



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2019년03월19일
(11) 등록번호 10-1960239
(24) 등록일자 2019년03월13일

(51) 국제특허분류(Int. Cl.)
G06F 9/455 (2018.01)
(52) CPC특허분류
G06F 9/45504 (2013.01)
G06F 9/45558 (2013.01)
(21) 출원번호 10-2018-7027979(분할)
(22) 출원일자(국제) 2012년05월29일
심사청구일자 2018년09월28일
(85) 번역문제출일자 2018년09월28일
(65) 공개번호 10-2018-0112078
(43) 공개일자 2018년10월11일
(62) 원출원 특허 10-2013-7031785
원출원일자(국제) 2012년05월29일
심사청구일자 2017년04월25일
(86) 국제출원번호 PCT/US2012/039878
(87) 국제공개번호 WO 2012/166741
국제공개일자 2012년12월06일
(30) 우선권주장
13/150,996 2011년06월01일 미국(US)
(56) 선행기술조사문헌
KR1020090026342 A
(뒷면에 계속)

(73) 특허권자
마이크로소프트 테크놀로지 라이선싱, 엘엘씨
미국 워싱턴주 (우편번호 : 98052) 레드몬드 원
마이크로소프트 웨이
(72) 발명자
교빈다라주 나가
미국 워싱턴주 98052-6399 레드몬드 원 마이크로
소프트 웨이 엘씨에이 - 인터내셔널 페이턴츠 마
이크로소프트 코포레이션
러시노비치 마크
미국 워싱턴주 98052-6399 레드몬드 원 마이크로
소프트 웨이 엘씨에이 - 인터내셔널 페이턴츠 마
이크로소프트 코포레이션
(74) 대리인
제일특허법인(유)

전체 청구항 수 : 총 19 항

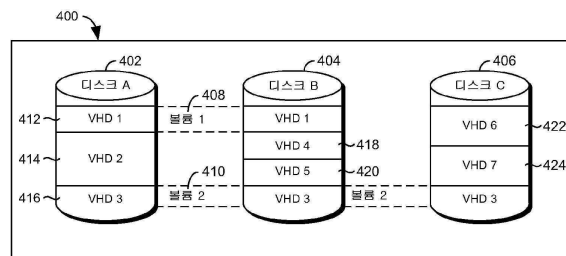
심사관 : 유진태

(54) 발명의 명칭 멀티디스크 호스트에서의 가상 머신 I/O의 분리

(57) 요약

본 발명의 실시예는 둘 이상의 컴퓨터 저장 매체에 걸쳐서 스페닝된 가상 하드 드라이브를 유지함과 동시에 컴퓨터 저장 매체 중 하나에서 스페닝되지 않은 가상 하드 드라이브를 유지하기 위한 시스템, 방법 및 컴퓨터 저장 매체에 관한 것이다. 이러한 방법은 볼륨 스페닝을 활용하는 컴퓨터 저장 매체에 걸쳐서 스페닝된 가상 하드 드라이브의 데이터를 저장하는 단계를 포함한다. 스페닝된 가상 하드 드라이브가 컴퓨터 저장 매체에 유지되는 동안, 방법은 컴퓨터 저장 매체 중 하나에서 스페닝되지 않은 가상 하드 드라이브의 데이터를 저장하는 단계를 포함한다.

대표도 - 도4



(52) CPC특허분류

G06F 2009/45579 (2013.01)

(56) 선행기술조사문헌

US7606868 B1

JP2009104321 A

W02010022099 A2

US6148369 A

KR1020090025204 A

명세서

청구범위

청구항 1

프로세서 및 메모리를 이용하는 컴퓨팅 환경에서 1) 둘 이상의 컴퓨터 저장 매체(computer storage media("CSM"))에 걸쳐서 스페닝된(spanned) 가상 하드 드라이브를 유지함과 동시에 2) 상기 둘 이상의 CSM 중 제 1 CSM에 스페닝되지 않은(non-spanned) 가상 하드 드라이브를 유지하는 방법으로서,

볼륨 스페닝을 이용하여 상기 둘 이상의 CSM에 걸쳐서 상기 스페닝된 가상 하드 드라이브의 데이터를 저장하는 단계 - 상기 둘 이상의 CSM은 호스트 파일 시스템에 기초하여 동작하는 노드 상의 물리적 하드 드라이브이고, 상기 둘 이상의 CSM은 호스트 디스크 드라이버에 기초하여 동작하고, 상기 스페닝된 가상 하드 드라이브는 가상 하드 드라이브 내의 데이터 액세스하는 하나 이상의 가상 머신과 연관됨 - 와,

상기 스페닝된 가상 하드 드라이브가 상기 둘 이상의 CSM에 유지되는 동안, 상기 제 1 CSM에 상기 스페닝되지 않은 가상 하드 드라이브의 데이터를 저장하는 단계를 포함하되,

상기 스페닝되지 않은 가상 하드 드라이브는 상기 둘 이상의 CSM을 스페닝하지 않고,

상기 스페닝되지 않은 가상 하드 드라이브는 상기 하나 이상의 가상 머신과 연관되고,

상기 하나 이상의 가상 머신은 가상 하드 드라이브(virtual hard drive ("VHD")) 드라이버에 기초하여 동작하고,

상기 스페닝된 가상 하드 드라이브 및 상기 스페닝되지 않은 가상 하드 드라이브의 데이터에 관한 상기 하나 이상의 가상 머신으로부터의 입력/출력(Input/Output("I/O")) 요청은 상기 호스트 파일 시스템을 통한 상기 VHD 드라이버 및 상기 호스트 디스크 드라이버 사이의 통신에 기초하여 처리되는

방법.

