)에 저장된 데이터의 메타 데이터를 저장할 수 있다.
- [0091] 호스트로부터 상기 제1 페이지(810), 제2 페이지(820), 제3 페이지(830) 및 제4 페이지(840)에 저장된 데이터의 삭제(delete) 커맨드를 수신하면 컨트롤러(710)는 제6 페이지(860)에 업데이트된 메타 데이터를 저장할 수 있다. 제6 페이지(860)에 저장된 메타 데이터는 제5 페이지(850)에 저장된 메타 데이터가 무효임을 나타낼 수 있다.
- [0092] 호스트는 상기 제1 페이지(810), 제2 페이지(820), 제3 페이지(830) 및 제4 페이지(840)에 저장된 데이터의 무효화 정보를 컨트롤러(710)에 전송하고, 컨트롤러(710)는 수신된 무효화 정보에 기초하여 상기 제1 페이지(810), 제2 페이지(820), 제3 페이지(830) 및 제4 페이지(840)에 저장된 데이터를 관리할 수 있다.
- [0093] 비휘발성 메모리(720)에 대한 가비지 콜렉션 연산 동안 컨트롤러(710)는 호스트로부터 수신된 무효화 정보에 기초하여 상기 제1 페이지(810), 제2 페이지(820), 제3 페이지(830) 및 제4 페이지(840)에 저장된 데이터를 카피하지 않고 상기 블록을 소거(erase)할 수 있다. 이 때 컨트롤러(710)는 제6 페이지(860)에 저장된 메타 데이터에 기초하여 제5 페이지(850)에 저장된 메타 데이터를 카피하지 않고 상기 블록을 소거할 수 있다.
- [0094] 도 9는 도 7의 스토리지 장치(700)의 동작의 일 예를 도시하는 도면이다.
- [0095] 도 9를 참조하면, 컨트롤러(710)는 비휘발성 메모리(720)의 저장 공간을 스왑 영역(910) 및 파일 시스템 영역(920)으로 분할한다. 파일 시스템 영역(920)은 데이터 영역이라 불릴 수도 있다. 스왑 영역(910)은 가상 메모리 기능을 위한 영역이고, 파일 시스템 영역(920)은 데이터를 저장하기 위한 영역이다.
- [0096] 컨트롤러(710)는 스왑 영역(910) 중 영역(911)을 호스트에 의해 수행되는 프로세스에 대하여 동적으로 할당할 수 있다. 컨트롤러(710)는 프로세스가 종료되면 프로세스에 할당된 영역(911)에 저장된 데이터를 무효화할 수 있다. 컨트롤러(710)는 호스트로부터 프로세스 종료 정보 또는 영역(911)의 무효화 정보를 수신할 수 있다. 컨트롤러(710)는 호스트로부터 수신된 프로세스 종료 정보 또는 영역(911)의 무효화 정보에 기초하여 영역(911)에 저장된 데이터를 무효화할 수 있다.
- [0097] 컨트롤러(710)는 파일 시스템 영역(920) 중 영역(921)을 프로세스에 의해 임시로 사용되는 파일에 할당할 수 있다. 컨트롤러(710)는 프로세스가 종료되면 프로세스에 의해 임시로 사용되는 파일에 할당된 영역(921)에 저장된 데이터를 무효화할 수 있다.
- [0098] 도 10은 본 발명의 또 다른 실시예에 따른 스토리지 장치(1000)를 도시하는 도면이다.
- [0099] 도 10을 참조하면, 스토리지 장치(1000)는 커맨드 큐(1010), 컨트롤러(1020) 및 N개의 메모리 채널들(1001 내지 1006)을 포함한다.
- [0100] 커맨드 큐(Command Queue)(1010)는 N개의 메모리 채널들(1001 내지 1006)에 대한 하나 이상의 커맨드를 순차적으로 수신하여 저장한다. 상기 하나 이상의 커맨드는 호스트로부터 전송된다.
- [0101] 컨트롤러(1020)는 커맨드 큐(1010)에 저장된 하나 이상의 커맨드 중 제1 커맨드를 선택한다. 컨트롤러(1020)는 커맨드 큐(1010)를 검색하여 가장 먼저 수신된 커맨드를 제1 커맨드로 선택할 수도 있고, 미리 결정된 우선 순위 기준에 따라 가장 높은 우선 순위를 가지는 커맨드를 제1 커맨드로 선택할 수도 있다.

