



(19) 대한민국특허청(KR)

(12) 등록특허공보(B1)

(45) 공고일자 2016년05월25일

(11) 등록번호 10-1624376

(24) 등록일자 2016년05월19일

(51) 국제특허분류(Int. Cl.)

G06F 9/06 (2006.01) G06F 15/16 (2006.01)

G06F 9/44 (2006.01)

(21) 출원번호 10-2011-7026772

(22) 출원일자(국제) 2010년05월10일

심사청구일자 2015년04월08일

(85) 번역문제출일자 2011년11월10일

(65) 공개번호 10-2012-0018314

(43) 공개일자 2012년03월02일

(86) 국제출원번호 PCT/US2010/034276

(87) 국제공개번호 WO 2010/132375

국제공개일자 2010년11월18일

(30) 우선권주장

12/464,329 2009년05월12일 미국(US)

(56) 선행기술조사문현

JP2005222123 A

JP2008040571 A

US07937545 B1

(73) 특허권자

마이크로소프트 테크놀로지 라이센싱, 엘엘씨

미국 워싱턴주 (우편번호 : 98052) 레드몬드 원
마이크로소프트 웨이

(72) 발명자

리오네티 크리스

미국 워싱턴주 98052-6399 레드몬드 원 마이크로
소프트 웨이 엘씨에이 - 인터내셔널 페이턴츠 마
이크로소프트 코포레이션

파이크 로버트

미국 워싱턴주 98052-6399 레드몬드 원 마이크로
소프트 웨이 엘씨에이 - 인터내셔널 페이턴츠 마
이크로소프트 코포레이션

(74) 대리인

제일특허법인

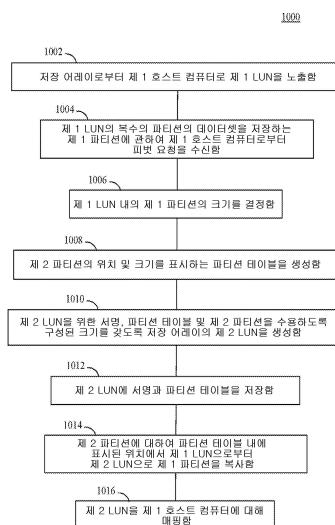
전체 청구항 수 : 총 18 항

심사관 : 권현수

(54) 발명의 명칭 LUN에서 파일로 또는 파일에서 LUN으로의 실시간 변환

(57) 요약

다수의 데이터셋(예를 들면, 가상 하드 드라이브 내에 저장된 가상 머신 등과 같은 데이터 및/또는 애플리케이션)을 저장할 수 있는 LUN이 제공된다. LUN은 다수의 파티션으로 분할된다. 하나 이상의 데이터셋은 각각의 파티션 내에 저장될 수 있다. 결과적으로, 다수의 데이터셋은 데이터셋의 개수에 비례하는 다수의 LUN을 통과하는 것 대신에 단일 LUN을 통해 액세스될 수 있다. 더욱이 LUN 내에 저장된 데이터셋은 피벗화(pivoted)될 수 있다. 제 1 LUN 내에 저장된 다수의 데이터셋 중 하나의 데이터셋을 저장하는 것을 전달하는 제 2 LUN이 생성될 수 있다. 데이터셋은 제 2 LUN으로 복사되고, 제 2 LUN은 호스트 컴퓨터에 노출되어 호스트 컴퓨터가 데이터셋과 대화할 수 있게 된다. 또한 데이터셋은 제 2 LUN으로부터 제 1 LUN의 파티션으로 되돌아가 피벗화될 수 있다.

대 표 도

명세서

청구범위

청구항 1

저장 어레이 제어기에서의 방법으로서,

저장 어레이로부터의 제 1 LUN(logical unit number)을 제 1 호스트 컴퓨터에 노출하는 단계와,

상기 제 1 LUN의 복수의 파티션 중 제 1 파티션과 관련된 피벗 요청을 상기 제 1 호스트 컴퓨터로부터 수신하는 단계- 상기 제 1 파티션은 상기 복수의 파티션 내에 저장된 복수의 데이터셋 중의 하나의 데이터셋을 저장함 - 와,

상기 제 1 LUN 내에서의 상기 제 1 파티션의 크기를 결정하는 단계와,

제 2 파티션에 대한 위치 및 크기를 표시하는 파티션 테이블을 생성하는 단계- 상기 제 2 파티션의 크기는 상기 결정된 제 1 파티션 크기보다 크거나 같음(≥) - 와,

제 2 LUN에 대한 서명, 상기 파티션 테이블 및 상기 제 2 파티션을 수용하도록 구성된 크기를 갖는 상기 저장 어레이의 상기 제 2 LUN을 생성하는 단계와,

상기 제 2 LUN에 상기 서명 및 상기 파티션 테이블을 저장하는 단계와,

상기 제 1 LUN으로부터 상기 제 1 파티션을 상기 제 2 LUN의 상기 파티션 테이블 내에서 상기 제 2 파티션에 대해 표시된 위치에 복사하는 단계와,

상기 제 2 LUN을 제 2 호스트 컴퓨터에 대해 매핑하는 단계를 포함하는

저장 어레이 제어기에서의 방법.

청구항 2

제 1 항에 있어서,

상기 데이터셋은 가상 머신을 위한 가상 하드 드라이브인

저장 어레이 제어기에서의 방법.

청구항 3

제 2 항에 있어서,

상기 매핑하는 단계는

상기 제 2 호스트 컴퓨터가 상기 제 2 LUN을 탑재(mount)할 수 있게 하는 단계와,

상기 제 2 호스트 컴퓨터가 상기 가상 머신을 실행하게 하고 상기 가상 머신을 위한 저장부로서 상기 가상 하드 드라이브에 액세스할 수 있게 하는 단계를 포함하는

저장 어레이 제어기에서의 방법.

청구항 4

제 1 항에 있어서,

상기 제 1 LUN 내에서의 상기 제 1 파티션의 크기를 결정하는 단계는

상기 파티션에 대해 상기 제 1 LUN 내에서의 시작 LBA(logical block address) 및 종료 LBA를 결정하는 단계를 포함하는

저장 어레이 제어기에서의 방법.

청구항 5

제 1 항에 있어서,

상기 복사하는 단계는

상기 제 1 파티션을 상기 제 1 LUN으로부터 상기 제 2 LUN으로 복사하기 위해 상기 제 1 파티션의 기록시 복사(copy-on-write) 스냅샷 또는 클론(clone)을 실행하는 단계를 포함하는

저장 어레이 제어기에서의 방법.

청구항 6

제 1 항에 있어서,

상기 제 2 LUN을 철회하는(retire) 요청을 상기 제 2 호스트 컴퓨터로부터 수신하는 단계와,

상기 제 2 파티션의 크기를 결정하는 단계와,

상기 제 1 LUN 내에 상기 제 2 파티션의 복사본을 저장하는 단계와,

상기 제 2 파티션의 상기 복사본이 저장되는 상기 제 1 LUN의 저장 영역에 대응하는 파티션을 포함하도록 상기 제 1 LUN의 파티션 테이블을 업데이트하는 단계를 더 포함하는

저장 어레이 제어기에서의 방법.

청구항 7

제 6 항에 있어서,

상기 제 1 LUN 내에 상기 제 2 파티션의 복사본을 저장하는 단계는

상기 제 2 파티션의 결정된 크기가 상기 제 1 LUN 내의 상기 제 1 파티션의 크기를 초과하지 않는다면, 상기 제 1 LUN의 상기 제 1 파티션 내에 상기 제 2 파티션의 상기 복사본을 저장하는 단계를 포함하는

저장 어레이 제어기에서의 방법.

청구항 8

제 7 항에 있어서,

상기 제 1 LUN 내에 상기 제 2 파티션의 복사본을 저장하는 단계는

상기 제 2 파티션의 결정된 크기가 상기 제 1 LUN 내의 상기 제 1 파티션의 크기를 초과하는 경우, 상기 제 1 LUN에 저장 영역을 추가함으로써 상기 제 1 LUN의 크기를 증대시키는 단계와, 상기 제 1 LUN의 상기 추가된 저장 영역 내에 상기 제 2 파티션의 상기 복사본을 저장하는 단계를 더 포함하는

저장 어레이 제어기에서의 방법.

청구항 9

제 1 호스트 컴퓨터에서의 방법으로서,

저장 어레이에 의해 상기 제 1 호스트 컴퓨터에 노출된 제 1 LUN(logical unit number)의 표시를 수신하는 단계와,

상기 제 1 LUN에 서명 및 파티션 테이블을 저장하는 단계와,

상기 제 1 LUN을 복수의 파티션으로 파티셔닝하는 단계와,

상기 복수의 파티션에 복수의 데이터셋을 저장하는 단계와,

상기 복수의 데이터셋 중의 하나의 데이터셋을 저장하는 상기 복수의 파티션 중의 하나의 파티션에 대한 액세스를 차단하는 단계와,

상기 데이터셋을 저장하는 상기 파티션과 관련된 퍼벗 요청을 상기 저장 어레이로 전송하는 단계- 상기 저장 어

레이이는 상기 피벗 요청에 응답하여 상기 파티션을 제2 LUN에 복사하고 상기 제2 LUN을 제2 호스트 컴퓨터에 맵핑하도록 구성됨 -

를 포함하는 제 1 호스트 컴퓨터에서의 방법.

청구항 10

제 9 항에 있어서,

상기 데이터셋은 가상 머신을 위한 가상 하드 드라이브인

제 1 호스트 컴퓨터에서의 방법.

청구항 11

제 10 항에 있어서,

상기 제 2 호스트 컴퓨터는 상기 제 2 LUN을 탐색(mount)하고, 상기 가상 머신을 실행하며, 상기 가상 머신에 대한 저장부로서 상기 가상 하드 드라이브에 액세스할 수 있는

제 1 호스트 컴퓨터에서의 방법.

청구항 12

제 9 항에 있어서,

상기 전송하는 단계는

상기 파티션에 대한 상기 제 1 LUN 내의 시작 LBA(logical block address) 및 종료 LBA를 상기 저장 어레이로 전송하는 단계를 포함하는

제 1 호스트 컴퓨터에서의 방법.

청구항 13

저장 어레이 제어기에서의 방법으로서,

저장 어레이의 제 1 LUN을 제 1 호스트 컴퓨터에 노출하고 상기 저장 어레이의 제 2 LUN을 제 2 호스트 컴퓨터에 노출하는 단계- 상기 제 1 LUN은 복수의 파티션을 포함하고 상기 제 2 LUN은 데이터셋을 저장하는 파티션을 포함함 -와,

상기 제 2 LUN을 철회하는(retire) 요청을 상기 제 2 호스트 컴퓨터로부터 수신하는 단계와,

상기 제 2 LUN에서의 상기 파티션의 크기를 결정하는 단계와,

상기 제 1 LUN 내에 상기 데이터셋의 복사본을 저장하는 단계와,

상기 데이터셋의 상기 복사본을 저장하는 상기 제 1 LUN의 저장 영역에 대응하는 파티션을 포함하도록 상기 제 1 LUN의 파티션 테이블을 업데이트하는 단계

를 포함하는 저장 어레이 제어기에서의 방법.

청구항 14

제 13 항에 있어서,

상기 제 1 LUN의 파티션에 상기 데이터셋의 복사본을 저장하는 단계는

상기 제 2 LUN에서의 상기 파티션의 결정된 크기가 상기 제 1 LUN 내의 제 1 저장 영역의 크기를 초과하지 않는다면, 상기 데이터셋을 이전에 저장한 상기 제 1 LUN의 상기 제 1 저장 영역에 상기 데이터셋의 상기 복사본을 저장하는 단계를 포함하는

저장 어레이 제어기에서의 방법.

청구항 15

제 14 항에 있어서,

상기 제 1 LUN의 파티션에 상기 데이터셋의 복사본을 저장하는 단계는

상기 제 2 LUN에서의 상기 파티션의 결정된 크기가 상기 제 1 LUN 내의 상기 제 1 저장 영역의 크기를 초과하는 경우, 상기 제 1 LUN에 제 2 저장 영역을 추가함으로써 상기 제 1 LUN의 크기를 증대시키는 단계와, 상기 제 1 LUN의 상기 제 2 저장 영역 내에 상기 데이터셋의 상기 복사본을 저장하는 단계를 더 포함하는

저장 어레이 제어기에서의 방법.

청구항 16

제 13 항에 있어서,

상기 데이터셋은 가상 머신을 위한 가상 하드 드라이브인

저장 어레이 제어기에서의 방법.

청구항 17

제 13 항에 있어서,

상기 제 2 LUN에서의 상기 파티션의 크기를 결정하는 단계는

상기 파티션에 대한 상기 제 2 LUN 내의 시작 LBA(logical block address) 및 종료 LBA를 결정하는 단계를 포함하는

저장 어레이 제어기에서의 방법.

청구항 18

제 13 항에 있어서,

상기 제 1 LUN에 상기 데이터셋의 복사본을 저장하는 단계는

상기 파티션을 상기 제 2 LUN으로부터 상기 제 1 LUN으로 복사하기 위해 상기 제 2 LUN 내의 상기 파티션의 기록시 복사(copy-on-write) 스냅샷 또는 클론을 수행하는 단계를 포함하는

저장 어레이 제어기에서의 방법.