청구항 2

제 1 항에 있어서,

상기 둘 이상의 CSM 중 어떤 CSM이 착신 I/O 요청과 관련된 데이터를 유지하는지를 식별하기 위해 하이퍼바이저 계층에서 상기 착신 I/O 요청을 검사하는 단계를 더 포함하는

방법.

청구항 3

제 1 항에 있어서,

상기 둘 이상의 CSM 중 어떤 CSM이 착신 I/O 요청과 관련된 데이터를 유지하는지를 식별하기 위해 상기 호스트 파일 시스템에서 상기 착신 I/O 요청을 검사하는 단계를 더 포함하는

방법.

청구항 4

제 1 항에 있어서,

상기 둘 이상의 CSM에 걸쳐 상기 스페닝되지 않은 가상 하드 드라이브를 스페닝("스페닝되지 않은 가상 하드 드라이브의 스페닝")하는 단계를 더 포함하는

방법.

청구항 5

제 4 항에 있어서,

상기 스페닝된 가상 하드 드라이브는 상기 스페닝되지 않은 가상 하드 드라이브의 스페닝보다 적어도 하나의 추가 CSM에 걸쳐 스페닝하는

방법.

청구항 6

제 1 항에 있어서,

상기 둘 이상의 CSM은 공통 노드 내에 통합되는

방법.

청구항 7

제 1 항에 있어서,

상기 스페닝되지 않은 가상 하드 드라이브는 별개의 CSM 상의 별개의 볼륨으로부터 액세스가능한

방법.

청구항 8

제 1 항에 있어서,

볼륨 스페닝은 단일 볼륨의 기능을 제공하면서 데이터가 둘 이상의 별개의 CSM에 저장되는 컨테이너인

방법.

청구항 9

프로세서 및 메모리를 갖는 컴퓨팅 시스템에 의해 실행될 때 컴퓨팅 시스템이 1) 둘 이상의 컴퓨터 저장 매체 (computer-storage medium("CSM"))에 걸쳐서 스페닝된 가상 하드 드라이브를 유지함과 동시에 2) 상기 둘 이상의 CSM 중 제 1 CSM에 스페닝되지 않은 가상 하드 드라이브를 유지하기 위한 방법을 수행하도록 하는 컴퓨터 실행 가능한 명령어를 포함하는 하나 이상의 컴퓨터 저장 매체로서,

상기 방법은,

상기 스페닝된 가상 하드 드라이브의 데이터에 대한 제 1 요청을 수신하는 단계와,

상기 제 1 요청을 상기 둘 이상의 CSM을 스페닝하는 상기 스페닝된 가상 하드 드라이브로 전송하는 단계 - 상기 둘 이상의 CSM은 호스트 파일 시스템에 기초하여 동작하는 노드 상의 물리적 하드 드라이브이고, 상기 둘 이상의 CSM은 호스트 디스크 드라이버에 기초하여 동작하고, 상기 스페닝된 가상 하드 드라이브는 가상 하드 드라이브 내의 데이터 액세스하는 하나 이상의 가상 머신과 연관됨 - 와,

상기 스페닝되지 않은 가상 하드 드라이브의 데이터에 대한 제 2 요청을 수신하는 단계 - 상기 스페닝되지 않은 가상 하드 드라이브는 상기 하나 이상의 가상 머신과 연관됨 - 와,

상기 제 2 요청을 상기 제 1 CSM으로 전송하는 단계 - 상기 스페닝되지 않은 가상 하드 드라이브 및 상기 스페

닝된 가상 하드 드라이브는 둘 다 적어도 부분적으로 상기 제 1 CSM에서 실질적으로 동시에 액세스할 수 있음 - 와,

상기 하나 이상의 가상 머신은 가상 하드 드라이브(virtual hard drive ("VHD")) 드라이버에 기초하여 동작하고,

상기 스페닝된 가상 하드 드라이브 및 상기 스페닝되지 않은 가상 하드 드라이브의 데이터에 관한 상기 하나 이상의 가상 머신으로부터의 입력/출력(Input/Output("I/O")) 요청은 상기 호스트 파일 시스템을 통한 상기 VHD 드라이버 및 상기 호스트 디스크 드라이버 사이의 통신에 기초하여 처리되는

컴퓨터 저장 매체.