- [0102] 컨트롤러(1020)는 제1 채널로 제1 커맨드를 전송한다. 제1 채널은 N개의 메모리 채널들(1001 내지 1006) 중 제1 커맨드가 접근하고자 하는 채널이다.
- [0103] N개의 메모리 채널들(1001 내지 1006) 각각은 하나 이상의 비휘발성 메모리 유닛을 포함할 수 있다.
- [0104] 실시예에 따라서는 스토리지 장치(1000)는 N개의 메모리 채널들(1001내지 1006)을 포함하는 SSD일 수 있다. 스토리지 장치(1000)는 하나 이상의 커맨드를 순차적으로 수신하여 저장하는 커맨드 큐(1010)를 포함함으로써 호스트로부터의 커맨드를 병렬적으로 처리할 수 있다.
- [0105] 도 11은 도 10의 스토리지 장치(1000)의 동작의 일 예를 도시하는 타이밍도이다.
- [0106] 도 11을 참조하면, 커맨드 라인(1110)을 경유하여 호스트로부터 3개의 읽기 커맨드들이 스토리지 장치(1000)로 전송된다.
- [0107] 예를 들어 커맨드 1은 메모리 채널 1(1001)에 접근하는 읽기 커맨드이고, 커맨드 2는 메모리 채널 2(1002)에 접근하는 읽기 커맨드이고, 커맨드 3은 커맨드 3은 메모리 채널 3(1003)에 접근하는 읽기 커맨드일 수 있다.
- [0108] 호스트는 커맨드 1, 커맨드 2, 커맨드 3을 연속적으로 커맨드 라인(1110)을 경유하여 스토리지 장치(1000)로 전송할 수 있다. 커맨드 큐(1010)는 연속적으로 전송된 커맨드 1, 커맨드 2, 커맨드 3을 순차적으로 저장하고, 컨트롤러(1020)는 수신된 순서대로 커맨드 1을 메모리 채널 1(1001)로 전송하고, 커맨드 2를 메모리 채널 2(1002)로 전송하고, 커맨드 3을 메모리 채널 3(1003)으로 전송할 수 있다.
- [0109] N개의 메모리 채널들(1001 내지 1006) 각각은 독립적으로 커맨드를 수행할 수 있다. 가장 먼저 커맨드 1을 수신한 메모리 채널 1(1001)은 커맨드 1에 대응하는 데이터 1을 데이터 라인(1120)을 경유하여 호스트로 전송할 수 있고, 메모리 채널 2(1002)는 커맨드 2에 대응하는 데이터 2를 데이터 1의 전송이 종료된 후에 전송할 수 있다. 메모리 채널 3(1003)은 커맨드 3에 대응하는 데이터 3을 데이터 2의 전송이 종료된 후에 전송할 수 있다.
- [0110] 스토리지 장치(1000)는 호스트가 커맨드 1에 대응하는 데이터 1을 기다리지 않고 커맨드 2를 전송할 수 있도록 하고 병렬적으로 동작하는 N개의 메모리 채널들(1001 내지 1006)을 포함함으로써 전반적인 데이터의 전송 밴드폭 (bandwidth)을 높일 수 있다.
- [0111] 도 12는 도 10의 스토리지 장치(1000)의 동작의 다른 예를 도시하는 동작 흐름도이다.
- [0112] 도 12를 참조하면, 스토리지 장치(1000)는 상기 선택된 제1 커맨드가 쓰기 커맨드인지 여부를 판정한다(S1210).
- [0113] 스토리지 장치(1000)는 상기 선택된 제1 커맨드가 쓰기 커맨드이면 상기 제1 채널에서 가비지 콜렉션이 수행 중인지 여부를 확인한다(S1220).
- [0114] 스토리지 장치(1000)는 상기 제1 채널에서 가비지 콜렉션이 수행 중이면 커맨드 큐(1010)를 검색한다(S1230).
- [0115] 스토리지 장치(1000)는 커맨드 큐(1010)에 저장된 하나 이상의 커맨드 중에서 제1 채널 이외의 제2 채널에 접근하는 제2 커맨드를 선택한다(S1240). 스토리지 장치(1000)는 상기 선택된 제2 커맨드를 제2 채널로 전송한다.
- [0116] 스토리지 장치(1000)는 제2 채널에서 상기 선택된 제2 커맨드를 수행한다(S1250).
- [0117] 스토리지 장치(1000)는 제1 커맨드가 쓰기 커맨드가 아니고 읽기 커맨드이면 제1 채널에서 제1 커맨드를 수행한다(S1260).
- [0118] 스토리지 장치(1000)는 제1 커맨드가 쓰기 커맨드이고 제1 채널에서 가비지 콜렉션이 수행 중이 아니면 제1 채널에서 제1 커맨드를 수행한다(S1260).
- [0119] 단계(S1240)에서 스토리지 장치(1000)는 제1 채널 및 제2 채널 이외의 제3 채널에 접근하는 제3 커맨드를 선택할 수 있다. 스토리지 장치(1000)는 병렬적으로 제2 커맨드를 제2 채널로 전송하고, 제3 커맨드를 제3 채널로 전송할 수 있다.
- [0120] 스토리지 장치(1000)는 병렬적으로 제2 채널에서 제2 커맨드를 수행하고, 제3 채널에서 제3 커맨드를 수행할 수 있다.
- [0121] 단계(S1260)에서 스토리지 장치(1000)는 제1 채널에서 제1 커맨드가 수행되는 동안 제1 채널 이외의 제4 채널에 접근하는 제4 커맨드를 커맨드 큐(1010)에 저장된 커맨드 중에서 선택할 수 있다.
- [0122] 스토리지 장치(1000)는 제1 채널에서 제1 커맨드가 수행되는 동안 제4 채널로 제4 커맨드를 전송할 수 있고 제4

커맨드를 제4 채널에서 수행할 수 있다.