발명의 설명

배경 기술

[0001]

컴퓨터 데이터 저장부는 소정 시간 동안 디지털 데이터를 유지하는 데 사용되는 구성 요소, 장치 및/또는 기록 매체를 지칭한다. RAM(random access memory) 장치 및 ROM(read only memory) 장치 등과 같은 메모리 장치(예를 들면, 반도체 저장부)와, 광학 디스크 및 자기 저장부(하드 디스크, 자기 테이프 등) 등과 같은 대용량 저장 장치와, 다른 종류의 저장부를 포함하는 여러 종류의 물리적 저장 장치가 존재한다. 이러한 저장 장치는 더 높은 레벨의 저장 및 더 높은 레벨의 저장 신뢰도를 제공하기 위해서 폴(pooled)될 수 있다. 예를 들면, 다수의 저장 장치를 결합하여 RAID(redundant array of independent disks) 어레이, SAN(storage area networks) 및 다른 집합 저장부를 형성할 수 있다.

[0002]

저장 가상화(storage virtualization)는 물리적 저장부로부터 가상 또는 논리 저장부를 추상화하는 하나의 방법이다. 가상 저장부를 제공하도록 구성된 저장 시스템은 컴퓨터에 컴퓨터가 액세스할 수 있는 논리적 저장 위치를 제공할 수 있다. 저장 시스템은 물리적 저장부 내의 물리적 저장 위치에 대해 논리적 저장 위치를 매핑하도록 구성된다. 한 가지 종류의 가상 저장부에서, 논리적 저장 장치는 LUN(logical unit number)으로 지칭될 수 있다. 컴퓨터는 컴퓨터가 액세스할 수 있는 저장 시스템에 의해 자신에게 노출된 하나 이상의 LUN을 가질 수 있다. 데이터의 블록은 LUN 내에서 LBA(logical block addresses)라고도 하는 해당 LUN 내의 오프셋(offsets) 또는 위치에 의해 어드레싱될 수 있다.

[0003]

가상 머신을 포함하는 여러 종류의 데이터가 가상 저장부 내에 저장될 수 있다. 가상 머신은 물리적 컴퓨터와 마찬가지로 프로그램을 실행하는 물리적 컴퓨터의 소프트웨어 구현이다. "시스템 가상 머신"은 운영 시스템을

포함하여 완전한 컴퓨터 플랫폼을 제공할 수 있다. 이러한 시스템 가상 머신은 하나 이상의 프로그램을 실행할 수 있다. "프로세스 가상 머신"은 단일 프로그램을 실행하도록 설계된다. 가상 머신을 나타내는 하나 이상의 파일은 가상 하드 드라이브로 지칭될 수 있다.

발명의 내용

해결하려는 과제

[0004]

일반적으로, 가상 머신에 대응하는 가상 이미지는 전형적으로 가상 저장 시스템의 격리된 LUN 내에 배치된다. 저장 시스템은 LUN을 호스트 컴퓨터에게 노출하여 호스트 컴퓨터가 가상 머신을 실행할 수 있게 한다. 많은 개수의 가상 머신(예를 들면, 수백, 수천 또는 그 이상의 가상 머신)이 호스트 컴퓨터에 의해 구현될 때, 호스트 컴퓨터는 가상 머신을 관리 및 인덱스화하기 위해 많은 양의 자원을 소모한다. 예를 들면, 하나의 현행 컴퓨터 파일 시스템은 255개 이상의 할당된 LUN을 지원할 수 없다. 더욱이 컴퓨터 파일 시스템의 성능은 50개의 LUN이 할당된 후에는 열악해질 수 있다. 따라서 종래의 기법에 따라서 LUN으로 구현된 가상 머신을 가지고 많은 개수의 가상 머신을 관리하는 것은 실제적으로 불가능하다.

과제의 해결 수단

[0005]

이 요약은 이하의 상세한 설명에서 더 설명되는 본 발명의 개념의 선택에 대해 단순화된 형태로 도입하고자 제시된 것이다. 이 요약은 청구 대상의 주요 특징 또는 근본 특징을 식별하도록 의도된 것이 아니고, 청구 대상의 범주를 제한하기 위해 사용하도록 의도된 것도 아니다.

[0006]

LUN(logical unit number) 내에 다수의 데이터셋을 저장하여 성능 향상을 가능하게 하는 방법, 시스템 및 컴퓨터 프로그램 제품이 제공된다. LUN은 다수의 데이터셋(예를 들면, 가상 하드 드라이브로서 저장된 가상 머신)을 저장할 수 있도록 제공된다. LUN은 다수의 파티션으로 분할될 수 있다. 하나 이상의 데이터셋은 각각의 파티션 내에 저장될 수 있다. 결과적으로, 다수의 데이터셋은 데이터셋의 개수에 비례하는 다수의 LUN을 통하는 것 보다는 단일 LUN을 통해 액세스될 수 있다.

[0007]

더욱이 LUN 내에 저장된 데이터셋은 피벗화될 수 있다. 일 실시예에서, 저장 어레이 제어기 내에서 실행하는 시스템 및 방법이 제공된다. 제 1 LUN은 저장 어레이로부터 제 1 호스트 컴퓨터로 노출된다. 피벗 요청은 제 1 LUN의 복수의 파티션 중 하나의 파티션(파티션은 복수의 파티션 중 어느 하나일 수 있음)에 관하여 제 1 호스트 컴퓨터로부터 수신된다. 파티션은 복수의 파티션 내에 저장된 복수의 데이터셋 중 하나의 데이터셋을 저장한다. 제 1 LUN 내에 파티션의 크기가 결정된다. 제 2 파티션의 위치 및 크기를 표시하는 파티션 테이블이 생성된다. 저장 어레이의 제 2 LUN은 제 2 LUN에 대한 서명, 파티션 테이블 및 제 2 파티션을 수용하도록 구성된 크기로 생성된다. 서명 및 파티션 테이블은 제 2 LUN 내에 저장된다. 제 1 LUN의 파티션은 제 1 LUN로부터 제 2 LUN로 복사되는데, 제 2 파티션에 대한 파티션 테이블 내에 표시된 위치로 복사된다. 제 2 LUN은 제 2 호스트 컴퓨터에 대해 매핑된다.

[0008]

다음에, 제 2 호스트 컴퓨터로부터 제 2 LUN을 철회(retire)하려는 요청이 수신될 수 있다. 제 2 파티션의 크기가 결정된다. 제 2 파티션의 복사본은 제 1 LUN 내에 저장된다. 제 1 LUN의 파티션 테이블은 제 2 파티션의 복사본이 저장되어 있는 제 1 LUN의 저장 영역에 대응하는 파티션을 표시하도록 업데이트된다.

[0009]

제 2 파티션의 크기가 제 1 LUN 내에 원래 피벗화된 파티션의 크기를 초과하지 않는다면, 제 2 파티션의 복사본은 제 1 LUN의 해당 파티션 내에 저장될 수 있다. 제 2 파티션의 크기가 제 1 LUN 내에 원래 피벗화된 파티션의 크기를 초과한다면, 제 1 LUN의 크기는 제 1 LUN에 저장 영역을 추가함으로써 성장할 수 있고, 제 2 파티션의 복사본은 제 1 LUN의 추가된 저장 영역 내에 저장될 수 있다.

[0010]

다른 실시예에서, 제 1 호스트 컴퓨터에서의 시스템 및 방법이 제공된다. 저장 어레이에 의해 제 1 호스트 컴퓨터에게 노출된 제 1 LUN(logical unit number)에 대한 표시가 수신된다. 서명 및 파티션 테이블은 제 1 LUN 내에 저장된다. 제 1 LUN은 복수의 파티션으로 분할된다. 복수의 데이터셋은 복수의 파티션 내에 저장된다. 복수의 데이터셋 중 하나의 데이터셋을 저장하는 복수의 파티션 중 하나의 파티션에 대한 액세스가 차단된다. 피벗 요청은 차단된 파티션에 관련된 저장 어레이로 송신된다. 저장 어레이는 피벗 요청에 응답하여 제 2 LUN

에 파티션을 복사하고, 제 2 호스트 컴퓨터에 제 2 LUN을 매핑하도록 구성된다.

[0011] 본 명세서 내의 상술된 부분 또는 다른 부분에서 설명된 방법을 실행 및/또는 가능하게 할 수 있고, 본 명세서에 설명된 다른 실시예를 구현하는 컴퓨터 시스템 및 컴퓨터 프로그램 제품(컴퓨터 판독 가능 매체에 저장됨)이 또한 본 명세서에 설명되어 있다.

[0012] 본 발명의 여러 실시예의 구조 및 동작뿐만 아니라 본 발명의 다른 특징 및 이점은 첨부된 도면을 참조하여 이하에 상세하게 설명되어 있다. 본 발명은 본 명세서에 설명된 특정 실시예로 한정되지 않는다는 것을 주지하라. 이러한 실시예는 오로지 예시를 목적으로 본 명세서에 제시되어 있다. 당업자라면 본 명세서에 포함된 개시 내용에 기초한 추가적인 실시예가 명확할 것이다.

[0013] 본 명세서에 포함되어 있고 본 명세서의 일부분을 형성하는 첨부된 도면은 본 발명을 설명하기 위한 것이고, 또한 그 상세한 설명과 함께 본 발명의 원리를 설명하고, 당업자들이 본 발명을 구성하고 사용할 수 있도록 하기 위해 제공되었다.

도면의 간단한 설명

[0014] 도 1은 복수의 컴퓨팅 장치가 공유형 저장부를 액세스하는 컴퓨팅 및 데이터 저장 시스템을 도시하는 도면.

도 2는 저장 어레이가 가상 저장부의 형태로 호스트 컴퓨터에 제시되어 있는 도 1에 도시된 컴퓨팅 및 데이터 저장 시스템을 도시하는 블록도.

도 3은 LUN(logical unit numbers)의 형태로 가상 저장부를 제공하는 저장 장치를 도시하는 블록도.

도 4는 복수의 가상 머신을 실행하는 호스트 컴퓨터를 도시하는 블록도.

도 5는 예시적인 실시예에 따라서 다수의 파티션을 갖는 LUN을 도시하는 블록도.

도 6은 예시적인 실시예에 따른 호스트 컴퓨터를 도시하는 블록도.

도 7은 예시적인 실시예에 따라서 LUN을 분할하는 프로세스를 제공하는 것을 도시하는 흐름도.

도 8은 예시적인 실시예에 따른 저장 어레이 제어기를 도시하는 블록도.

도 9는 제 1 호스트 컴퓨터로부터 제 2 호스트 컴퓨터로 데이터셋에 대한 액세스를 피벗화하는 프로세스를 도시하는 도면.

도 10은 예시적인 실시예에 따라서 제 1 호스트 컴퓨터로부터 제 2 호스트 컴퓨터로 데이터셋을 피벗화하는 프로세스를 제공하는 것을 도시하는 흐름도.

도 11은 예시적인 실시예에 따른 어레이 제어기를 도시하는 블록도.

도 12는 예시적인 실시예에 따른 컴퓨팅 및 데이터 저장 시스템을 도시하는 블록도.

도 13은 예시적인 실시예에 따라서 파티션이 피벗화될 수 있게 하는 프로세스를 제공하는 것을 도시하는 흐름도.

도 14는 예시적인 실시예에 따라서 피벗화된 LUN에 대한 액세스를 수신하는 프로세스가 제 2 호스트 컴퓨터에서 가능하게 되도록 제공하는 것을 도시하는 흐름도.

도 15는 예시적인 실시예에 따라서 제 2 호스트 컴퓨터로부터 제 1 호스트 컴퓨터로 데이터셋에 대한 액세스를 피벗화하는 프로세스를 도시하는 도면.

도 16은 예시적인 실시예에 따라서 제 2 호스트 컴퓨터로부터 제 1 호스트 컴퓨터로 데이터셋을 피벗화하는 프로세스를 제공하는 것을 도시하는 흐름도.

도 17은 예시적인 실시예에 따라서 도 12에 도시된 컴퓨팅 및 데이터 저장 시스템을 도시하는 블록도.

도 18은 예시적인 실시예에 따라서 파티션이 피벗화될 수 있게 하는 프로세스를 제공하는 것을 도시하는 흐름도.

도 19는 본 발명의 실시예를 구현하는데 사용될 수 있는 예시적인 컴퓨터를 도시하는 블록도.

발명을 실시하기 위한 구체적인 내용

[0015]

본 발명의 특징 및 이점은 그 전체에 걸쳐 동일한 참조 부호가 대응하는 구성 요소를 식별하는 도면과 함께 이하에 제시된 상세한 설명을 숙지함으로써 더 명확해질 것이다. 도면 내에서, 동일한 참조 번호는 일반적으로 동일한, 기능적으로 유사한 및/또는 구조적으로 유사한 구성 요소를 나타낸다. 소정의 구성 요소가 첫 번째로 나타나는 도면은 대응하는 참조 번호 내에 가장 좌측 숫자로 표시되어 있다.

I. 도입

[0016]

본 명세서는 본 발명의 특징부를 구체화하는 하나 이상의 실시예에 관해 개시한다. 개시된 실시예는 단지 본 발명의 예시에 불과하다. 본 발명의 범주는 개시된 실시예로 한정되지 않는다. 본 발명은 본 명세서에 첨부된 청구항에 의해 정의된다.

[0017]

명세서 내에서 "일실시예", "실시예", "예시적인 실시예" 등에 대한 참조는 개시된 실시예가 특정한 특징, 구조 또는 특성을 포함할 수 있다는 것을 나타내지만, 모든 실시예가 반드시 특정한 특징, 구조 또는 특성을 포함해야 한다는 것은 아니다. 더욱이 이러한 문구는 반드시 동일한 실시예를 지칭하지는 않는다. 또한 일실시예와 관련하여 특정한 특징, 구조 또는 특성이 기술될 때, 그것이 명확히 기술되었는지에 관계없이 당업자가 가질 수 있는 지식의 한도 내에서 다른 실시예와 관련하여 이러한 특징 구조 또는 특성이 달성될 수 있다는 것을 제안하는 것이다.