청구항 10

제 9 항에 있어서,

상기 제 1 요청은 호스트 상에서 동작하는 가상 머신으로부터의 I/O 요청으로서, 상기 호스트가 상기 둘 이상의 CSM을 제어하는

컴퓨터 저장 매체.

청구항 11

제 9 항에 있어서,

상기 제 1 요청을 전송하는 단계는 호스트의 하이퍼바이저 계층에서 수행되는

컴퓨터 저장 매체.

청구항 12

제 9 항에 있어서,

상기 제 1 요청을 전송하는 단계는 호스트의 호스트 파일 시스템 계층에서 수행되는

컴퓨터 저장 매체.

청구항 13

제 9 항에 있어서,

상기 방법은, 스페닝된 볼륨을 상기 제 1 CSM에 동적으로 추가하여, 상기 제 1 CSM이 스페닝된 볼륨 및 스페닝되지 않은 볼륨 둘 다를 동시에 유지하는 단계를 더 포함하는

컴퓨터 저장 매체.

청구항 14

제 9 항에 있어서,

상기 방법은, 볼륨이 스페닝할 상기 둘 이상의 CSM 중 하나의 CSM을 선택하는 단계 - 상기 하나의 CSM은 상기 둘 이상의 CSM의 각각을 이용하는 다수의 가상 머신에 기초하여 선택됨 - 를 더 포함하는

컴퓨터 저장 매체.

청구항 15

제 9 항에 있어서,

상기 방법은, 볼륨이 스페닝할 상기 둘 이상의 CSM 중 하나의 CSM을 선택하는 단계 - 상기 둘 이상의 CSM은 스페닝된 볼륨을 배타적으로 유지하도록 재구성되지 않음 - 를 더 포함하는

컴퓨터 저장 매체.

청구항 16

1) 둘 이상의 컴퓨터 저장 매체(computer-storage medium("CSM"))에 걸쳐서 스페닝된 가상 하드 드라이브를 유지함과 동시에 2) 상기 둘 이상의 CSM 중 제 1 CSM에 스페닝되지 않은 가상 하드 드라이브를 유지하는 시스템으로서,

호스트를 포함하되, 상기 호스트는

(1) 가상 머신과,

(2) 가상 하드 드라이브 드라이버와,

(3) 호스트 파일 시스템과,

(4) 스페닝된 가상 하드 드라이브의 적어도 부분 및 스페닝되지 않은 가상 하드 드라이브의 전체 둘 다를 유지하는 제 1 CSM - 상기 둘 이상의 CSM은 호스트 파일 시스템에 기초하여 동작하는 노드 상의 물리적 하드 드라이브이고, 상기 둘 이상의 CSM은 호스트 디스크 드라이버에 기초하여 동작하고, 상기 스페닝된 가상 하드 드라이브는 가상 하드 드라이브 내의 데이터 액세스하는 하나 이상의 가상 머신과 연관됨 - 과,

(5) 상기 스페닝된 가상 하드 드라이브의 적어도 부분을 유지하며 상기 스페닝되지 않은 가상 하드 드라이브를 유지하지 않는 제 2 CSM을 포함하되,

상기 스페닝되지 않은 가상 하드 드라이브는 상기 하나 이상의 가상 머신과 연관되고,

상기 하나 이상의 가상 머신은 가상 하드 드라이브(virtual hard drive ("VHD")) 드라이버에 기초하여 동작하고,

상기 스페닝된 가상 하드 드라이브 및 상기 스페닝되지 않은 가상 하드 드라이브의 데이터에 관한 상기 하나 이상의 가상 머신으로부터의 입력/출력(Input/Output("I/O")) 요청은 상기 호스트 파일 시스템을 통한 상기 VHD 드라이버 및 상기 호스트 디스크 드라이버 사이의 통신에 기초하여 처리되는

시스템.

청구항 17

제 16 항에 있어서,

상기 VHD 드라이버는 I/O 요청을 상기 제 1 CSM 및 상기 제 2 CSM에 재전송하는

시스템.

청구항 18

제 16 항에 있어서,

상기 호스트 파일 시스템은 I/O 요청을 상기 제 1 CSM 및 상기 제 2 CSM에 재전송하는 VHD 필터를 포함하는

시스템.

청구항 19

제 16 항에 있어서,

상기 제 1 CSM 및 상기 제 2 CSM은 하나 이상의 추가의 스페닝된 볼륨 및 하나 이상의 스페닝되지 않은 볼륨을 포함하도록 동적으로 조정하는

시스템.