- [0123] 메모리 채널들 각각은 복수의 메모리 뱅크들을 포함할 수 있다. 스토리지 장치(1000)는 메모리 뱅크들 각각에 대하여 병렬적으로 커맨드를 전송하고, 메모리 뱅크들 각각에 대하여 병렬적으로 커맨드를 수행할 수 있다.
- [0124] 도 13은 도 10의 메모리 채널 1(1001)의 일 예를 도시하는 도면이다.
- [0125] 도 13을 참조하면, 메모리 채널 1(1001)은 공통의 데이터 버스에 연결되는 8개의 비휘발성 메모리들(1310 내지 1380)을 포함한다.
- [0126] 비휘발성 메모리 1(1310), 비휘발성 메모리 2(1320), 비휘발성 메모리 3(1330) 및 비휘발성 메모리 4(1340)는 메모리 그룹 1을 형성할 수 있고, 비휘발성 메모리 5(1350), 비휘발성 메모리 6(1360), 비휘발성 메모리 7(1370) 및 비휘발성 메모리 8(1380)는 메모리 그룹 2를 형성할 수 있다.
- [0127] 도 14는 도 13의 비휘발성 메모리 1(1310)의 일 예를 도시하는 도면이다.
- [0128] 도 14를 참조하면, 비휘발성 메모리 1(1310)는 버퍼 메모리(1410) 및 비휘발성 메모리 유닛(1420)을 포함한다.
- [0129] 버퍼 메모리(1410)는 비휘발성 메모리 유닛(1420)의 데이터 입력/출력 속도보다 빠른 데이터 입력/출력 속도를 가질 수 있다.
- [0130] 버퍼 메모리(1410)는 비휘발성 메모리 유닛(1420)에 저장된 데이터 중 호스트에 의해 접근된 데이터를 저장할 수 있다.
- [0131] 도 15는 도 10의 스토리지 장치(1000)의 동작의 또 다른 예를 도시하는 동작 흐름도이다.
- [0132] 도 15를 참조하면, 스토리지 장치(1000)는 커맨드 큐(1010)를 검색한다(S1510).
- [0133] 스토리지 장치(1000)는 커맨드 큐(1010)에 저장된 커맨드 중 하나를 선택하고, 상기 선택된 커맨드가 읽기 커맨드인지 여부를 판정한다(S1520).
- [0134] 스토리지 장치(1000)는 상기 선택된 커맨드가 읽기 커맨드이면 상기 선택된 커맨드가 버퍼 메모리(1410)에 저장된 데이터를 데이터를 읽는 읽기 커맨드인지 여부를 판정한다(S1530).
- [0135] 스토리지 장치(1000)는 상기 선택된 커맨드가 버퍼 메모리(1410)에 저장된 데이터를 읽는 읽기 커맨드이면 상기 선택된 커맨드에 우선 순위 1을 부여한다(S1550).
- [0136] 스토리지 장치(1000)는 상기 선택된 커맨드가 비휘발성 메모리 유닛(1420)에 저장된 데이터를 읽는 읽기 커맨드이면 상기 선택된 커맨드에 우선 순위 2를 부여한다(S1560).
- [0137] 스토리지 장치(1000)는 커맨드 큐(1010)에 저장된 하나 이상의 커맨드 중 버퍼 메모리(1410)에 저장된 데이터를 읽을 수 있는 읽기 커맨드에 비휘발성 메모리 유닛(1420)에 저장된 데이터를 읽는 읽기 커맨드보다 높은 우선 순위를 부여할 수 있다(우선 순위 1은 우선 순위 2보다 높은 우선 순위를 가지는 것으로 본다).
- [0138] 스토리지 장치(1000)는 상기 선택된 커맨드가 쓰기 커맨드이면 상기 선택된 커맨드에 의해 쓰여지는 데이터의 크기가 기준치 이하인지 여부를 판정한다(S1540).
- [0139] 스토리지 장치(1000)는 상기 선택된 커맨드에 의해 쓰여지는 데이터의 크기가 기준치 이하이면 상기 선택된 커맨드에 우선 순위 3을 부여한다(S1570).
- [0140] 스토리지 장치(1000)는 상기 선택된 커맨드에 의해 쓰여지는 데이터의 크기가 기준치보다 크면 상기 선택된 커맨드에 우선 순위 4를 부여한다(S1580).
- [0141] 스토리지 장치(1000)는 커맨드 큐(1010)에 저장된 하나 이상의 커맨드 중 읽기 커맨드에 쓰기 커맨드보다 높은 우선 순위를 부여할 수 있다.
- [0142] 스토리지 장치(1000)는 커맨드 큐(1010)에 저장된 하나 이상의 커맨드 중 기준치 이하의 크기의 데이터를 쓰는 커맨드에 기준치보다 큰 크기의 데이터를 쓰는 커맨드보다 높은 우선 순위를 부여할 수 있다.
- [0143] 스토리지 장치(1000)는 기준치 이하의 크기의 데이터를 쓰는 커맨드의 데이터를 버퍼 메모리(1410)에 저장할 수 있다. 스토리지 장치(1000)는 버퍼 메모리(1410)에 저장된 데이터를 일정 주기마다 비휘발성 메모리 유닛(1420)에 저장할 수 있다. 실시예에 따라서는 스토리지 장치(1000)는 기준치 이하의 크기의 데이터를 쓰는 커맨드의 데이터를 버퍼 메모리(1410) 및 비휘발성 메모리 유닛(1420)에 동시에 저장할 수도 있다.