[0018]

더욱이, 본 명세서에 사용된 공간적 서술(예를 들면, "위", "아래", "상측", "좌측", "우측", "하부", "상부", "하부", "수직", "수평" 등)은 오로지 설명을 목적으로 한 것이고, 본 명세서에 기술된 구조에 대한 실제적인 구현은 임의의 방향 또는 방식으로 공간 정렬될 수 있다는 것을 이해할 것이다.

[0019]

II. 저장 시스템의 실시예

[0020]

본 발명의 실시예는 컴퓨터 데이터 저장부 내에 데이터를 저장하고, 저장된 데이터를 액세스하는 기법에 관한 것이다. 예를 들면, 도 1은 복수의 컴퓨팅 장치가 공유형 저장부를 액세스하는 컴퓨팅 및 데이터 저장 시스템(100)을 도시한다. 도 1에 도시된 바와 같이, 시스템(100)은 제 1 내지 제 3 호스트 컴퓨터(102a-102c), 통신 네트워크(104) 및 저장 어레이(106)를 포함한다. 저장 어레이(106)는 복수의 저장 유닛/저장 장치(114a-114n) 및 저장 통신 네트워크(108)를 포함한다. 도 1의 예시에서, 제 1 내지 제 3 호스트 컴퓨터(102a-102c)는 저장 어레이(106)의 저장 장치(114a-114n) 내에 데이터를 저장할 수 있는 것과, 그 곳에 저장된 데이터를 검색할 수 있는 것을 포함하여 저장 어레이(106)에 대한 액세스를 공유한다.

[0021]

도 1에는 3개의 호스트 컴퓨터(102a-102c)가 도시되어 있으나, 수백, 수천 또는 그 이상의 개수의 컴퓨팅 장치를 포함하는 임의의 개수의 호스트 컴퓨터(102)가 네트워크(104)에 결합되어 저장 어레이(106)를 공유할 수 있다. 호스트 컴퓨터(102)의 예시는 고정형 및 이동형 컴퓨팅 장치를 포함한다. 예를 들면, 각각의 호스트 컴퓨터(102a-102c)는 서버, 데스크탑 컴퓨터(예를 들면, 퍼스널 컴퓨터), 모바일 컴퓨터(예를 들면, PDA(personal digital assistant), 랩톱 컴퓨터, 노트북 컴퓨터, 스마트폰 등) 또는 다른 종류의 컴퓨팅 장치일 수 있다.

[0022]

각각의 호스트 컴퓨터(102a-102c)는 네트워크(104) 및 대응하는 통신 링크를 통해 저장 어레이(106)와 통신하는 것으로 도 1에 도시되어 있다. 예를 들면, 도 1에 도시된 바와 같이 호스트 컴퓨터(102a)는 제 1 통신 링크(110a)를 통해 네트워크(104)에 통신 가능하게 결합되고, 호스트 컴퓨터(102b)는 제 2 통신 링크(110b)를 통해 네트워크(104)에 통신 가능하게 결합되고, 호스트 컴퓨터(102c)는 제 3 통신 링크(110c)를 통해 네트워크(104)에 통신 가능하게 결합된다. 저장 어레이(106)는 제 4 통신 링크(110d)를 통해 네트워크(104)에 통신 가능하게 결합된 것으로 도시되었다. 네트워크(104)는 LAN, WAN(wide area network)이거나, 인터넷 등과 같은 네트워크의 조합일 수 있다. 제 1 내지 제 4 통신 링크(110a-110d)는 IEEE 802.11 WLAN(wireless LAN) 무선 링크, Wi-MAX(Worldwide Interoperability for Microwave Access) 링크, 셀룰라(cellular) 네트워크 링크, 무선 PAN(personal area network) 링크(예를 들면, 블루투스™ 링크), 이더넷(Ethernet) 링크, USB 링크 등과 같은 유선 및/또는 무선 링크를 포함하는 임의의 종류의 통신 링크 또는 통신 링크의 조합을 포함할 수 있다.

[0023]

저장 장치(114a-114n)가 도 1에 도시되어 있으나, 수백, 수천 또는 그 이상의 개수의 저장 장치를 포함하는 임의의 개수의 저장 장치(114)가 데이터를 저장하기 위해 저장 어레이(106) 내에 포함될 수 있다. 어레이 제어기

(112) 및 저장 장치(114a-114n)는 저장 통신 네트워크(108)에 의해 서로 통신 가능하게 결합된 것으로 도 1에 도시되어 있다. 저장 장치(114a-114n) 및 저장 통신 네트워크(108)는 SAN(storage area network), 섬유 채널(Fibre Channel) 어레이, NAS(network attached storage) 등을 포함하는 임의의 구성으로 저장 어레이(106) 내에 정렬 또는 연관될 수 있다. 저장 장치(114)는 메모리 장치 및/또는 대용량 저장 장치를 포함하는 임의의 종류의 저장 장치, 휘발성 및/또는 비휘발성일 수 있다. 저장 장치(114)의 예시는 RAM(random access memory) 장치(휘발성 또는 비휘발성) 및 ROM(read only memory) 장치 등과 같은 메모리 장치(예를 들면, 반도체 저장부)와, 광학 디스크 및 자기 저장부(하드 디스크, 자기 테이프 등) 등과 같은 대용량 저장 장치 및 다른 종류의 저장부를 포함한다.

[0025] 저장부 가상화는 물리적 저장부로부터 논리적 저장부를 추상화하는 방법이다. 가상 저장부를 제공하도록 구성된 저장 시스템은 컴퓨터가 액세스할 수 있는 논리적 저장 위치를 컴퓨터에 제공할 수 있다. 저장 시스템은 논리적 저장 위치를 물리적 저장부 내의 물리적 저장 위치로 매핑하도록 구성된다.

[0026] 예를 들면, 도 1에 도시된 바와 같이 저장 어레이(106)는 어레이 제어기(112)를 포함할 수 있다. 어레이 제어기(112)는 저장 장치(114a-114n)의 물리적 저장부에 기초하여 호스트 컴퓨터(102a-102c)에 논리적 저장부를 할당하도록 구성될 수 있다. 예를 들면, 어레이 제어기(112)는 저장 어레이(106)의 물리적 저장 장치의 세트를 결합하여 RAID(redundant array of independent disks) 어레이 또는 세트를 생성하도록 구성될 수 있다. RAID 세트는 저장 장치(114a-114n)의 논리적 구성체이다. 저장 어레이(106)는 저장 어레이(106)를 위한 여러 종류의 RAID 세트로 구성될 수 있다. RAID 세트의 하나의 예시로는 "RAID 1"이 있고, 여기에서 저장 장치(114a-114n)는 거울 형상(mirror imaging) 저장 장치를 포함한다. 다른 종류의 RAID 세트는 이와 다르게 RAID 0, RAID 5, RAID 6 및 RAID 10 구성을 포함하는 저장 어레이(106) 내에 포함될 수 있는데, 이들은 당업자들에게 잘 알려져 있다. 어레이 제어기(112)는 RAID 세트 중 사용 가능한 저장 공간의 몇몇 부분을 이용하여 LUN(logical unit number)로 지칭되는 하드 드라이브의 가상 표현을 생성할 수 있다. 각각의 컴퓨터(102a-102c)는 저장 액세스를 위해서 어레이 제어기(112)에 의해 자신에게 노출된 하나 이상의 LUN을 가질 수 있다.

[0027] 예를 들면, 도 2는 도 1에 도시된 컴퓨팅 및 데이터 저장 시스템(100)에 대한 블록도를 도시하는데, 여기에서 저장 어레이(106)는 가상 저장부(202)의 형태로 어레이 제어기(112)에 의해 호스트 컴퓨터(102a-102c)로 제공된다. 도 2에 도시된 바와 같이, 가상 저장부(202)는 복수의 LUN(204a-204m)을 포함한다. LUN(204a-204m)은 컴퓨터(102a-102c) 중 하나에 대해 가상 저장부로서 제공될 수 있는 저장 장치(114a-114n)(도 1에 도시됨)의 저장부의 일부분으로서 어레이 제어기(112)에 의해 각각 생성된다. 예를 들면, 도 3은 도 1에 도시된 저장 장치(114a-114n) 중 하나에 대한 예시인 저장 장치(300)에 대한 블록도를 도시한다. 도 3에 도시된 바와 같이, 저장 장치(300)는 복수의 LUN(302a-302o)으로 논리적으로 분리된다. 각각의 LUN(302)은 전체 하드 디스크, 하드 디스크의 부분(예를 들면, 블록 범위), 스팬들(spindle), 메모리 장치의 메모리 영역 등을 포함하는 저장 장치(300)의 임의의 부분일 수 있다. 각각의 LUN(302)은 호스트 컴퓨터에게 노출될 수 있다. 호스트 컴퓨터의 운영 시스템은 그것이 물리적 저장부인 것처럼 LUN(302)과 대화한다.

[0028] 다시 도 2를 참조하면, LUN(204a-204m)이 도시되어 있으나, 수백, 수천 또는 그 이상의 개수의 LUN(204)을 포함하는 임의의 개수의 LUN(204)이 저장 어레이(106) 내에 포함될 수 있다. 더욱이 임의의 개수의 LUN(204)이 컴퓨터(102a-102c) 각각에 대해 제공될 수 있다. 각각의 LUN(204)은 복수의 어드레스 가능한 데이터 블록을 포함한다. LUN(204a-204m) 내의 특정한 데이터 블록은, 특정 LUN을 식별함으로써, 또한 하나 이상의 LBA(logical block addresses)의 형태로 특정 LUN 내에 오프셋 또는 위치를 제공함으로써 컴퓨터(102)에 의해 어드레스될 수 있다.

[0029] 여러 종류의 데이터 및/또는 애플리케이션은 가상 저장부(202) 내에 저장될 수 있다. 이러한 데이터 및 애플리케이션은 본 명세서에서 "데이터셋"으로 지칭된다. 데이터셋의 하나의 예시로는 가상 머신이 있다. 가상 머신은 물리적 컴퓨터와 유사하게 프로그램을 실행하는 물리적 컴퓨터의 소프트웨어 구현이다. "시스템 가상 머신"은 운영 시스템을 포함하는 완전한 컴퓨터 플랫폼을 제공할 수 있다. 이러한 시스템 가상 머신은 하나 이상의 프로그램을 실행할 수 있다. "프로세스 가상 머신"은 단일 프로그램을 실행하도록 설계된다. 가상 머신의 예시는 SQL 서버, 웹 서버 등을 포함한다. 가상 머신을 나타내는 하나 이상의 파일은 "가상 이미지"로 지칭될 수 있다.

[0030] 일반적으로, 가상 머신에 대응하는 가상 이미지는 가상 저장 시스템의 격리형 LUN(예를 들면, 도 2에 도시된 LUN(204) 중 하나) 내에 배치된다. 이러한 가상 저장 시스템에서, LUN은 호스트 컴퓨터가 가상 머신을 실행할 수 있도록 호스트 컴퓨터에 대해 노출된다. 많은 개수의 가상 머신(예를 들면, 수백, 수천 또는 그 이상의 가

상 머신)이 호스트 컴퓨터에 의해 구현될 때, 동일한 개수의 LUN은 가상 머신 이미지를 포함하는 호스트 컴퓨터에 대해 노출된다.

[0031] 예를 들면, 도 4는 복수의 가상 머신(402a-402p)을 실행하는 호스트 컴퓨터(102a)에 대한 블록도를 도시한다. 더욱이 도 4는 가상 머신(402a-402p)을 위한 저장부로서 컴퓨터(102a)에 할당된 복수의 LUN(204a-204p)을 포함하는 가상 저장부(202)를 도시한다. 도 4에 도시된 바와 같이, 각각의 LUN(204a-204p)은 가상 하드 드라이브(404a-404p) 중 대응하는 것을 포함한다. 가상 하드 드라이브(404a-404p)는 가상 머신(402a-402p)에 대응하는 데이터를 위한 가상 저장 위치이다. 예를 들면, 가상 하드 드라이브(404a-404p)는 가상 머신(402a-402p) 중 대응하는 것으로서 컴퓨터(102a)에 의해 실행되는 하나 이상의 파일을 각각 저장할 수 있다. 더욱이 가상 하드 드라이브(404a-404p)는 가상 머신(402a-402p) 중 대응하는 것을 위한 가상 저장부를 각각 제공할 수 있다.

[0032] 가상 머신은 가상 저장부로부터 물리적 저장부로 몇몇 레벨의 추상화를 통해 저장부를 액세스한다. 몇몇 가상 머신은 또한 "가상 서버"로 지칭될 수 있다. 작동 중에, 가상 머신(402a)은 그 가상 하드 드라이브(404a)에 데이터의 블록을 기록하도록 시도할 것이다. 컴퓨터(102a)의 운영 시스템(OS)은 데이터 블록 기록 동작을 차단할 수 있고, 데이터의 블록이 LUN(204a)에 대하여 기록되어야 하는 위치를 결정하기 위해 변환을 수행할 수 있다. 예를 들면, 가상 머신(402a)은 가상 하드 드라이브(404a)에 대해 394의 LBA를 갖는 데이터 블록을 기록하도록 시도할 수 있다. OS는 가상 하드 드라이브(404a) 내의 데이터 블록 오프셋(394)이 LUN(204a) 내에서 9942의 LBA에 대응한다고 결정할 수 있다. 따라서 OS는 LUN(204a)의 LBA 9942에 데이터 블록을 기록하도록 시도할 수 있다. 어레이 제어기(112)(도 1)는 호스트 컴퓨터(102a)의 OS로부터 LBA 9942에 대한 기록 요청을 수신할 수 있고, 또한 LBA 9942에 대응하는 저장 어레이(106) 내의 실제 물리적 위치(예를 들면, 저장 장치(114a)의 스팬들 47 및 스팬들 48 모두에 있는 실제 스팬들 블록 12345 등)에 대해 데이터 블록을 기록할 수 있다.