발명의 설명

기술 분야

배경 기술

[0001] 전통적으로, 가상 머신의 호스트는 가상 머신에 의해 액세스할 수 있는 데이터를 저장하기 위해 물리적 디스크를 제공한다. 호스트의 관점에서 데이터는 디스크 상에 파일로 보이지만, 가상 머신의 관점에서 데이터는 하드 드라이브에 유지되는 것으로 보인다. 이러한 파일은 호스트의 특정 디스크에 유지되는 가상 하드 드라이브이거나 복수의 호스트 디스크에 걸쳐 스페닝(spanning)되는 가상 하드 드라이브로 지칭된다. 가상 머신은 전통적으로 공통 호스트의 물리적 자원을 확장하기 위해 구현되기 때문에, 호스트마다 몇 개의 가상 머신이 있을 수 있다. 전통적으로, 주어진 호스트에서, 호스트 상의 모든 가상 머신은 (예를 들어, 모두 스페닝되거나(spanned) 모두 스페닝되지 않은(non-spanned)) 배타적 데이터 구조 구성으로 제한된다. 특정 타입의 데이터 구조의 배타성은 I/O 분리 가변성, 내결함성(fault tolerance) 조정 및 액세스 조정의 가능성을 방지할 수 있다.

발명의 내용

해결하려는 과제

과제의 해결 수단

[0002] 본 발명의 실시예는 둘 이상의 컴퓨터 저장 매체에 걸쳐 스페닝된 가상 하드 드라이브를 유지함과 동시에 컴퓨터 저장 매체 중 하나에서 스페닝되지 않은 가상 하드 드라이브를 유지하기 위한 시스템, 방법 및 컴퓨터 저장 매체에 관한 것이다. 방법은 볼륨 스페닝을 이용하여 컴퓨터 저장 매체에 걸쳐서 스페닝된 가상 하드 드라이브의 데이터를 저장하는 단계를 포함한다. 스페닝된 가상 하드 드라이브가 컴퓨터 저장 매체에서 유지되는 동안, 방법은 컴퓨터 저장 매체 중 하나에 스페닝되지 않은 가상 하드 드라이브의 데이터를 저장하는 단계를 포함한다.

[0003] 이러한 요약은 상세한 설명에서 아래에 추가로 설명되는 개념의 선택을 단순화된 형태로 소개하기 위해 제공된다. 이러한 요약은 청구된 발명 대상의 중요한 특징 또는 필수적인 특징을 식별하도록 의도되지 않으며, 청구된 발명 대상의 범위를 결정하는데 도움을 주는 것으로 사용되도록 의도되지도 않는다.

도면의 간단한 설명

[0004] 본 발명의 예시적인 실시예는 첨부 도면을 참조하여 아래에 상세히 설명되며, 첨부 도면들은 참조에 의해 본 명세서에 통합된다.

도 1은 본 발명의 실시예를 구현하는 데 적절한 예시적인 컴퓨팅 장치를 도시한다.

도 2는 본 발명의 실시예에 따른 복수의 스페닝된 디스크를 가진 예시적인 노드를 도시한다.

도 3은 본 발명의 실시예에 따른 복수의 별개의 디스크를 가진 예시적인 노드를 도시한다.

도 4는 본 발명의 실시예에 따른 하이브리드 데이터 구조로 데이터를 유지하는 복수의 디스크를 가진 예시적인 노드를 도시한다.

도 5는 본 발명의 실시예에 따른 하이퍼바이저 계층(hypervisor layer)에서 VHD 드라이버를 가진 예시적인 호스

트를 도시한다.

도 6은 본 발명의 실시예에 따른 호스트 파일 시스템 계층에서 VHD 필터를 가진 예시적인 호스트를 도시한다.

도 7은 본 발명의 실시예에 따른 둘 이상의 컴퓨터 저장 매체(computer-storage media("CSM"))에 걸쳐서 스페닝된 가상 하드 드라이브를 유지함과 동시에 둘 이상의 CSM 중 제 1 CSM에서 스페닝되지 않은 가상 하드 드라이브를 유지하기 위한 방법을 도시한다.

도 8은 본 발명의 실시예에 따른 둘 이상의 CSM에 걸쳐서 스페닝된 가상 하드 드라이브를 유지함과 동시에 둘 이상의 CSM 중 제 1 CSM에서 스페닝되지 않은 가상 하드 드라이브를 유지하기 위한 다른 방법을 도시한다.