- [0144] 스토리지 장치(1000)는 기준치보다 큰 크기의 데이터를 쓰는 커맨드의 데이터를 버퍼 메모리(1410)를 경유하지 않고 비휘발성 메모리 유닛(1420)에 저장할 수 있다.
- [0145] 기준치는 버퍼 메모리(1410)의 크기 또는 버퍼 메모리(1410)의 한 행(row)의 크기에 기초하여 결정될 수 있다. 스토리지 장치(1000)는 일정 크기 이하의 데이터를 버퍼 메모리(1410)에 저장함으로써 상기 데이터를 호스트가 요청하는 경우에 빠르게 호스트로 전송할 수 있다.
- [0146] 스토리지 장치(1000)는 커맨드 큐(1010)에 저장된 커맨드 간의 우선 순위를 결정하는 스케줄링(scheduling)을 수행할 수 있다. 스케줄링의 기준은 다음과 같을 수 있다.
- [0147] 1. 읽기 커맨드의 데이터를 버퍼 메모리(1410)로부터 빠르게 읽을 수 있는 경우
- [0148] 2. 읽기 커맨드의 데이터를 비휘발성 메모리 유닛(1420)로부터 읽어야 하는 경우
- [0149] 3. 쓰기 커맨드의 데이터를 버퍼 메모리(1410)에 빠르게 저장할 수 있는 경우(쓰기 커맨드의 데이터의 크기가 기준치 이하인 경우)
- [0150] 4. 쓰기 커맨드의 데이터를 비휘발성 메모리(1420)에 저장해야 하는 경우
- [0151] 스토리지 장치(1000)는 가장 높은 우선 순위를 가지는 커맨드를 제1 커맨드로 선택하고, 제1 커맨드에 대응하는 제1 채널에서 제1 커맨드를 수행할 수 있다. 스토리지 장치(1000)는 제1 채널에서 제1 커맨드가 수행되는 동안 제1 채널에서 수행되는 커맨드가 아닌 커맨드 중에서 가장 높은 우선 순위를 가지는 커맨드를 제2 커맨드로 선택하고, 제2 커맨드에 대응하는 제2 채널에서 제2 커맨드를 수행할 수 있다. 이 때, 스토리지 장치(1000)는 제1 커맨드 및 제2 커맨드를 병렬적으로 수행할 수 있다.
- [0152] SSD는 상대적으로 많은 수의 채널에서 커맨드를 병렬적으로 수행할 수 있다. SSD는 높은 데이터 입력/출력 병렬성(parallelism)을 가진다고 할 수 있다.
- [0153] 스토리지 장치(1000)는 SSD일 수 있고, 스토리지 장치(1000)는 커맨드 큐(1010)를 이용함으로써 SSD 내부의 높은 데이터 입력/출력 병렬성을 호스트가 이용하도록 할 수 있다.
- [0154] 호스트는 커맨드 큐(1010)의 존재로 인해 복수의 커맨드들을 병렬적으로 스토리지 장치(1000)로 전송할 수 있으므로 SSD 내부의 높은 데이터 입력/출력 병렬성을 이용할 수 있다. 스토리지 장치(1000)는 호스트 및 스토리지 장치(1000) 간의 데이터 전송률을 높일 수 있다.
- [0155] 본 발명의 실시예들에 따른 스토리지 장치들은 메모리 관리 기능을 통합하여 관리함으로써 시스템 인 패키지(System-in Package, SiP)를 용이하게 구현하도록 할 수 있다.
- [0156] 본 발명의 실시예들에 따른 플래시 메모리 장치 그리고/또는 메모리 컨트롤러는 다양한 형태들의 패키지를 이용하여 구현될 수 있다. 예를 들면, 본 발명의 실시예들에 따른 플래시 메모리 장치 그리고/또는 메모리 컨트롤러는 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 패키지들을 이용하여 구현될 수 있다.
- [0157] 플래시 메모리 장치와 메모리 컨트롤러는 메모리 카드를 구성할 수 있다. 이러한 경우, 메모리 컨트롤러는 USB, MMC, PCI-E, SATA, PATA, SCSI, ESDI, 그리고 IDE 등과 같은 다양한 인터페이스 프로토콜들 중 하나를 통해 외부(예를 들면, 호스트)와 통신하도록 구성될 수 있다.
- [0158] 플래시 메모리 장치는 전력이 차단되어도 저장된 데이터를 유지할 수 있는 비휘발성 메모리 장치이다. 셀룰러 폰, PDA 디지털 카메라, 포터블 게임 콘솔, 그리고 MP3P와 같은 모바일 장치들의 사용 증가에 따라, 플래시 메모리 장치는 데이터 스토리지 뿐만 아니라 코드 스토리지로서 보다 널리 사용될 수 있다. 플래시 메모리 장치는, 또한, HDTV, DVD, 라우터, 그리고 GPS와 같은 홈 어플리케이션에 사용될 수 있다.
- [0159] 본 발명의 실시예들에 따른 컴퓨팅 시스템 또는 호스트는 버스에 전기적으로 연결된 마이크로프로세서, 사용자 인터페이스, 베이스밴드 칩셋(baseband chipset)과 같은 모뎀, 메모리 컨트롤러, 그리고 플래시 메모리 장치를 포함한다. 플래시 메모리 장치에는 마이크로프로세서에 의해서 처리된/처리될 N-비트 데이터(N은 1 또는 그 보



다 큰 정수)가 메모리 컨트롤러를 통해 저장될 것이다. 본 발명의 실시예에 따른 컴퓨팅 시스템이 모바일 장치인 경우, 컴퓨팅 시스템의 동작 전압을 공급하기 위한 배터리가 추가적으로 제공될 것이다.