[0033] 호스트 컴퓨터(102a)는 LUN(204a-204p) 및 가상 머신(402a-402p)을 관리 및 인덱스화하기 위해, 수백, 수천 또는 그 이상의 숫자가 될 수 있는 많은 양의 자원을 소모한다. LUN(204a-204p)을 물리적 저장부인 것처럼 다른 호스트 컴퓨터(102a)는, LUN(204a-204p)을 모니터링하여, 본질적으로 LUN(204a-204p)을 관리하는 것에 대해 적절히 동작하는지 확인한다. 더욱이, 저장 어레이(106) 내에서 각각의 저장 장치로 향하는 여러 경로가 존재할 수 있고, 호스트 컴퓨터(102)는 각각의 경로를 또한 관리하도록 시도할 수 있다. 이것의 결과로 관리하는데 있어서 호스트 컴퓨터(102a)에게 있어서 많은 양의 비용이 소모된다. 서버가 단일 첨부된 LUN을 가지고 8개의 관련 장치를 관리하는 것은 일반적이다. 많은 개수의 LUN(및 대응하는 다수의 경로)가 서버에게 노출되면, 서버는 경로 및 저장 장치의 세트를 발견하는데 많은 양의 시간을 소요할 것이다. 충분히 많은 개수의 노출된 LUN을 가지고, 서버는 무반응 및/또는 충돌하게 될 수 있다. 따라서 격리형 LUN으로 구현된 가상 머신을 가지고 매우 많은 개수의 가상 머신을 관리하는 것은 실제적으로 불가능하다.

[0034] 일실시예에서, LUN은 다수의 가상 머신(가상 하드 드라이브로서 저장됨)을 포함하여 다수의 데이터셋을 저장하도록 제공될 수 있다. LUN은 다수의 파티션으로 분할될 수 있다. 하나 이상의 데이터셋은 각각의 파티션 내에 저장될 수 있다. 결과적으로, 다수의 데이터셋에 액세스할 필요가 있는 호스트 컴퓨터는 데이터셋의 개수에 비례하는 개수의 LUN을 통하는 것이 아니라 단일 LUN을 통해서 다수의 데이터셋을 액세스할 수 있다. 이러한 방식으로, 호스트 컴퓨터에 의해 추적되는 하드웨어(예를 들면, 호스트 컴퓨터에 의해 하드웨어로 간주되는 LUN)의 양이 감소된다.

[0035] 예를 들면, 도 5는 예시적인 실시예에 따른 LUN(500)에 대한 블록도를 도시한다. 도 5에 도시된 바와 같이, LUN(500)은 복수의 파티션(506a-506p)을 포함한다. 더욱이 각각의 파티션(506)은 데이터셋(508a-508p) 중 대응하는 것을 저장한다. 이와 다르게, 하나 이상의 데이터셋(508)이 각각의 파티션(506) 내에 저장될 수 있다. 다수의 데이터셋이 LUN 내에 저장되며 함으로써, 호스트 컴퓨터가 동일한 개수 또는 (예를 들면, LUN 당 파티션의 개수를 인수로 하는) 더 많은 개수의 데이터셋을 액세스할 수 있도록 호스트 컴퓨터에 더 적은 개수의 LUN이 할당될 수 있다.

[0036] 실시예에서, 어레이 제어기 또는 호스트 컴퓨터는 도 5에 도시된 바와 같이 LUN을 분할할 수 있다. 예를 들면, 도 6은 예시적인 실시예에 따른 호스트 컴퓨터(602)의 블록도를 도시한다. 호스트 컴퓨터(602)는 컴퓨터(102a-102c) 중 하나에 대한 예시이다. 도 6에 도시된 바와 같이, 호스트 컴퓨터(602)는 LUN 파티셔너(604)를 포함한다. LUN 파티셔너(604)는 호스트 컴퓨터(602)가 도 5의 LUN(500) 등과 같은 LUN을 다수의 파티션(506)으로 분할할 수 있게 하도록 구성된다.

[0037] LUN 파티셔너(604)는 여러 방식으로 LUN을 분할하도록 구성될 수 있다. 예를 들면 도 7은 예시적인 실시예에 따라서 LUN을 분할하는 프로세스를 제공하는 흐름도(700)를 도시한다. 일실시예에서, 흐름도(700)는 도 6에 도

시된 호스트 컴퓨터(602)에 의해 실행될 수 있다. 당업자라면 흐름도(700)와 관련한 설명에 기초하여 다른 구조적 및 동작적 실시예가 명확할 것이다. 흐름도(700)는 다음과 같이 설명된다.

[0038] 도 7에 도시된 바와 같이, 흐름도(700)는 단계(702)에서 시작된다. 단계(702)에서, LUN(logical unit number)이 저장 어레이에 의해 제 1 호스트 컴퓨터에 대해 노출되었다는 통지를 수신한다. 예를 들면, 일실시예에서 호스트 컴퓨터(602)는 도 5에 도시된 LUN(500) 등과 같은 LUN이 호스트 컴퓨터(602)에 대해 노출되었다는 통지를 도 2에 도시된 어레이 제어기(112) 등과 같은 어레이 제어기로부터 수신할 수 있다. 호스트 컴퓨터(602)가 가상 저장부로서 LUN(500)을 액세스할 수 있도록, LUN(500)이 호스트 컴퓨터(602)에 노출된다. LUN(500)은 LUN 식별자(예를 들면, 식별 번호) 및 LUN(500)에 의해 제공된 저장의 양으로 호스트 컴퓨터(602)에 표시될 수 있다. 예를 들면, LBA(logical block addresses)의 범위는 LUN(500)의 저장의 양을 표시하기 위해서 호스트 컴퓨터(602)에 제공될 수 있다.

[0039] 단계(704)에서, 서명 및 파티션 테이블이 LUN 내에 저장된다. 예를 들면, 도 5에 도시된 바와 같이, LUN 파티셔너(604)는 LUN(500) 내에 서명(502) 및 파티션 테이블(504)을 저장할 수 있다. 서명(502)은 LUN(500)을 위한 디지털 및/또는 디스크 서명이고, LUN(500)을 위한 식별 정보를 포함하는 여러 정보를 포함할 수 있다. 파티션 테이블(504)은 LUN 파티셔너(604)에 의해 LUN(500) 내에 형성된 파티션의 위치 및 크기를 표시하도록 구성되는 데이터 구조이다.

[0040] 단계(706)에서, LUN은 복수의 파티션으로 분할된다. 일실시예에서, LUN 파티셔너(604)는 파티션 LUN(500)을 복수의 파티션(506a-506p)으로 분할하도록 구성된다. LUN(500)은 호스트 컴퓨터(602)에 의해 요구되는 것만큼 또한 LUN(500)의 저장 양에 맞는 만큼의 임의의 개수의 파티션(506)을 구비할 수 있다. LUN 파티셔너(604)는 파티션(506)을 생성하고, 파티션 테이블(504) 내에 파티션(506)을 표시한다. 예를 들면, LUN 파티셔너(604)는 베이스 LBA(예를 들면, LBA 0)로부터의 LBA 오프셋의 형태로 파티션 테이블(504) 내에 파티션(506a-506p)을 표시할 수 있다. 예를 들면 파티션 테이블(504)은 서명(502)을 위한 LUN(500) 내에서의 LBA 범위(예를 들면, LBA 0 내지 LBA 600), 파티션 테이블(504)을 위한 LBA 범위(예를 들면, LBA 601 내지 LBA 10000), 파티션(506a)을 위한 LBA 범위(예를 들면, LBA 10001 내지 LBA 123456), 파티션(506b)을 위한 LBA 범위(예를 들면, LBA 123457 내지 LBA 654321), 파티션(506c)을 위한 LBA 범위(예를 들면, LBA 654322 내지 LBA 999999) 등을 표시할 수 있다.

[0041] 단계(708)에서, 복수의 데이터셋은 복수의 파티션 내에 저장된다. 예를 들면, 도 5에 도시된 바와 같이 호스트 컴퓨터(602)는 각각의 파티션(506a-506p) 내에 하나의(하나 이상의) 데이터셋(508a-508p)을 저장할 수 있다. 상술된 바와 같이, 데이터셋(508)은 가상 머신(가상 머신 하드 드라이브) 등과 같이 데이터 및/또는 애플리케이션을 형성하는 하나 이상의 파일일 수 있다. 예를 들면, 도 4에 도시된 호스트 컴퓨터(102a)와 마찬가지로, 호스트 컴퓨터(602)는 도 5에 도시된 데이터셋(508a-508p) 중 하나로서 저장된 대응하는 가상 하드 드라이브를 각각 구비하는 복수의 가상 머신(402a-402p)을 작동시킬 수 있다. 다수의 가상 하드 드라이브가 (데이터셋(508)으로서) LUN(500) 내에 저장될 수 있으므로, 호스트 컴퓨터(602)는 더 적은 개수의 LUN을 가지고 동일한 또는 훨씬 더 많은 개수의 가상 머신을 관리하여, 호스트 컴퓨터(602)가 더 높은 레벨의 저장 성능을 가지고 더 많은 개수의 가상 머신을 지원할 수 있게 한다.

[0042] 더욱이 다수의 데이터셋을 저장하는 LUN으로부터 데이터셋 및/또는 파티션을 "폐벽화"하는 기능이 가능해진다. 예를 들면, LUN(500)은 데이터 저장부의 형태로서 다수의 파티션 내에 데이터셋(508a-508p)을 저장할 수 있고, 개별적인 데이터셋(508a-508p)은 LUN(500)으로부터 다른 호스트 컴퓨터와 연관된 각각의 LUN으로 복사될 수 있다. 다른 호스트 컴퓨터는 각각의 LUN에서 데이터셋을 액세스할 수 있다. 이와 같이, 데이터셋의 소유권은 제 1 호스트 컴퓨터로부터, 전용 LUN의 형태로 데이터셋에 대한 액세스를 제공받은 제 2 호스트 컴퓨터로 폐벽화될 수 있다. 더욱이, 제 2 호스트 컴퓨터가 더 이상 데이터셋에 대한 액세스를 필요로 하지 않을 때, 데이터셋을 전용 LUN으로부터 멀티-파티션 LUN으로 다시 복사함으로써, 제 2 호스트 컴퓨터는 데이터셋의 소유권을 제 1 호스트 컴퓨터로 다시 폐벽화할 수 있다.

[0043] LUN 내에 다수의 데이터셋을 저장하고, 데이터셋에 대한 액세스를 다른 LUN으로 폐벽화하는 기능은, 호스트 컴퓨터의 가상 머신 관리자가 비교적 적은 개수의 가상 머신을 관리하는 것으로부터 수천, 수십만 또는 그 이상의 개수의 가상 머신을 관리하는 것으로 확대될 수 있게 한다. 더욱이 저장 어레이에는 멀티-파티션 LUN 내에서 현재 사용되지 않는 LUN을 대기시킬 수 있다. 일반적으로, 상업적으로 입수 가능한 저장 어레이에는 고정된 개수의 LUN을 저장할 수 있다. 예를 들면, 미국 메사츠세츠주의 흉킨톤에 소재한 이엠씨 코포레이션(EMC Corporation)에서 생산된 이엠씨 클라리온(EMC CLARIION)™ 제품군은 오로지 4096개의 LUN만을 지원할 수 있다. 필요할

때까지 데이터셋이 대기될 수 있는 본 명세서에 개시된 실시예에 따르면, 저장 어레이는 통상적으로 가능한 것 보다 10배 이상의 LUN을 저장할 수 있다. 예를 들면, 특정 시간에 4096개의 LUN이 저장 어레이에 의해 노출될 수 있으나, 장기간 저장부에 있어서는 40,000개의 데이터셋이 가능해진다.

[0044] 표준 파일 서버는 데이터셋의 배포에 사용될 수 있는 네트워크 공유로 많은 데이터셋을 저장할 수 있다는 것을 주지하라. 그러나 네트워크 부착형 파일 시스템은 상술된 바와 같이 파일 시스템에 대한 원래의(예를 들면, 블록 레벨의) 액세스에 비교할 때 그 속도 및 기능이 결여되어 있다. 네트워크 부착형 파일 시스템에서 이용 가능하지 않은 여러 고도의 기능이 블록 레벨 액세스 시스템에서 가능하게 된다. 예를 들면, 네트워크 부착형 파일 시스템은 가상 하드 드라이브가 더 이상 순차적이지 않게 하는 방식으로 드라이브의 세트에 걸쳐 가상 하드 드라이브의 세트를 분산할 수 있다. 이와는 대조적으로 블록 레벨 액세스 시스템에서, 서로 다른 가상 하드 드라이브는 LUN의 각각의 파티션 내에 배치될 수 있다. 각각의 파티션에 대한 시작 블록 및 종료 블록이 알려져 있으므로, 완전한 블록 레벨 지역성(locality)이 달성된다. 결과적으로, 호스트 컴퓨터에 노출된 LUN의 개수가 감소되고, 순차성과 성능의 지역성이 획득된다.