발명을 실시하기 위한 구체적인 내용

- [0005] 본 발명의 실시예의 발명 대상은 법정 요건을 충족시키기 위해 본 명세서에서 특정하게 설명된다. 그러나, 설명 자체는 본 특허의 범위를 제한하도록 의도되지 않는다. 오히려, 발명가는 청구된 발명 대상이 또한 다른 현재 또는 미래의 기술과 연계하여 본 문서에 설명된 것과 다른 단계 또는 유사한 단계의 조합을 포함하도록 다른 방식으로 실시될 수 있다는 것을 고려하였다.
- [0006] 본 발명의 실시예는 둘 이상의 컴퓨터 저장 매체에 걸쳐서 스페닝된 가상 하드 드라이브를 유지함과 동시에 컴퓨터 저장 매체 중 하나에서 스페닝되지 않은 가상 하드 드라이브를 유지하기 위한 시스템, 방법 및 컴퓨터 저장 매체에 관한 것이다. 이러한 방법은 볼륨 스페닝을 이용하여 컴퓨터 저장 매체에 걸쳐서 스페닝된 가상 하드 드라이브의 데이터를 저장하는 단계를 포함한다. 스페닝된 가상 하드 드라이브가 컴퓨터 저장 매체에서 유지되는 동안, 방법은 컴퓨터 저장 매체 중 하나에 스페닝되지 않은 가상 하드 드라이브의 데이터를 저장하는 단계를 포함한다.
- [0007] 다른 양태에서, 본 발명은 다중 디스크에 걸쳐 스페닝된 가상 하드 드라이브를 유지함과 동시에 이러한 디스크 중 하나에서 스페닝되지 않은 가상 하드 드라이브를 둘 다 유지하기 위한 다른 방법을 제공한다. 이러한 방법은 스페닝된 가상 하드 드라이브의 데이터에 대한 I/O 요청과 같은 제 1 요청을 수신하는 단계를 포함한다. 방법은 또한 제 1 요청을 스페닝된 가상 하드 드라이브로 전송하는 단계를 포함한다. 방법은 또한 스페닝되지 않은 가상 하드 드라이브의 데이터에 대한 제 2 요청을 수신하는 단계를 포함한다. 또한, 방법은 제 2 요청을 스페닝되지 않은 가상 하드 드라이브를 유지하는 디스크로 전송하는 단계를 포함한다. 스페닝되지 않은 가상 하드 드라이브 및 스페닝된 가상 하드 드라이브는 둘 다 적어도 부분적으로 공통 디스크에서 동시에 액세스할 수 있다.
- [0008] 본 발명의 제 3 양태는 둘 이상의 디스크에 걸쳐 스페닝된 가상 하드 드라이브를 유지함과 동시에 둘 이상의 디스크 중 제 1 디스크에서 스페닝되지 않은 가상 하드 드라이브를 유지하기 위한 시스템을 제공한다. 이러한 시스템은 호스트를 포함한다. 호스트는 가상 머신, 가상 하드 드라이브 드라이버, 및 호스트 파일 시스템으로 구성된다. 추가적으로, 호스트는 스페닝된 가상 하드 드라이브 중 적어도 부분 및 스페닝되지 않은 가상 하드 드라이브의 전체 둘 다를 동시에 유지하는 제 1 컴퓨터 저장 매체로 구성된다. 호스트는 또한 스페닝된 가상 하드 드라이브의 적어도 부분을 유지하고 스페닝되지 않은 가상 하드 드라이브를 유지하지 않는 제 2 컴퓨터 저장 매체로 구성된다.
- [0009] 본 발명의 실시예의 개요를 간단히 설명하였으므로, 본 발명의 실시예를 구현하는 데 적절한 예시적인 운영 환경이 아래에 설명된다.
- [0010] 전반적으로 도면을 참조하되 먼저 특허 도 1을 참조하면, 본 발명의 실시예를 구현하는 데 적절한 예시적인 운영 환경이 도시되며 일반적으로 컴퓨팅 장치(100)로 지칭된다. 컴퓨팅 장치(100)는 적절한 컴퓨팅 환경의 일례에 불과하며, 본 발명의 용도 또는 기능의 범위에 관해 어떠한 제한을 제공하도록 의도되지 않는다. 컴퓨팅 장치(100)는 예시된 구성 요소 중 어느 하나 또는 조합에 관한 어떤 종속성 또는 요구를 갖는 것으로 해석되지 않아야 한다.
- [0011] 실시예는 개인 휴대 정보 단말기 또는 다른 핸드헬드 장치와 같은 컴퓨터 또는 다른 머신에 의해 실행되는 프로그램 모듈과 같은 컴퓨터 실행 가능 명령어를 포함하는 컴퓨터 코드 또는 머신 사용 가능 명령어의 일반적인 맥락에서 설명될 수 있다. 일반적으로, 루틴, 프로그램, 객체, 구성 요소, 데이터 구조 등을 포함하는 프로그램 모듈은 특정 태스크를 수행하거나 특정 추상 데이터 타입을 구현하는 코드를 나타낸다. 실시예는 핸드헬드 장치, 가전 제품, 범용 컴퓨터, 전문 컴퓨팅 장치 등을 포함하는 다양한 시스템 구성에서 실시될 수 있다. 실시예는 또한 태스크가 통신 네트워크를 통해 연결되는 원격 처리 장치에 의해 수행되는 분산 컴퓨팅 환경에서 실

시될 수 있다.