[0160] 본 발명의 실시예들에 따른 컴퓨팅 시스템에는 응용 칩셋(application chipset), 카메라 이미지 프로세서(Camera Image Processor: CIS), 모바일 디램, 등이 더 제공될 수 있음은 이 분야의 통상적인 지식을 습득한 자들에게 자명하다. 메모리 컨트롤러와 플래시 메모리 장치는, 예를 들면, 데이터를 저장하는 데 비휘발성 메모리를 사용하는 SSD(Solid State Drive/Disk)를 구성할 수 있다.

[0161] 이상과 같이 본 발명은 비록 한정된 실시예와 도면에 의해 설명되었으나, 본 발명은 상기의 실시예에 한정되는 것은 아니며, 본 발명이 속하는 분야에서 통상의 지식을 가진 자라면 이러한 기재로부터 다양한 수정 및 변형이 가능하다.

[0162] 그러므로, 본 발명의 범위는 설명된 실시예에 국한되어 정해져서는 아니 되며, 후술하는 특허청구범위뿐 아니라 이 특허청구범위와 균등한 것들에 의해 정해져야 한다.

**도면의 간단한 설명**

[0163] 도 1은 본 발명의 일 실시예에 따른 스토리지 장치(100)를 도시하는 도면이다.

[0164] 도 2는 도 1의 스토리지 장치(100)의 동작의 일 예를 도시하는 도면이다.

[0165] 도 3은 도 1의 혼성 파일 시스템 관리부(130)의 동작의 일 예를 도시하는 동작 흐름도이다.

[0166] 도 4는 도 3의 단계(S350)의 일 예를 상세히 도시하는 동작 흐름도이다.

[0167] 도 5는 본 발명의 다른 실시예에 따른 스토리지 장치(500)를 도시하는 도면이다.

[0168] 도 6은 도 5의 스토리지 장치(500)의 동작의 일 예를 도시하는 도면이다.

[0169] 도 7은 본 발명의 또 다른 실시예에 따른 스토리지 장치(700)를 도시하는 도면이다.

[0170] 도 8은 도 7의 비휘발성 메모리(720)의 한 블록에 저장된 데이터의 일 예를 도시하는 도면이다.

[0171] 도 9는 도 7의 스토리지 장치(700)의 동작의 일 예를 도시하는 도면이다.

[0172] 도 10은 본 발명의 또 다른 실시예에 따른 스토리지 장치(1000)를 도시하는 도면이다.

[0173] 도 11은 도 10의 스토리지 장치(1000)의 동작의 일 예를 도시하는 타이밍도이다.

[0174] 도 12는 도 10의 스토리지 장치(1000)의 동작의 다른 예를 도시하는 동작 흐름도이다.

[0175] 도 13은 도 10의 메모리 채널 1(1001)의 일 예를 도시하는 도면이다.

[0176] 도 14는 도 13의 비휘발성 메모리 1(1310)의 일 예를 도시하는 도면이다.

[0177] 도 15는 도 10의 스토리지 장치(1000)의 동작의 또 다른 예를 도시하는 동작 흐름도이다.