[0045] 멀티-파티션 LUN의 데이터셋/파티션을 다른 LUN에 대해 피벗화하는 것의 예시는 다음의 단락에서 설명되어 있고, LUN의 데이터셋을 멀티-파티션 LUN의 파티션으로 피벗화하는 것의 예시는 그 다음 단락에 설명되어 있다.

III. 데이터셋/LUN 파티션을 LUN으로 피벗화하는 예시적인 실시예

[0047] 일실시예에서, 데이터셋에 대한 액세스는 제 1 호스트 컴퓨터로부터 제 2 호스트 컴퓨터로 피벗화될 수 있다. 예를 들면, 제 2 호스트 컴퓨터가 제 1 호스트 컴퓨터에 비해서 덜 바쁘다는 이유 또는 다른 이유로, 데이터셋의 피벗화는 제 2 호스트 컴퓨터에 있는 데이터셋에 대한 액세스를 제공하기 위해 실행될 수 있다. 이러한 피벗화는 저장 어레이의 어레이 제어기가 피벗화를 실행하도록 구성하는 것을 포함하여 여러 방식으로 실행될 수 있다. 예를 들면, 도 8은 예시적인 실시예에 따른 어레이 제어기(802)에 대한 블록도를 도시한다. 어레이 제어기(802)는 상술된 어레이 제어기(112)(도 1 및 도 2)와 유사하지만 다음에 설명되는 차이점을 갖고 있다. 도 8에 도시된 바와 같이, 어레이 제어기(802)는 데이터셋 피벗 모듈(804)을 포함한다. 데이터셋 피벗 모듈(804)은 어레이 제어기(802)가 데이터셋에 대한 액세스를 피벗화할 수 있게 한다. 예를 들면, 데이터셋 피벗 모듈(804)은 도 9에 도시된 단계(902)를 실행하도록 구성될 수 있다. 단계(902)에서, 데이터셋을 포함하는 LUN 파티션을 제 2 LUN으로 매핑함으로써 데이터셋에 대한 액세스는 제 1 호스트 컴퓨터로부터 제 2 호스트 컴퓨터로 피벗화된다.

[0048] 어레이 제어기(802)는 여러 방식으로 단계(902)를 실행하도록 구성될 수 있고, 단계(902)는 여러 방식으로 실행될 수 있다. 예를 들면 도 10은 예시적인 실시예에 따라서 제 1 호스트 컴퓨터로부터 제 2 호스트 컴퓨터로 데이터셋을 피벗화하는 프로세서를 제공하는 흐름도(1000)를 도시한다. 흐름도(1000)는 도 9에 도시된 단계(902)의 예시이고, 일실시예에서 도 8에 도시된 어레이 제어기(802)에 의해 실행될 수 있다. 예를 들면, 도 11은 예시적인 실시예에 따른 어레이 제어기(802)의 블록도를 도시한다. 흐름도(1000)는 오로지 예시를 위해서 도 11에 도시된 어레이 제어기(802)의 예시를 참조하여 설명된다. 도 11에 도시된 바와 같이, 어레이 제어기(802)는 LUN 할당기(1102) 및 데이터셋 피벗 모듈(804)을 포함한다. 데이터셋 피벗 모듈(804)은 피벗 요청 수신기(1104), LUN 생성기(1106) 및 파티션 복사기(1108)를 포함한다. 당업자라면 흐름도(1000)에 관련된 설명을 기초하여 다른 구조적 및 동작적 실시예가 명확할 것이다.

[0049] 설명을 목적으로, 흐름도(1000)는 예시적인 실시예에 따라서 컴퓨팅 및 데이터 저장 시스템(1200)의 블록도를 도시하는 도 12와 관련하여 다음과 같이 설명되어 있다. 도 12에 도시된 바와 같이, 시스템(1200)은 제 1 및 제 2 호스트 컴퓨터(102a, 102b) 및 저장 어레이(1202)를 포함한다. 저장 어레이(1202)는 가상 저장부(1204) 및 어레이 제어기(802)를 포함한다. 가상 저장부(1204)는 제 1 LUN(500) 및 제 2 LUN(1206)(추가적인 LUN은 간결성을 목적으로 도 12에 도시되지 않은 상태로 표시될 수 있음)을 포함한다. 도 12는 제 1 호스트 컴퓨터(102a)로부터 제 2 호스트 컴퓨터(102b)로 피벗화된 데이터셋(508a)을 도시한다.

[0050] 도 10에 도시된 바와 같이, 흐름도(1000)는 단계(1002)에서 시작한다. 단계(1002)에서, 제 1 LUN(logical unit number)은 저장 어레이로부터 제 1 호스트 컴퓨터로 노출된다. 예를 들면, 일실시예에서 단계(1002)는 LUN 할당기(1102)에 의해 실행될 수 있다. 도 12를 참조하면, LUN 할당기(1102)는 (예를 들면, 도 1 및 도 2에 도시된 것과 같은 통신 매체를 통해서) 컴퓨터(102a)로 노출된 LUN 표시 신호(1218)를 송신할 수 있는데, 이 신호는 제 1 LUN(500)이 컴퓨터(102a)에 할당되었다는 것을 표시한다. 예를 들면, 노출된 LUN 표시 신호(1218)는

제 1 LUN(500)을 위한 식별자(예를 들면, 식별 번호, 어드레스 등) 및 제 1 LUN(500) 내에 포함된 저장의 양을 포함할 수 있다.

[0051] 단계(1002) 이후에, 제 1 호스트 컴퓨터(102a)는 LUN(500) 내에 데이터를 저장하고, LUN(500)으로부터 데이터를 검색하는 것을 포함하여 제 1 LUN(500)의 저장부를 액세스할 수 있다. 예를 들면, 제 1 호스트 컴퓨터(102a)는 LUN(500)의 표시를 수신하는 단계(단계(702)), LUN(500) 내에 서명(502) 및 파티션 테이블(504)을 저장하는 단계(단계(704)), LUN(500)을 분할하여 파티션(506a, 506b, ...)을 생성하는 단계(단계(706)) 및 파티션(506a, 506b, ...) 내에 데이터셋(508a, 508b, ...)을 각각 저장하는 단계(단계(708))를 포함하는 상술된 흐름도(700)(도 7)를 실행할 수 있다.

[0052] 특정한 시간에, 제 1 호스트 컴퓨터(102a)는 LUN(500)의 데이터셋이 다른 호스트 컴퓨터로 피벗화된다고 결정할 수 있다. 예를 들면, 덜 바쁜 호스트 컴퓨터로 데이터셋을 피벗화하는 것이 바람직할 것이다. 이와 같이, 제 1 호스트 컴퓨터(102a)는 도 13에 도시된 흐름도(1300)를 실행할 수 있다. 흐름도(1300)는 다음과 같이 설명된다. 흐름도(1300)의 단계(1302)에서, 복수의 데이터셋 중 하나의 데이터셋을 저장하는 복수의 파티션 중 하나의 파티션에 대한 액세스가 차단된다. 예를 들면, 제 1 호스트 컴퓨터(102a)는 데이터셋(508a)을 저장하는 제 1 파티션(506a)에 대한 액세스를 차단하여, 데이터셋(508a)이 피벗화되도록 준비할 수 있다. 일실시예에서, 제 1 호스트 컴퓨터(102a)는 데이터셋(508a)에 대한 자신의 개방된 핸들을 차단할 수 있고, 데이터셋(508a)이 더 이상 액세스되지 않을 때, 제 1 호스트 컴퓨터(102a)의 논리적 디스크 관리자는 드라이브 문자(drive letter)를 제거하고 제 1 파티션(506a)으로 매핑할 수 있다. 이것은 제 1 호스트 컴퓨터(102a)가 제 1 파티션(506a)에 대한 판독 및/또는 기록을 중지하도록 보장한다. 단계(1304)에서, 피벗 요청은 데이터셋을 저장하는 파티션에 판하여 저장 어레이로 송신된다. 예를 들면, 도 12에 도시된 바와 같이, 제 1 호스트 컴퓨터(102a)는 (예를 들면, 도 1 및 도 2에 도시된 바와 같은 통신 매체를 통해서) 저장 어레이(1202)로 피벗 요청 신호를 송신하여 데이터셋(508a)이 다른 호스트 컴퓨터로 피벗화되도록 요청할 수 있다.

[0053] 다시 흐름도(1000)(도 10)를 참조하면, 단계(1004)에서 제 1 LUN의 복수의 파티션의 데이터셋을 저장하는 제 1 파티션과 관련하여 피벗 요청이 제 1 호스트 컴퓨터로부터 수신된다. 예를 들면, 일실시예에서 단계(1004)는 피벗 요청 수신기(1104)에 의해 실행될 수 있다. 도 12에 도시된 바와 같이, 피벗 요청 수신기(1104)는 제 1 파티션(506a)의 데이터셋(508a)이 피벗화된다는 것을 표시하는 피벗 요청 신호(1220)가 제 1 호스트 컴퓨터(102a)로부터 수신될 수 있다. 결과적으로, 피벗 요청 수신기(1104)는 LUN 생성기(1106)에게 데이터셋(508a)을 수용하기 위해 LUN이 생성된다는 것을 표시할 수 있다.

[0054] 단계(1006)에서, 제 1 LUN 내에서 제 1 파티션의 크기가 결정된다. 예를 들면, 일실시예에서 단계(1006)는 LUN 생성기(1106)에 의해 실행될 수 있다. LUN 생성기(1106)는 임의의 방식으로 제 1 파티션(506a)의 크기를 결정할 수 있다. 예를 들어, 도 12를 참조하면 LUN 생성기(1106)는 파티션 테이블(504)을 액세스하여 오프셋 및/또는 제 1 파티션(506a)에 할당된 LBA의 범위를 결정한다. 다른 실시예에서, 제 1 호스트 컴퓨터(102a)는 피벗 요청 수신기(1104)에 의해 수신된 피벗 요청 신호(1220) 내의 제 1 파티션(506a)의 크기에 대한 표시를 포함할 수 있고, 피벗 요청 수신기(1104)는 LUN 생성기(1106)에 대해 크기 표시를 제공할 수 있다.

[0055] 단계(1008)에서, 제 2 파티션의 위치 및 크기를 표시하는 파티션 테이블이 생성된다. 예를 들면, 일실시예에서, 단계(1008)는 LUN 생성기(1106)에 의해 실행될 수 있다. LUN 생성기(1106)는 제 2 파티션이 새로운 LUN 내에 데이터셋(508a)을 저장하도록 구성되었다는 것을 표시하는 파티션 테이블을 생성하도록 구성될 수 있다. 파티션 테이블은 새로운 LUN 내의 제 2 파티션에 대한 위치(예를 들면, LBA 오프셋) 및 크기(예를 들면, 시작 LBA 및 종료 LBA)를 표시하도록 생성된다. 일실시예에서, 파티션 테이블은 다른 파티션에 대한 참조가 아닌 제 2 파티션에 대한 참조를 포함하고, 생성된 파티션 테이블 내에서 제 2 파티션에 대한 표시는 새로운 LUN에 대해 생성된 서명 바로 뒤에 위치한다.

[0056] 단계(1010)에서, 저장 어레이의 제 2 LUN은 제 2 LUN을 위한 서명, 파티션 테이블 및 제 2 파티션을 수용하도록 구성된 크기를 갖도록 생성된다. 예를 들면, 일실시예에서, 단계(1010)는 LUN 생성기(1106)에 의해 실행될 수 있다. LUN 생성기(1106)는 가상 저장부(1204) 내의 제 2 LUN에 포함될 저장 어레이(1202)의 물리적 저장의 양을 할당하도록 구성된다. 예를 들면, 도 12에 도시된 바와 같이 LUN 생성기(1106)는 제 2 LUN(1206)을 생성할 수 있다. 제 2 LUN(1206)은 서명, 단계(1008)에서 생성된 파티션 테이블 및 제 2 파티션을 수용하도록 구성된 크기를 갖도록 생성된다.

[0057] 단계(1012)에서, 서명 및 파티션 테이블은 제 2 LUN에 저장된다. 예를 들면, 일실시예에서 단계(1012)는 LUN 생성기(1106)에 의해 실행될 수 있다. 예를 들면, 도 12에 도시된 바와 같이 LUN 생성기(1106)는 제 2

LUN(1206) 내의 서명(1208) 및 파티션 테이블(1210)을 저장하였다. 서명(1208)은 제 2 LUN(1206)에 대한 식별 정보를 포함하고, 파티션 테이블(1210)은 단계(1008)에서 생성되었다.