- [0012] 도 1을 계속 참조하면, 컴퓨팅 장치(100)는 메모리(112), 하나 이상의 프로세서(114), 하나 이상의 프리젠테이션 모듈(116), 입력/출력(I/O) 포트(118), 입력/출력 모듈(120) 및 예시적인 전력 공급 장치(122)를 직간접적으로 결합하는 버스(110)를 포함한다. 버스(110)는 (주소 버스, 데이터 버스 또는 이의 조합과 같은) 하나 이상의 버스일 수 있는 것을 나타낸다. 도 1의 다양한 블록이 명료함을 위해 선으로 도시되지만, 실제로 다양한 모듈을 묘사하는 것은 그렇게 명확하지 않으며, 비유적으로 그 선은 더욱 정확하게는 희색이고 흐릿할 수 있다. 예를 들면, 디스플레이 장치와 같은 프리젠테이션 모듈은 I/O 모듈인 것으로 여겨질 수 있다. 또한, 프로세서는 메모리를 갖는다. 본 발명의 발명가는 그러한 것이 본 기술의 본질임을 인식하고 도 1의 다이어그램이 단지 하나 이상의 실시예와 관련하여 사용될 수 있는 예시적인 컴퓨팅 장치를 단지 예시하는 것임을 반복해서 언급한다. 모두가 도 1의 범위 및 "컴퓨터" 또는 "컴퓨팅 장치"에 대한 언급 내에서 고려되는 바와 같이, "워크 스테이션", "서버", "랩톱", "블레이드", "노드" 등과 같은 그런 카테고리들은 서로 구별되지 않는다.
- [0013] 컴퓨팅 장치(100)는 일반적으로 다양한 컴퓨터 판독 가능 매체를 포함한다. 제한이 아닌 예로서, 컴퓨터 판독 가능 매체는 랜덤 액세스 메모리(RAM), 판독 전용 메모리(ROM), EEPROM(Electronically Erasable Programmable Read Only Memory), 플래시 메모리 또는 다른 메모리 기술, CDROM, 디지털 다기능 디스크(DVD) 또는 다른 광 디스크 저장 장치, 마그네틱 카세트, 마그네틱 테이프, 마그네틱 디스크 저장 또는 다른 마그네틱 저장 장치, 또는 원하는 정보를 인코딩하는 데 사용되고 컴퓨팅 장치(100)에 의해 액세스될 수 있는 어떤 다른 매체를 포함할 수 있다.
- [0014] 메모리(112)는 휘발성 및/또는 비휘발성 메모리의 형태의 비일시적(non-transitory) 컴퓨터 저장 매체와 같은 컴퓨터 저장 매체("CSM")를 포함한다. 메모리는 이동식, 비이동식 또는 이의 조합일 수 있다. 예시적인 하드웨어 장치는 고체 메모리, 하드 드라이브, 광 디스크 드라이브 등을 포함한다. 컴퓨팅 장치(100)는 메모리(112) 또는 I/O 구성 요소(120)와 같은 다양한 엔티티로부터 데이터를 판독하는 하나 이상의 프로세서를 포함한다. 프리젠테이션 모듈(116)은 데이터 표시를 사용자 또는 다른 장치에 제공한다. 예시적인 프리젠테이션 모듈은 디스플레이 장치, 스피커, 인쇄 모듈, 진동 모듈 등을 포함한다. I/O 포트(118)는 컴퓨팅 장치(100)가 I/O 모듈(120)을 포함하는 다른 장치에 논리적으로 결합되도록 하며, 다른 장치 중 일부는 내장될 수 있다. 예시적인 모듈은 마이크, 조이스틱, 게임 패드, 위성 안테나, 스캐너, 프린터, 무선 장치 등을 포함한다.
- [0015] 도 2는 본 발명의 실시예에 따른 복수의 CSM을 가진 예시적인 노드(200)를 도시한다. 노드는 도 1의 컴퓨팅 장치(100)와 같은 컴퓨팅 장치일 수 있다. 예를 들면, 노드는 블레이드, 서버, 또는 네트워크를 통해 액세스할 수 있는 하나 이상의 데이터 부분을 유지하기 위해 활용되는 다른 분산 저장 구성 요소일 수 있다.
- [0016] 노드(200)와 같은 노드는 소정 개수 및/또는 타입의 CSM으로 구성될 수 있다는 점이 이해된다. 예를 들면, 예시적인 노드는 4개의 물리적 하드 드라이브와 같은 4개의 별개의 CSM을 포함한다. 예시적인 데이터 저장 구조에서, 제 1 디스크(디스크 0)는 호스트 시스템 및/또는 하나 이상의 가상 머신(virtual machine("VM"))과 같은 루트 운영 체제("OS")을 위해 예약될 수 있다. 이러한 예를 계속 설명하면, 나머지 3개의 CSM은 게스트 컨테이너 자원(예를 들어, VM에 의해 인식되는 바와 같은 볼륨)을 위해 예약된다. 이러한 예시적인 데이터 저장 구조의 구성은 3개의 CSM을 스페닝하는 단일 볼륨을 노출시키며, 이것은 볼륨 스페닝의 일례이다.
- [0017] 다수의 CSM에 걸쳐 볼륨을 스페닝하는 것은 3개의 별개의 CSM(또는 이의 부분)이 연속 볼륨인 모습을 VM(또는 하이퍼바이저 계층에서나 위의 어떤 계층)에 제공한다. 그러나, 호스트 레벨에서(예를 들어, 하드웨어 관점으로부터), 실제로 다수의 볼륨을 갖는 3개의 별개의 CSM이 있다. 따라서, 본 명세서에서 사용되는 용어는 그것을 보는 관점에 따라 변할 수 있다. 예를 들면, VM은 VHD를 인식할 수 있지만, 호스트는 파일(또는 볼륨, 또는 파티션)과 같은 동일한 데이터를 인식할 수 있다. 추가적인 예는 도 4에 대해서 이후에 논의될 것이다.
- [0018] 도 2로 복귀하면, 노드(200)는 디스크 A(202), 디스크 B(204) 및 디스크 C(206)를 포함한다. 디스크(202-206) 각각은 어떤 형태의 CSM(예를 들어, 휘발성, 비휘발성, RAM, 광학 하드 디스크 등)일 수 있다. 용어 "디스크"는 본 명세서에서 CSM에 대한 일반적인 용어로서 사용되며, CSM의 특정 형태로 제한하도록 의도되지 않는다. 노드(200)에서 디스크(202-204)는 볼륨/컨테이너가 모든 디스크에 걸쳐 스페닝하도록 허용하는 볼륨 스페닝을 지원하는 디스크이다. 예를 들면, 볼륨 1(208), 볼륨 2(210), 볼륨 3(212) 및 볼륨 4(214)은 모두 디스크(202-206)를 스페닝한다. 하드웨어의 관점으로부터, 볼륨 1(208)은 3개의 별개의 파일, 즉 파일 1(218), 파일 2(220) 및 파일 3(222)이다. 그러나, VM의 관점으로부터, 볼륨 1(208)은 VHD 1(216)과 같은 가상 하드 드라이브("VHD")를 유지하는 연속 볼륨이다. (또한 가상 하드 디스크로 지칭되는) VHD는 디스크 파티션 및 파일 시스템과 같은 물리적 하드 디스크 드라이브(예를 들어, CSM)에서 발견되는 것을 포함할 수 있는 가상 하드 디스크