[0178] <도면의 주요 부분에 대한 부호의 설명>

[0179] 110: 제1 저장 매체

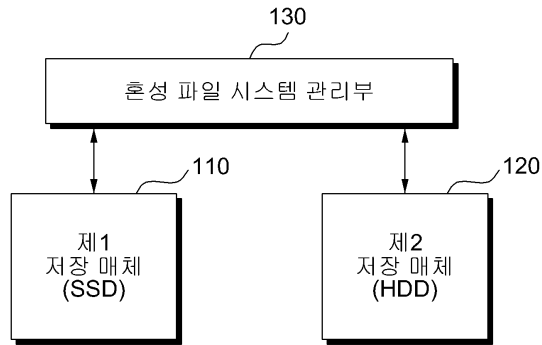
[0180] 120: 제2 저장 매체

[0181] 130: 혼성 파일 시스템 관리부

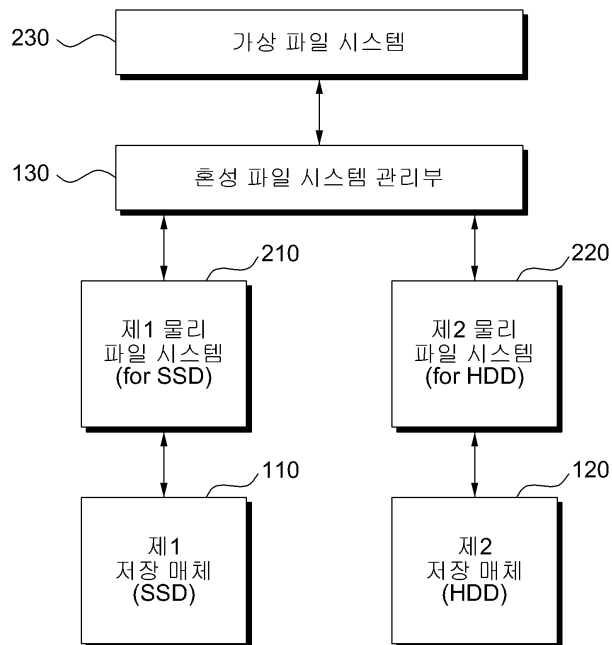
도면

도면1

100

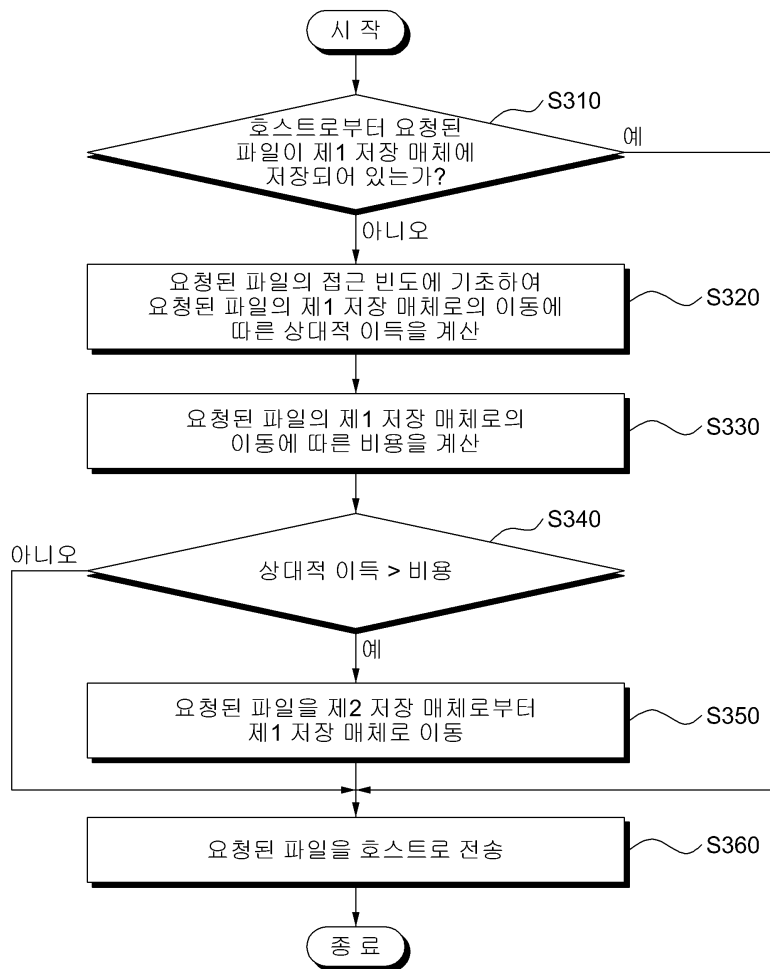


도면2