[0058] 단계(1014)에서, 제 1 파티션은 제 2 파티션을 위한 파티션 테이블 내에 표시된 위치에서 제 1 LUN로부터 제 2 LUN으로 복사된다. 예를 들면, 일실시예에서 단계(1014)는 파티션 복사기(1108)에 의해 실행될 수 있다. 파티션 복사기(1108)는 소스 LUN으로부터 목적지 LUN으로 파티션을 복사하도록 구성된다. 복사를 개시하기 위해서, LUN 생성기(1106)는 목적지 LUN이 생성되었다는 것을 파티션 복사기(1108)에 표시할 수 있다. 예를 들어, 도 12를 참조하면 파티션 복사기(1108)는 제 1 LUN(500)의 제 1 파티션(506a)(데이터셋(508a)을 포함함)을 제 2 LUN(1206) 내의 제 2 파티션(1212)로 복사한다. 도 12에 도시된 바와 같이, 데이터셋(508a)은 제 2 LUN(1206)의 파티션(1212)으로 복사된다(점선으로 표시되어 있음). 파티션 복사기(1108)는 당업자에게 알려진 기법에 의한 것을 포함하는 임의의 방식으로 파티션 복사를 실행할 수 있다. 예를 들면, 파티션 복사기(1108)는 제 1 파티션(506a)의 스냅샷(예를 들면, 기록시 복사(copy-on-write) 스냅샷 기술을 이용함) 또는 클론(clone)을 실행하여 파티션(1212)을 생성할 수 있다. 예를 들면, 복사 명령어는 "snapshot LUN 500: Partition 506a" 또는 "LUN 500: LBA start-LBA end"의 형태로 발행될 수 있고, 여기에서 "LBA start" 및 "LBA end"는 각각 제 1 LUN(500) 내에서 제 1 파티션(506a)에 대한 시작 및 종료 LBA이다. 이러한 복사 동작은 수 초의 단위 등과 같이 비교적 빠르게 실행될 수 있다. 일실시예에서, LUN(1206)은 서명(1208), 파티션 테이블(1210)(서명(1208) 뒤에 올 수 있음) 및 파티션(1212)(파티션 테이블(1210) 뒤에 올 수 있음)으로 완전히 채워질 수 있지만, 다른 실시예에서, LUN(1206)은 추가적인 저장 공간을 포함할 수 있다.

[0059] 단계(1016)에서, 제 2 LUN은 제 2 호스트 컴퓨터에 대해 매핑된다. 예를 들면, 일실시예에서 단계(1016)는 LUN 할당기(1102)에 의해 실행될 수 있다. 예를 들면, 단계(1002)와 마찬가지로, LUN 할당기(1102)는 노출된 LUN 표시 신호를 (예를 들면, 도 1 및 도 2에 도시된 것과 같은 통신 매체를 통해서) 제 2 호스트 컴퓨터(102b)로 송신할 수 있고, 이것은 제 2 LUN(1206)이 제 2 호스트 컴퓨터(102b)에 할당되었다는 것을 나타낸다. 예를 들면, 노출된 LUN 표시 신호(1222)는 제 2 LUN(1206)에 대한 식별자(예를 들면, 식별 번호, 어드레스 등) 및 제 2 LUN(1206) 내에 포함된 저장의 양을 포함할 수 있다. 일실시예에서, 제 2 호스트 컴퓨터(102b)는 제 2 LUN(1206)을 검출하기 위해서 재스캔 명령어(rescan command)를 발행할 수 있고, 노출된 LUN 표시 신호(1222)에 의해 제 2 LUN(1206)을 주입하도록 요청받을 수 있다.

[0060] 제 2 LUN(1206)을 주입한 후, 제 2 LUN(1206)는 직접 액세스 장치와 동일한 방식으로 제 2 호스트 컴퓨터(102b)에 의해 탑재 가능하고 사용 가능하다. 예를 들면, 일실시예에서 제 2 호스트 컴퓨터(102b)는 도 14에 도시된 흐름도(1400)를 실행할 수 있다. 흐름도(1400)는 제 2 호스트 컴퓨터(102b)에서 가능하게 된 예시적인 프로세스를 제공한다. 흐름도(1400)의 단계(1402)에서, 제 2 호스트 컴퓨터는 제 2 LUN을 탑재할 수 있게 된다. 예를 들면, LUN 할당기(1102)는 제 2 호스트 컴퓨터(102b)가 제 2 LUN(1206)을 탑재할 수 있게 한다. 이와 같이 제 2 호스트 컴퓨터(102b)는 제 2 LUN(1206)의 파티션(1212) 내에 저장된 데이터셋(508a)을 액세스할 수 있다. 예를 들어, 데이터셋(508a)이 가상 머신 하드 드라이브라면, 제 2 호스트 컴퓨터(102b)는 대응하는 가상 머신을 작동하기 위해서 데이터셋(508a)의 가상 머신 파일을 실행할 수 있다. 이와 같이, 흐름도(1400)의 단계(1404)에서, 제 2 호스트 컴퓨터는 가상 머신을 실행하고, 가상 머신에 대한 저장부로서 가상 하드 드라이브를 액세스할 수 있게 된다.

IV. 데이터셋/LUN 파티션에 대해 LUN을 피벗화하는 예시적인 실시예

[0061] 일실시예에서, 데이터셋에 대한 액세스는 제 2 호스트 컴퓨터로부터 제 1 호스트 컴퓨터로(또는 다른 호스트 컴퓨터로) 다시 피벗화될 수 있다. 이러한 데이터셋의 피벗화는 제 2 호스트 컴퓨터가 데이터셋과 더 대화할 필요가 없고, 그에 따라 데이터셋이 (예를 들면, 일시적으로 또는 완전히) 철회(retired)되기 때문이거나, 다른 이유 때문에 실행될 수 있다. 이러한 피벗화는 저장 어레이의 어레이 제어기를 이용하여 피벗화를 실행하는 것을 포함하는 여러 방식으로 실행될 수 있다. 예를 들면, 도 8에 도시된 어레이 제어기(802)는 일실시예에서 피벗화를 실행하도록 구성될 수 있다.

[0062] 데이터셋 피벗 모듈(804)은 어레이 제어기(802)가 제 1 호스트 컴퓨터(또는 다른 호스트 컴퓨터)로 다시 데이터셋에 대한 액세스를 피벗화하게 할 수 있다. 예를 들면, 데이터셋 피벗 모듈(804)은 도 15에 도시된 단계(1502)를 실행하도록 구성될 수 있다. 단계(1502)에서, 데이터셋에 대한 액세스는 데이터셋을 포함하는 LUN을 LUN 파티션에 대해 매핑함으로써 제 2 호스트 컴퓨터로부터 제 1 호스트 컴퓨터로 피벗화된다.

- [0064] 데이터셋 피벗 모듈(804)은 여러 방식으로 단계(1502)를 실행하도록 구성될 수 있고, 단계(1502)는 여러 방식으로 실행될 수 있다. 예를 들면, 도 16은 예시적인 실시예에 따라서 제 2 호스트 컴퓨터로부터 제 1 호스트 컴퓨터(또는 다른 호스트 컴퓨터)로 데이터셋을 피벗화하는 프로세스를 제공하는 흐름도(1600)를 도시한다. 흐름도(1600)는 도 15에 도시된 단계(1502)의 예시이고, 일실시예에서 도 11에 도시된 어레이 제어기(802)에 의해 실행될 수 있다. 당업자라면 흐름도(1600)에 관련된 설명을 기초로 할 때 다른 구조적 및 동작적 실시예가 명확해질 것이다.
- [0065] 설명을 목적으로, 흐름도(1600)는 도 17과 관련하여 다음과 같이 설명되는데, 도 17은 예시적인 실시예에 따라서 도 12에 도시된 컴퓨팅 및 데이터 저장 시스템(1200)에 대한 블록도를 도시한다. 도 17은 제 2 호스트 컴퓨터(102b)로부터 제 1 호스트 컴퓨터(102a)로 피벗화되는 데이터셋(508a)을 도시한다.
- [0066] 도 16에 도시된 바와 같이, 흐름도(1600)는 단계(1602)에서 시작한다. 단계(1602)에서, 저장 어레이의 제 1 LUN은 제 1 호스트 컴퓨터에 대해 노출되고, 저장 어레이의 제 2 LUN은 제 2 호스트 컴퓨터에 대해 노출된다. 예를 들어, 도 17을 참조하면 제 1 LUN(500)은 제 1 호스트 컴퓨터(102a)에 대해 노출될 수 있고, 제 2 LUN(1206)은 제 2 호스트 컴퓨터(102b)에 대해 노출될 수 있다. 일실시예에서, 제 1 LUN(500)은 흐름도(1000)(도 10)의 단계(1002)에 따라서 제 1 호스트 컴퓨터(102a)에 대해 노출될 수 있고, 제 2 LUN(1206)은 상술된 바와 같이 흐름도(1000)의 단계(1016)에 따라서 제 2 호스트 컴퓨터(102b)에 대해 노출될 수 있다.
- [0067] 특정한 시간에, 제 2 호스트 컴퓨터(102b)는 데이터셋(508a)이 제 1 호스트 컴퓨터(102a)로(또는 다른 호스트 컴퓨터로) 다시 피벗화되는지 결정한다. 이와 같이, 제 2 호스트 컴퓨터(102b)는 도 18에 도시된 흐름도(1800)를 실행할 수 있다. 흐름도(1800)는 다음과 같이 설명된다. 흐름도(1800)의 단계(1802)에서, 제 2 호스트에 대해 노출된 LUN(logical unit number) 내에 데이터셋을 저장하는 파티션에 대한 액세스가 차단된다. 예를 들면, 제 2 호스트 컴퓨터(102b)는 데이터셋(508a)을 저장하는 제 2 LUN(1206)의 파티션(1212)에 대한 액세스를 차단하여, 데이터셋(508a)이 피벗화될 준비를 할 수 있다. 일실시예에서, 제 2 호스트 컴퓨터(102b)는 데이터셋(508a)에 대한 임의의 개방된 핸들을 차단할 수 있고, 데이터셋(508a)이 더 이상 액세스되지 않으면, 제 2 호스트 컴퓨터(102b)의 논리적 디스크 관리자는 드라이브 문자(drive letter)를 제거하고 파티션(1212)으로 매핑할 수 있다. 이것은 제 2 호스트 컴퓨터(102b)가 파티션(1212)에 대한 판독 및/또는 기록을 중지하도록 보장한다. 단계(1804)에서, LUN을 철회하려는 요청이 저장 어레이로 송신된다. 예를 들면, 도 18에 도시된 바와 같이 제 2 호스트 컴퓨터(102b)는 (예를 들면, 도 1 및 도 2에 도시된 바와 같은 통신 매체를 통해서) 저장 어레이(1202)로 철회 또는 피벗 요청 신호(1702)를 송신하여 데이터셋(508a)이 다른 호스트 컴퓨터로 철회 및/또는 피벗화되도록 요청할 수 있다.
- [0068] 다시 흐름도(1600)(도 16)를 참조하면, 단계(1604)에서 제 2 LUN의 철회에 대한 요청이 제 2 호스트 컴퓨터로부터 수신된다. 예를 들면, 일실시예에서 단계(1604)는 피벗 요청 수신기(1104)에 의해 실행될 수 있다. 도 17에 도시된 바와 같이, 피벗 요청 수신기(1104)는 철회 요청 신호(1702)를 제 2 호스트 컴퓨터(102b)로부터 수신할 수 있다. 철회 요청 신호(1702)는 파티션(1212) 내에 저장된 데이터셋(508a)이 저장 라이브러리를 작동하는 멀티-파티션 LUN에 대해 다시 피벗화된다는 것을 표시할 수 있다.
- [0069] 단계(1606)에서, 제 2 LUN 내에서 파티션의 크기가 결정된다. 예를 들면, 일실시예에서 단계(1606)는 LUN 생성기(1106)에 의해 실행될 수 있다. LUN 생성기(1106)는 임의의 방식으로 파티션(예를 들면, 파티션(1212))의 크기를 결정할 수 있다. 예를 들어, 도 17을 참조하면 LUN 생성기(1106)는 LUN(1206)의 파티션 테이블(1210)을 액세스하여 오프셋 및/또는 파티션(1212)에 할당된 LBA의 범위를 결정한다. 다른 실시예에서, 제 2 호스트 컴퓨터(102b)는 피벗 요청 수신기(1104)에 의해 수신된 철회 요청 신호(1702) 내의 파티션(1212)의 크기에 대한 표시를 포함할 수 있고, 피벗 요청 수신기(1104)는 LUN 생성기(1106)에 대해 크기 표시를 제공할 수 있다.
- [0070] 단계(1608)에서, 데이터셋의 복사본이 제 1 LUN 내에 저장된다. 예를 들면, 일실시예에서 단계(1608)는 파티션 복사기(1108)에 의해 실행될 수 있다. 상술된 바와 같이, 파티션 복사기(1108)는 소스 LUN으로부터 목적지 LUN으로 파티션을 복사하도록 구성될 수 있다. 예를 들어, 도 17을 참조하면, 파티션(1212)이 파티션(506a)의 크기를 초과하지 않는(예를 들면 같거나 작은) 크기(단계(1606)에서 결정됨)를 갖는 경우에 파티션 복사기(1108)는 제 2 LUN(1206)의 파티션(1212)(데이터셋(508a)을 포함함)을 제 1 LUN(500)의 제 1 파티션(506a)으로 복사하거나, 제 1 LUN(500)의 다른 파티션으로 파티션(1212)을 복사할 수 있다. 일실시예에서, 제 2 LUN(1206) 내의 파티션(1212)의 크기가 제 1 LUN(500) 내의 제 1 파티션(506a)의 크기를 초과한다면, 제 1 LUN(500)에 저장 어레이(1202)의 저장 영역을 첨부함으로써 제 1 LUN(500)의 크기는 어레이 제어기(802)만큼 증가(예를 들면, 제 1 LUN(500)은 그 크기의 성장이 초래될 수 있음)될 수 있다. 이러한 경우에, 데이터셋(508a)의 복사본은 제 1

LUN(500)에 첨부된 저장 영역 내에 저장될 수 있다. 제 1 호스트 컴퓨터(102a)는 제 1 LUN(500)을 탑재 해제 및 재탑재하여 첨부된 저장 영역에 대한 액세스가 가능해지게 될 수 있다.