파일 형식이다. VHD는 일반적으로 VM에 대한 "하드 디스크"로 사용된다. 따라서, 물리적 컴퓨팅 장치가 물리적 하드 디스크를 사용할 수 있는 것처럼, VM은 VHD를 유사하게 사용할 수 있다.

[0019] 도 2가 다른 구성 요소(예를 들어, 프로세서, 통신 구성 요소, 전력 공급 장치 등)를 명시적으로 도시하지 않지만, 그러한 구성 요소는 도 1에 대해 논의된 것과 같이 고려되며 컴퓨팅 장치에 전통적이라는 점이 이해된다. 따라서, 도 2(및 본 출원의 추가적인 도면)는 특정 정보를 전달하는 방식으로 예시되지만, 추가적인 구성 요소, 단계 및 항목이 본 명세서에서 고려될 수 있다는 점이 이해된다. 더욱이, 도 2가 스트립 타입의 데이터 저장 구성을 도시하지만, 임의의 볼륨 스페닝 구성이 고려되며 그 결과 본 발명의 실시예와 관련하여 이용될 수 있다는 점이 이해된다. 예를 들면, 제 1 볼륨 및 제 2 볼륨은 제 1 CSM에서 제 1 파티션을 이용할 수 있는 반면에 제 1 볼륨 및 제 2 볼륨은 또한 제 2 CSM에서 제 2 파티션을 이용한다. 이러한 예에서, 제 1 볼륨은 제 1 CSM에서 공통 파티션을 제 2 볼륨과 공유하면서 제 1 CSM 및 제 2 CSM을 스페닝한다. 상술한 것은 일례이며 추가적인 배치가 고려된다는 점이 고려된다.

[0020] 도 3은 본 발명의 실시예에 따른 복수의 CSM을 갖는 노드(300)를 도시한다. 노드(300)는 디스크 A(302), 디스크 B(304) 및 디스크 C(306)를 갖는다. 그러나, 도 2의 노드(200)과는 달리, 노드(300)는 스페닝을 이용하지 않는다. 따라서, 노드(300)에서, 각 VHD는 별개의 디스크로 한정된다. 노드(300)는 복수의 VHD(308-328)로 구성되며, 이의 각각은 다수의 물리적 디스크를 스페닝하지 않는다.

[0021] 그러나, 도 2의 노드(200)에 이용된 구조 및 도 3의 노드(300)에 이용된 구조는 둘 다 노드의 둘 이상의 디스크에 걸쳐 전용 저장 솔루션을 구현한다는 점에서 유사하다. 도 2의 경우에, 둘 이상의 디스크는 스페닝(예를 들어, 볼륨 스페닝)으로 배타적으로 구조화된다. 도 3은 VHD(예를 들어, 각 VHD는 특정 물리적 디스크로 제한된다)에 의해 별개의 디스크 이용을 배타적으로 구현하는 둘 이상의 디스크를 제공한다. "just a bunch of disks" ("JBOD") 구성은 도 3에 예시되는 것의 일례이다.