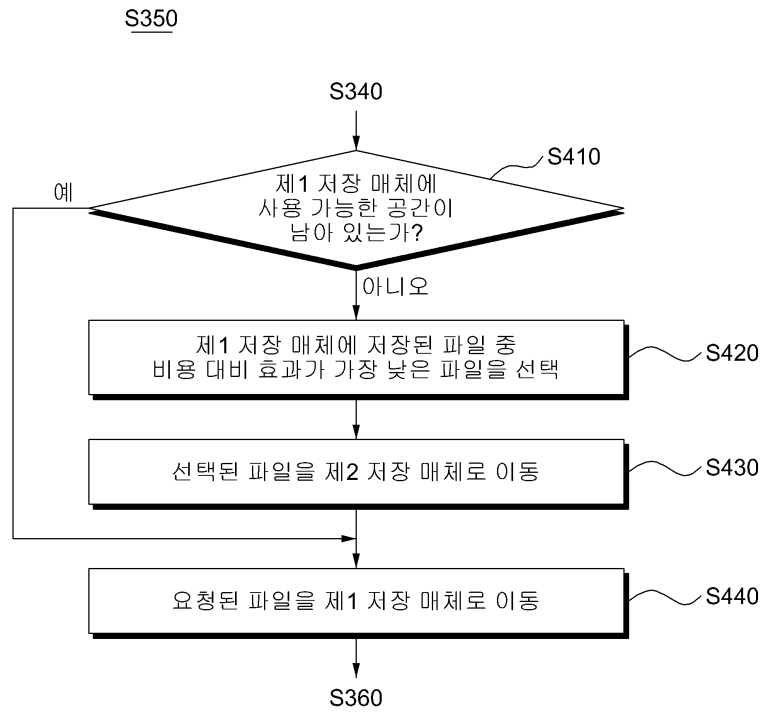




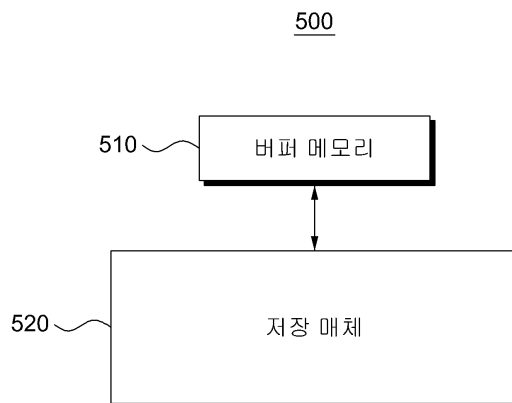
도면3



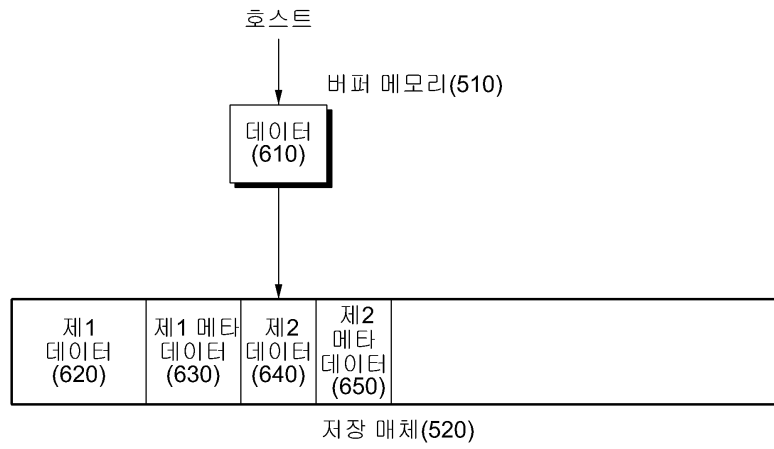
도면4



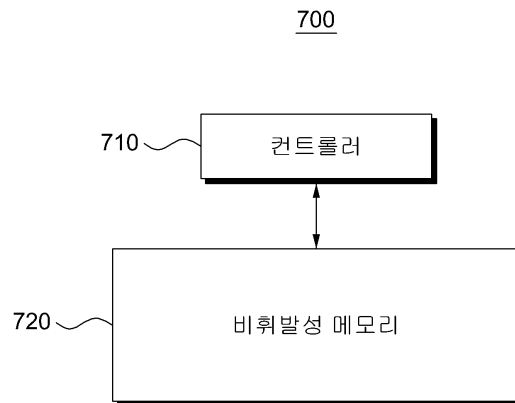
도면5



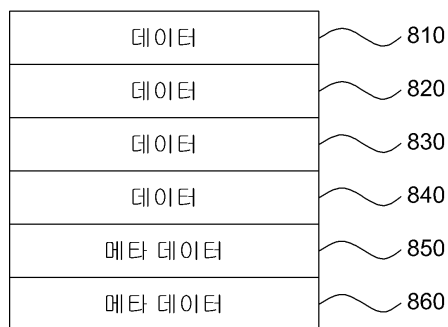
도면6



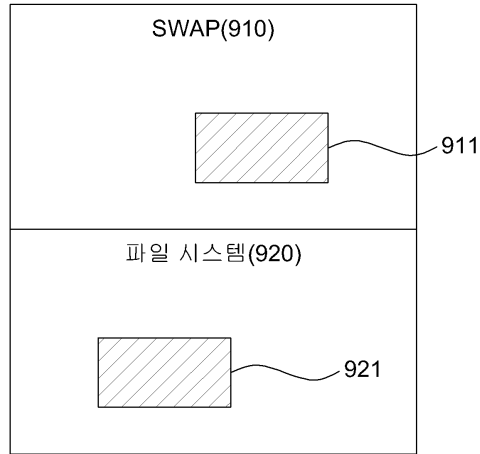
도면7