[0071] 도 17의 예시에서, 데이터셋(508a)은 제 1 LUN(500)의 제 1 파티션(506a)으로 복사된다(점선으로 표시됨). 파티션 복사기(1108)는 흐름도(1000)(도 10)의 단계(1014)와 관련하여 상술된 기법을 포함하고, 당업자에게 알려져 있을 수 있는 임의의 방식으로 파티션 복사를 실행할 수 있다. 이러한 복사 동작은 초 단위로 실행되는 것을 포함하여 비교적 빠르게 실행될 수 있다.

[0072] 단계(1610)에서, 제 1 LUN의 파티션 테이블은 데이터셋의 복사본을 저장하는 제 1 LUN의 저장 영역에 대응하는 파티션을 포함하도록 업데이트된다. 예를 들면, 일실시예에서 단계(1610)는 LUN 생성기(1106)에 의해 실행될 수 있다. 도 17의 예시에서, LUN 생성기(1106)는 제 1 파티션(506a) 또는 제 1 LUN(500)의 다른 파티션으로 데이터셋(508a)을 복사하는 것에 따라서 제 1 LUN(500)의 파티션 테이블(504)을 업데이트하도록 구성될 수 있다. 제 1 LUN(500)이 데이터셋(508a)을 저장하기 위해 새로운 파티션을 생성하도록 성장되었다면, 새로운 파티션의 위치 및/또는 크기는 LUN 생성기(1106)에 의해 파티션 테이블(504) 내에 표시된다. 제 1 호스트 컴퓨터(102a)는 제 1 LUN(500)를 탑재 해제 및 재탑재하여, 필요한 경우에 제 1 호스트 컴퓨터(102a)의 메모리/캐시 내에 파티션 테이블을 업데이트할 수 있다.

[0073] 본 명세서에 설명된 실시예는 가상 머신 파일 등과 같은 데이터셋이 멀티-파티션 LUN 내에 저장되게 하고, 호스트 컴퓨터에 노출될 수 있는 하나 이상의 파일로서 전용 LUN 내에 실시간으로 저장되게 한다. 멀티-파티션 LUN은 가상 하드 드라이브의 통합형 라이브러리로서 동작할 수 있다. 데이터셋이 배치되는 것이 바람직할 때, 데이터셋은 LUN으로 변환되고 목적지 호스트로 할당된다. 바람직한 데이터셋을 포함하는 멀티-파티션 LUN의 파티션은 전용 LUN으로 (예를 들면, 테라바이트 당 초 단위로) 비교적 빠르게 복제될 수 있다(예를 들면, 기록 스탯샷 기술에서의 하드웨어 복사를 이용함). 다음에 전용 LUN으로 복사된 데이터셋은 새로운 서버에 대해 노출되고, 직접 부착된 LUN으로서 액세스될 수 있다. 이 기술은 가상 머신의 신속한 배치를 허용하기 위해 사용, 더 큰 데이터베이스에서의 데이터 마이닝(data mining)을 가능하게 하는 데 사용 및/또는 저장 어레이가 그의 통상적인 LUN 맵핑 제한의 10배 또는 100배로 확대할 수 있도록 허용하는데 사용되어, 저장 어레이가 장기간 저장소로서 사용될 수 있게 한다. 이러한 저장 어레이에는 오토로더 명령어(autoloader commands)가 저장 어레이 내에서 재매핑(remapped)되어 파티션이 분리된 LUN으로 다시 재반영될 수 있게 하기 때문에, 가상 테이프 라이브러리를 시뮬레이션하는 것이 가능하게 될 것이다.

V. 다른 예시적인 실시예

[0074] 어레이 제어기(112)(도 1, 도 2), LUN 파티셔너(604)(도 6), 어레이 제어기(802)(도 8, 도 11, 도 12, 도 17), 데이터셋 피벗 모듈(804)(도 8, 도 11, 도 12, 도 17), LUN 할당기(1102)(도 11, 도 12, 도 17), 피벗 요청 수신기(1104)(도 11, 도 12, 도 17), LUN 생성기(1106)(도 11, 도 12, 도 17) 및 파티션 복사기(1108)(도 11, 도 12, 도 17)는 하드웨어, 소프트웨어, 펌웨어 또는 그 임의의 조합으로 구현될 수 있다. 예를 들면, 어레이 제어기(112), LUN 파티셔너(604), 어레이 제어기(802), 데이터셋 피벗 모듈(804), LUN 할당기(1102), 피벗 요청 수신기(1104), LUN 생성기(1106) 및/또는 파티션 복사기(1108)는 하나 이상의 프로세서 내에 실행되도록 구성된 컴퓨터 프로그램 코드로서 구현될 수 있다. 이와 다르게, 어레이 제어기(112), LUN 파티셔너(604), 어레이 제어기(802), 데이터셋 피벗 모듈(804), LUN 할당기(1102), 피벗 요청 수신기(1104), LUN 생성기(1106) 및/또는 파티션 복사기(1108)는 하드웨어 로직/전기 회로로 구현될 수 있다.

[0075] [0076] 도 19는 본 발명의 실시예가 구현될 수 있는 컴퓨터(1900)의 예시적인 구현을 도시한다. 예를 들면, 컴퓨터(102a-102c)(도 1, 도 12, 도 17) 및/또는 컴퓨터(602)(도 6)는 컴퓨터(1900)와 마찬가지로 구현될 수 있고, 컴퓨터(1900)의 하나 이상의 특정부 및/또는 다른 특정부를 포함할 수 있다. 컴퓨터(1900)는 예를 들면, 통상적인 퍼스널 컴퓨터, 모바일 컴퓨터 또는 워크스테이션(workstation)의 형태를 갖는 범용 컴퓨팅 장치일 수 있거나, 컴퓨터(1900)는 전용 컴퓨팅 장치일 수 있다. 본 명세서에 제공된 컴퓨터(1900)에 관한 설명은 설명을 목적으로 제공된 것이고, 제한하려는 의도가 아니다. 본 발명의 실시예는 당업자에게 공지된 바와 같이 다른 종류의 컴퓨터 시스템에서 구현될 수 있다.

[0077] 도 19에 도시된 바와 같이, 컴퓨터(1900)는 처리 장치(1902), 시스템 메모리(1904) 및 시스템 메모리(1904)를 처리 장치(1902)로 결합하는 것을 포함하여 여러 시스템 구성 요소를 결합하는 버스(1906)를 포함한다. 버스(1906)는 메모리 버스 또는 메모리 제어기, 주변 버스, 가속 그래픽 포트(accelerated graphics port) 및 프로

세서 또는 다양한 버스 아키텍처 중 임의의 것을 사용하는 지역 버스를 포함하는 몇몇 종류의 버스 구조 중 임의의 하나 이상의 것을 표시한다. 시스템 메모리(1904)는 ROM(read only memory)(1908) 및 RAM(random access memory)(1910)을 포함한다. 기본 입/출력 시스템(1912)(BIOS)은 ROM(1908) 내에 저장된다.

[0078] 컴퓨터(1900)는 또한 하드 디스크를 판독 및 기록하는 하드 디스크 드라이브(1914), 탈착 가능 자기 디스크(1918)를 판독 또는 기록하는 자기 디스크 드라이브(1916) 및 CD ROM, DVD ROM 또는 다른 광학 매체 등과 같은 탈착 가능 광학 디스크(1922)를 판독 또는 기록하는 광학 디스크 드라이브(1920) 중 하나 이상을 포함한다. 하드 디스크 드라이브(1914), 자기 디스크 드라이브(1916) 및 광학 디스크 드라이브(1920)는 각각 하드 디스크 드라이브 인터페이스(1924), 자기 디스크 드라이브 인터페이스(1926) 및 광학 드라이브 인터페이스(1928)에 의해 버스(1906)에 접속된다. 드라이브 및 그 연관된 컴퓨터 판독 가능 매체는 컴퓨터-판독 가능 명령어, 데이터 구조, 프로그램 모듈 및 컴퓨터를 위한 다른 데이터의 비휘발성 저장부를 제공한다. 하드 디스크, 탈착 가능 자기 디스크 및 탈착 가능 광학 디스크가 설명되어 있으나, 플래시 메모리 카드, 디지털 비디오 디스크, RAM(random access memories), ROM(read only memories) 등과 같은 다른 종류의 컴퓨터 판독 가능 매체를 사용하여 데이터를 저장할 수 있다.

[0079] 다수의 프로그램 모듈은 하드 디스크, 자기 디스크, 광학 디스크, ROM 또는 RAM에 저장될 수 있다. 이러한 프로그램은 운영 시스템(1930), 하나 이상의 애플리케이션 프로그램(1932), 다른 프로그램 모듈(1934) 및 프로그램 데이터(1936)를 포함한다. 애플리케이션 프로그램(1932) 또는 프로그램 모듈(1934)은 예를 들면, 어레이 제어기(112), LUN 파티셔너(604), 어레이 제어기(802), 데이터셋 피벗 모듈(804), LUN 할당기(1102), 피벗 요청 수신기(1104), LUN 생성기(1106) 및/또는 파티션 복사기(1108), 흐름도(700), 단계(902), 흐름도(1000), 흐름도(1300), 흐름도(1400), 단계(1502), 흐름도(1600), 흐름도(1800), (흐름도(700, 1000, 1300, 1400, 1600, 1800)의 임의의 단계를 포함함), 및/또는 상술된 임의의 다른 실시예를 구현/가능화하는 컴퓨터 프로그램 로직을 포함할 수 있다.

[0080] 사용자는 키보드(1938) 및 포인팅 장치(1940) 등과 같은 입력 장치를 통해 컴퓨터(1900)에 명령어 및 정보를 입력할 수 있다. 다른 입력 장치(도시하지 않음)는 마이크로폰, 조이스틱, 게임패드, 위성 접시, 스캐너 등을 포함할 수 있다. 이를 및 다른 입력 장치는 때때로 버스(1906)에 결합된 직렬 포트 인터페이스(1942)를 통해 처리 장치(1902)에 접속되지만, 병렬 포트, 게임 포트 또는 USB(universal serial bus) 등과 같은 다른 인터페이스에 의해 접속될 수도 있다.

[0081] 모니터(1944) 또는 다른 종류의 디스플레이 장치는 또한 비디오 어댑터(1946) 등과 같은 인터페이스를 통해 버스(1906)에 접속된다. 모니터에 추가하여, 컴퓨터(1900)는 스피커 및 프린터 등과 같은 다른 주변 출력 장치(도시하지 않음)를 포함할 수 있다.

[0082] 컴퓨터(1900)는 네트워크 어댑터 또는 인터페이스(1950), 모뎀(1952) 또는 네트워크를 통해 통신을 형성하는 다른 수단을 통해 네트워크(1948)(예를 들면, 인터넷)에 접속된다. 내부형이거나 외부형일 수 있는 모뎀(1952)은 직렬 포트 인터페이스(1942)를 통해 버스(1906)에 접속된다.

[0083] 본 명세서에 사용된 바와 같이, "컴퓨터 프로그램 매체" 및 "컴퓨터 판독 가능 매체"라는 용어는 일반적으로 플래시 메모리 카드, 디지털 비디오 디스크, RAM(random access memories), ROM(read only memories) 등과 같은 다른 매체뿐만 아니라 하드 디스크 드라이브(1914)와 연관된 하드 디스크, 탈착 가능 자기 디스크(1918), 탈착 가능 광학 디스크(1922) 등과 같은 매체를 지칭하는 데 사용된다.

[0084] 상술된 바와 같이, 컴퓨터 프로그램 및 모듈(애플리케이션 프로그램(1932) 및 다른 프로그램 모듈(1934)을 포함함)은 하드 디스크, 자기 디스크, 광학 디스크, ROM 또는 RAM에 저장될 수 있다. 이러한 컴퓨터 프로그램은 또한 네트워크 인터페이스(1950) 또는 직렬 포트 인터페이스(1942)를 통해 수신될 수 있다. 이러한 컴퓨터 프로그램은, 애플리케이션에 의해 실행 또는 로딩될 때 컴퓨터(1900)가 본 명세서에 설명된 본 발명의 실시예의 특징을 구현하게 한다. 따라서 이러한 컴퓨터 프로그램은 컴퓨터(1900)의 제어기를 나타낸다.

[0085] 본 발명은 또한 임의의 컴퓨터 사용 가능 매체에 저장된 소프트웨어를 포함하는 컴퓨터 프로그램 제품에 관한 것이다. 이러한 소프트웨어는, 하나 이상의 데이터 처리 장치에서 실행될 때 데이터 처리 장치가 상술된 바와 같이 동작하도록 유도한다. 본 발명의 실시예는 현재 또는 미래에 공지되는 임의의 컴퓨터 사용 가능 또는 컴퓨터 판독 가능 매체를 이용한다. 컴퓨터 판독 가능 매체의 예시는 RAM, 하드 드라이브, 플로피 디스크, CD ROM, DVD ROM, 접 디스크, 테이프, 자기 저장 장치, 광학 저장 장치, MEM, 나노 기술 기반의 저장 장치 등과 같은 저장 장치를 포함하지만 이것으로 한정되지 않는다.

[0086]

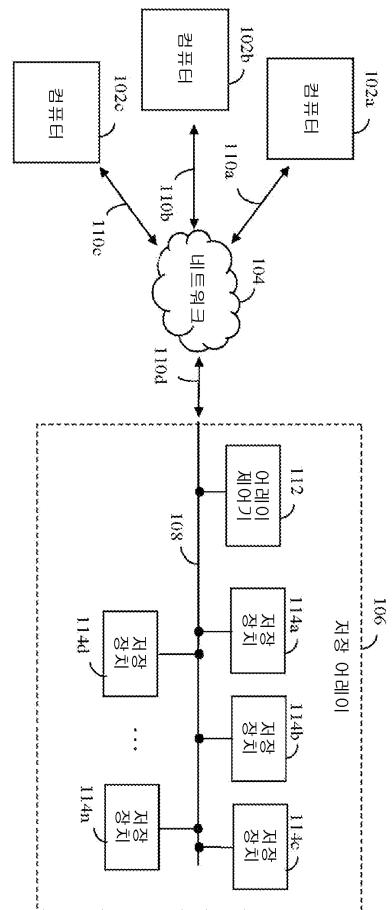
VI. 결론

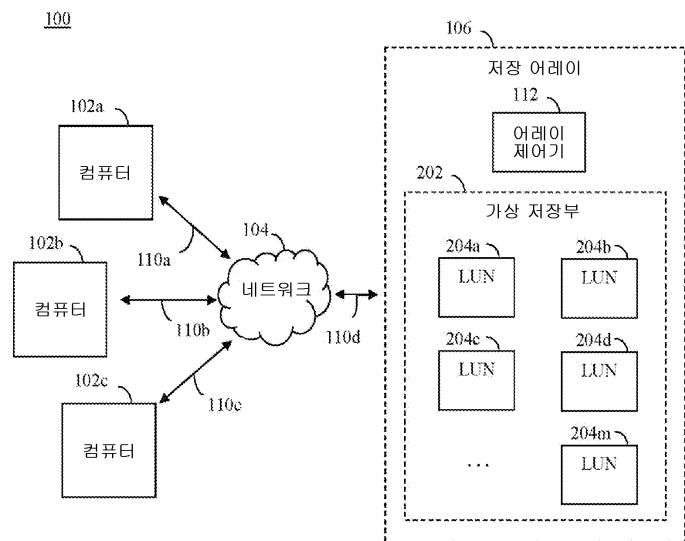
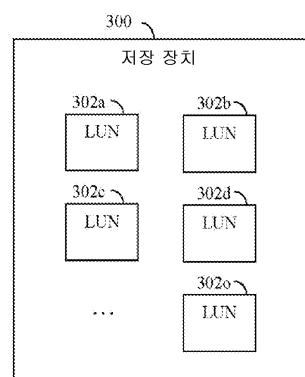
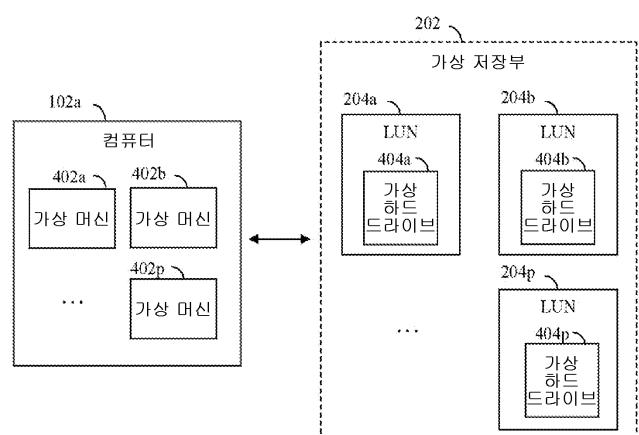
[0087]

본 발명의 여러 실시예가 상술되어 있으나, 이것은 오로지 예시를 목적으로 제시된 것이고, 제한하기 위한 것이 아님을 이해할 것이다. 당업자라면 첨부된 청구항에서 정의된 본 발명의 정신 및 범주를 벗어나지 않으면서 그에 대한 형태 및 세부 사항에서의 다양한 변형이 이루어질 수 있다는 것을 이해할 것이다. 따라서 본 발명의 범위 및 범주는 상술된 예시적인 실시예 중 어느 것으로도 한정되지 않지만, 오로지 이하의 청구항 및 그 등가물에 따라서만 정의되어야 한다.

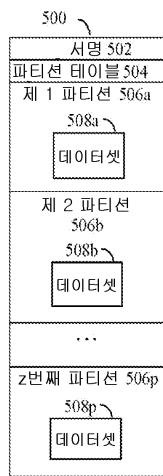
도면

도면1



도면2**도면3****도면4**

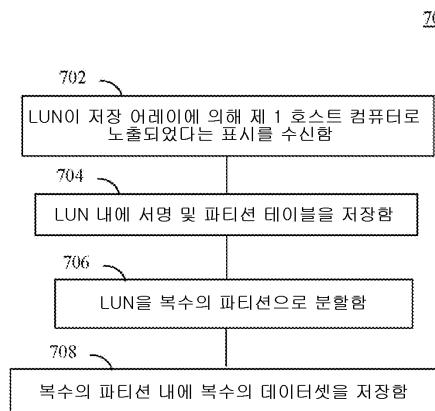
도면5



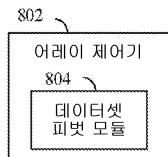
도면6



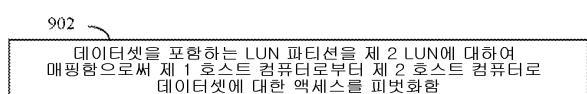
도면7

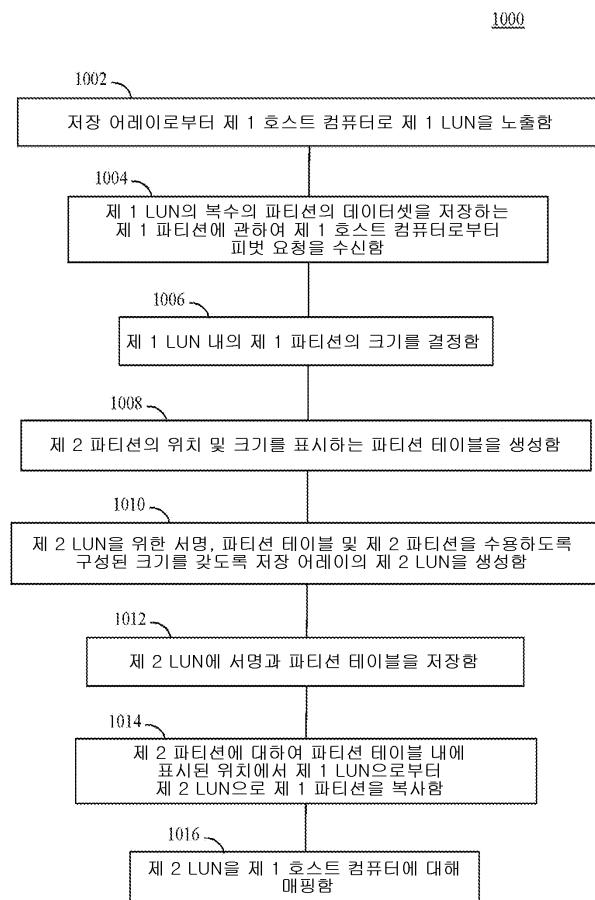
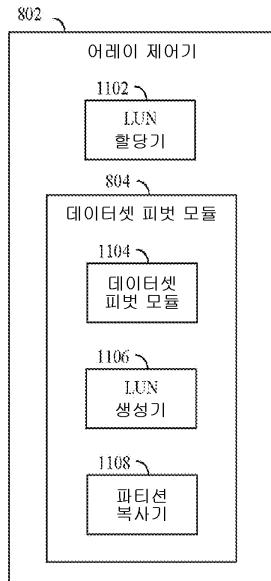


도면8

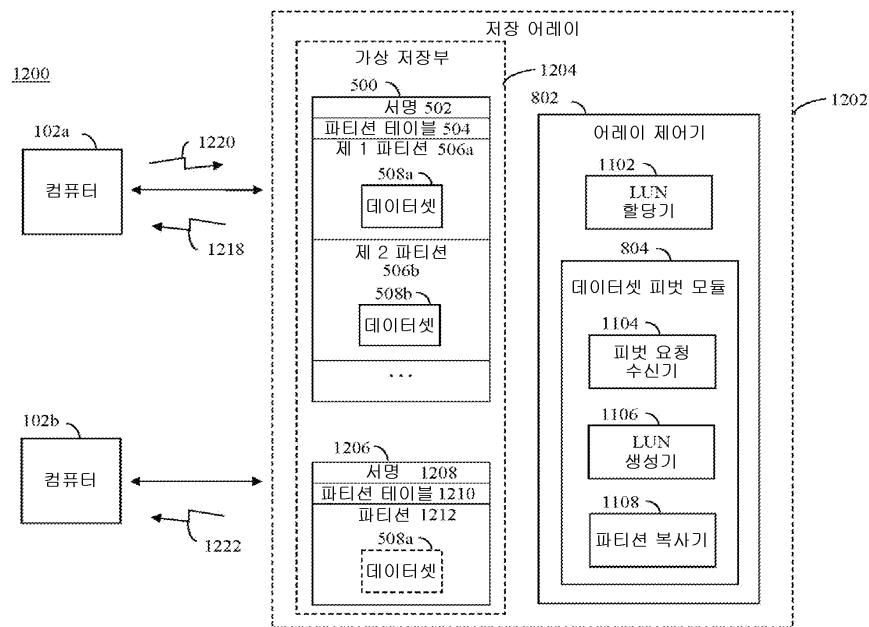


도면9

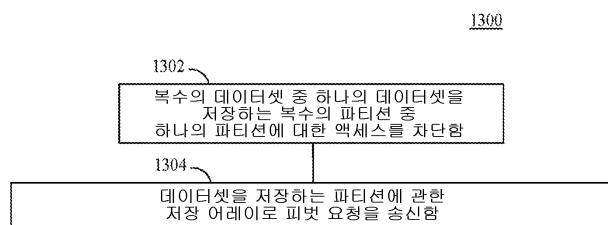


도면10**도면11**

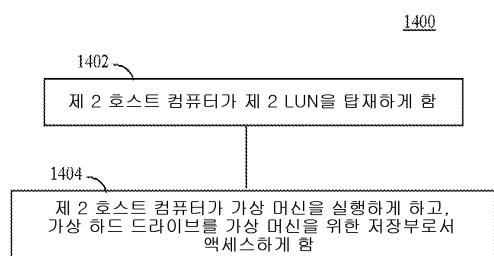
도면12



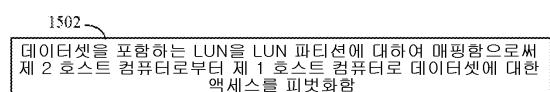
도면13



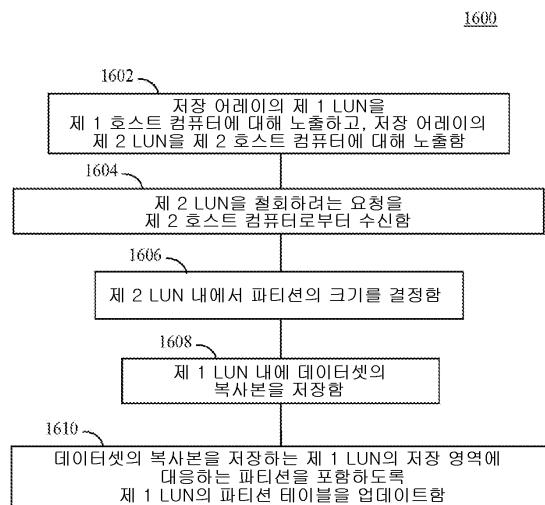
도면14



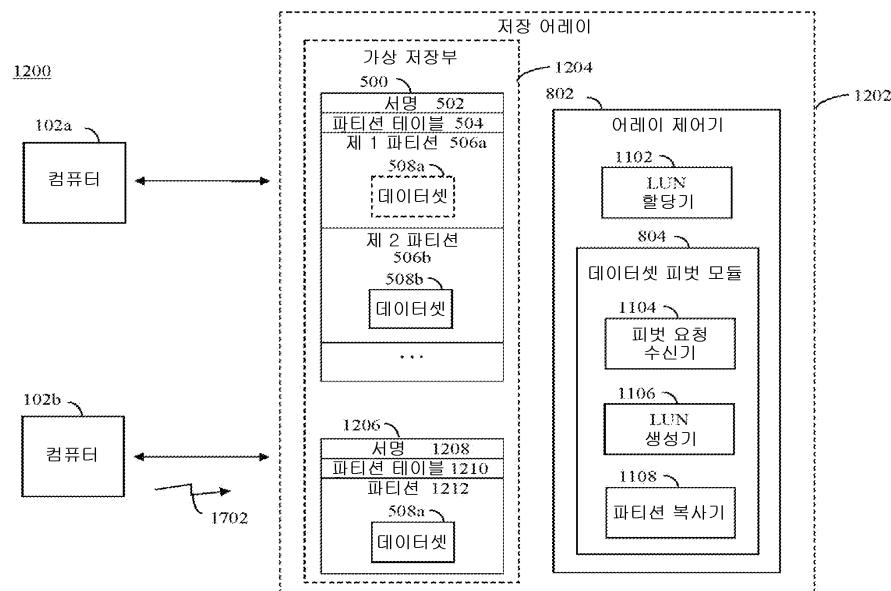
도면15



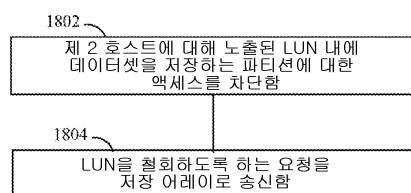
도면16



도면17



도면18



도면19