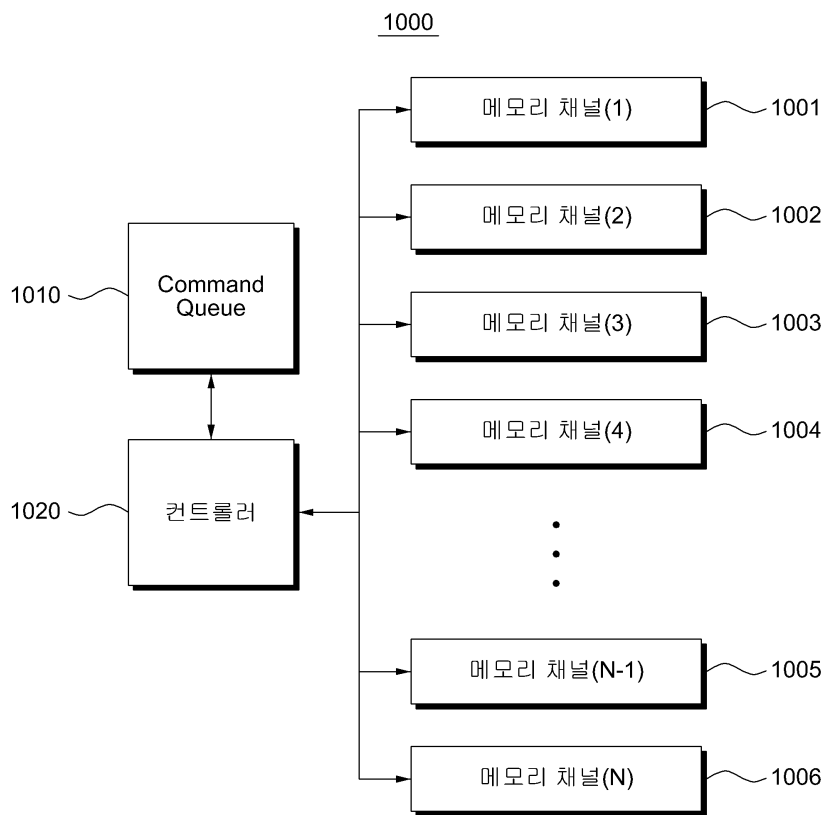
도면8



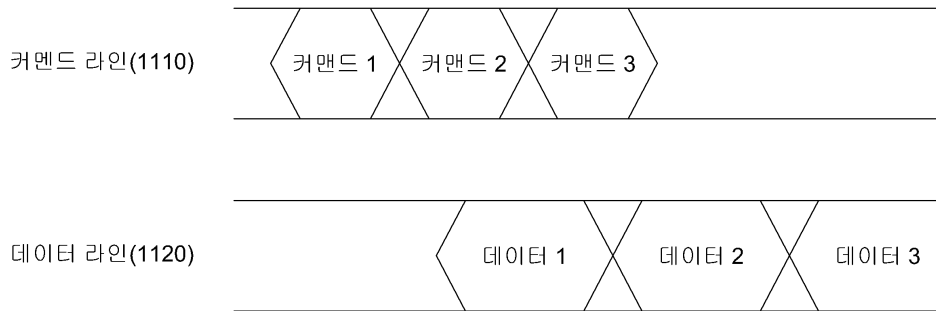
도면9



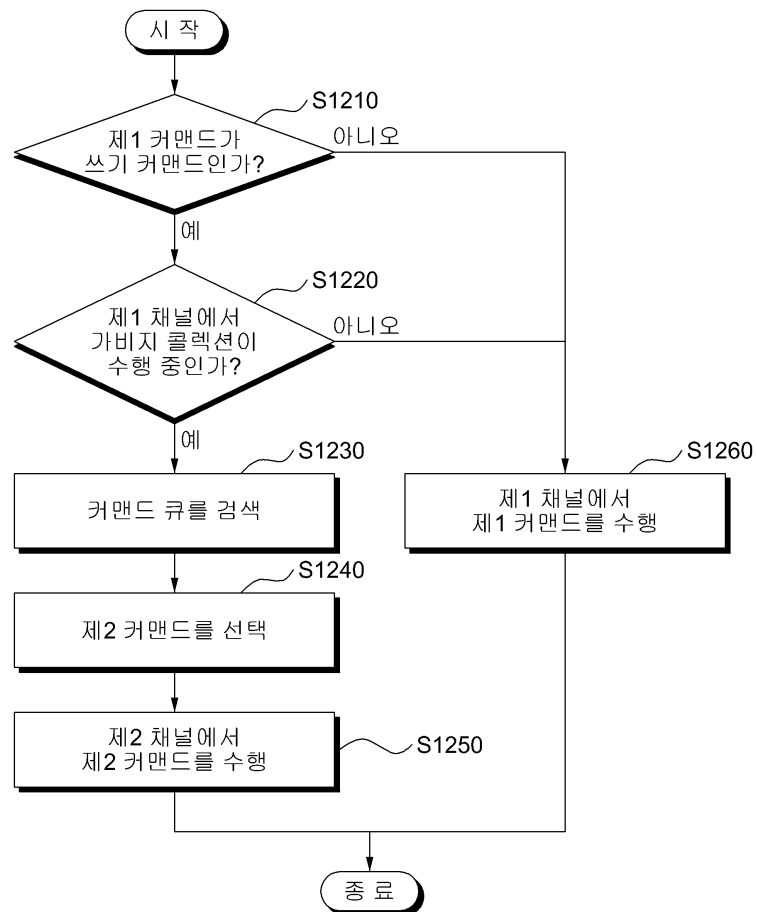
도면10



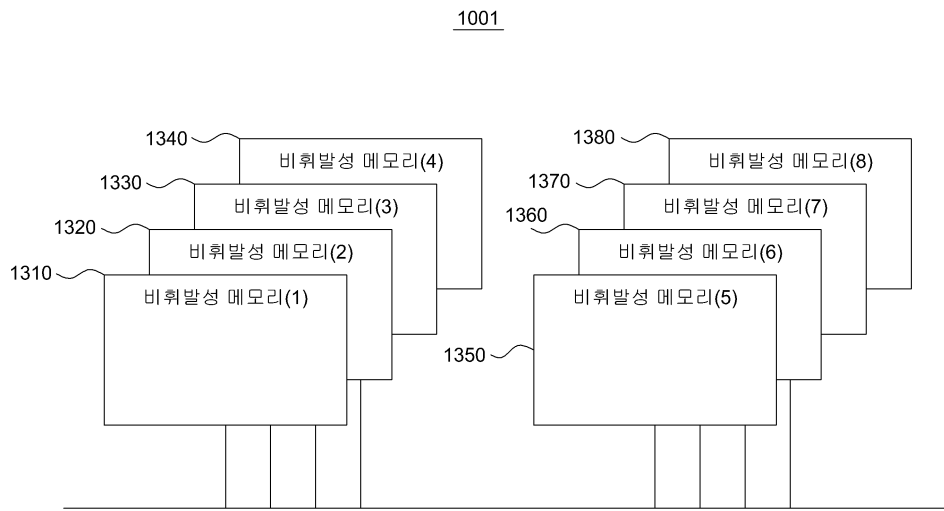
도면11



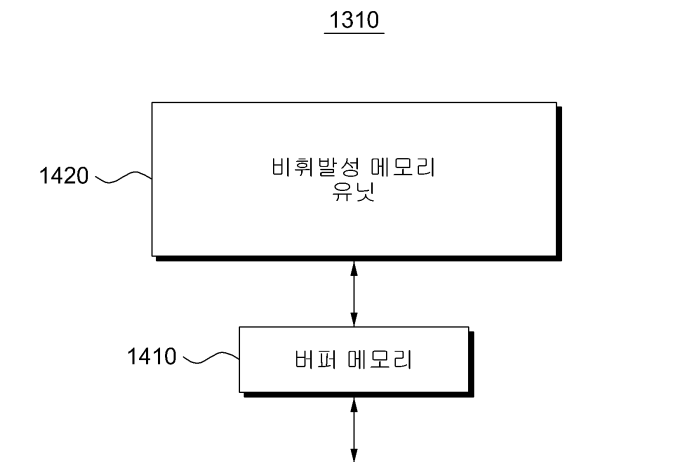
도면12



도면13



도면14



도면15