[0022] 도 2의 구성은 가장 큰 가능한 컨테이너가 저장되도록 허용한다. 추가적으로, 도 2의 저장 추상화(storage abstraction)는 개개의 디스크에 컨테이너 파일의 능동적 관리 없이 디스크에 걸쳐 컨테이너 자원의 할당을 단순화한다. 그러나, 도 2의 구성은 또한 결점을 가질 수 있다. (이러한 예에서) 3개의 디스크 중 하나에 장애가 발생하면, 전체 노드(200)에 대한 데이터가 손실될 수 있다. 이것은 디스크의 각각에서 단일 장애 포인트(single point of failure)를 가져온다. 이것은 동일한 데이터를 저장하는 데 개별적으로 필요한 디스크의 개수로 나누어지는 사용 디스크의 수의 팩터(factor)만큼 데이터 손실의 가능성을 증가시킨다. 추가적으로, 배타적으로 스페닝된 솔루션은 입력/출력("I/O") 분리를 제공하지 않는다. 예를 들면, 스페닝된 디스크 상의 VM의 집중적인(intensive) I/O 동작은 동일한 디스크를 이용하는 모든 다른 VM에 대한 응답 시간에 영향을 미친다. 따라서, 도 2의 구성은 파일 관리 및 컨테이너 크기 유지 옵션에서 효율성을 제공하며, 또한 장애 및 I/O 분리 우려를 가져온다.

[0023] 도 3에 의해 제공되는 구성은 더욱 높은 레벨의 I/O 분리 및 내결함성을 제공한다. 그러나, 단일 컨테이너(예를 들어, VHD에 사용 가능한 저장소의 부분)의 크기는 특정 디스크의 크기로 제한된다. 따라서, VM이 어떤 하나의 디스크(또는 단일 디스크에 이용 가능한 공간)보다 큰 저장 컨테이너를 필요로 할 때, 도 3의 구성은 비효율적이다.

[0024] 결과적으로, 도 2의 구성의 결함 및 도 3의 구성의 결점을 극복하기 위해, 동적이고 유연한 구성이 바람직하다. 도 4는 본 발명의 실시예에 따른 복수의 CSM을 가진 노드(400)의 하이브리드 저장 구성을 도시한다. 노드(400)는 디스크 A(402), 디스크 B(404) 및 디스크 C(406)를 포함한다. 노드(400)에 제공된 저장 솔루션은 배타적으로 스페닝하지 않고 배타적으로 별개의 것이지도 않으며, 대신에 CSM의 별개의 부분적 스페닝 및/또는 완전한 스페닝의 조합이 구현될 수 있다.

[0025] 예를 들면, VHD 1(412)은 VHD 1(412)이 VM에 의해 알 수 있듯이 볼륨 1(408)로서 컨테이너에 상주하도록 디스크 A(402) 및 디스크 B(404)의 부분적 스페닝이다. 마찬가지로, VHD 2(414), VHD 4(418), VHD 5(420), VHD 6(422) 및 VHD 7(424)은 별개의 저장 솔루션의 예이며, 여기서 이러한 VHD의 각각이 포함되는 컨테이너는 별개의 물리적 디스크로 제한된다. 더욱이, VHD 3(416)을 포함하는 컨테이너는 볼륨 2(410)로서 VM에 의해 인식되는 노드(400)의 액세스 가능한 디스크 모두를 스페닝한다.

[0026] 볼륨 1(408)은 노드(400) 내에서 CSM 모두를 스페닝하지 않고 둘 이상의 CSM을 스페닝한다. 볼륨 2(410)는 노드(400)의 디스크 모두를 스페닝한다. 따라서, 도 4에서, 예시적인 실시예는 이전의 전용 스토리지 솔루션이 공통 노드에서 공존하도록 하여, VM 이용 기준에 관계없이 VM에 의해 이용되는 어떤 자원에 대해 스페닝에서 비

스페닝으로 런타임(runtime) 또는 동적 시프트하는 것을 허용한다. 도 5 및 6에 대해 아래에서 논의되는 바와 같이, 이것은 하이퍼바이저 계층에서의 VHD 드라이버 및/또는 호스트 파일 시스템 계층에서의 VHD 필터의 사용을 통해 적어도 부분적으로 이러한 예시적인 실시예에 대해 가능하다.

[0027] 예시적인 실시예에서, VM에 대한 자원을 포함하는 컨테이너는 필요에 따라 확장되는 것이 고려된다. 따라서, 필요한 자원이 증가할 때, 컨테이너는 별개의 컨테이너(예를 들어, VHD 2(414))에서 부분적으로 스페닝된 컨테이너(예를 들어, 볼륨 1(408))로 동적으로 전환될 수 있다. 다시 말하면, 필요한 자원(예를 들어, 현재 사용되는 CSM, I/O 요청 등에 이용 가능한 공간)이 추가적인 임계 값을 초과하여 증대하면, 부분적으로 스페닝된 컨테이너는 추가적인 디스크(예를 들어, 볼륨 2(410))를 스페닝할 수 있다. 반대로, VM이 필요로 하는 자원이 감소하면, 컨테이너는 스페닝되거나 부분적으로 스페닝된 컨테이너에서 별개의 컨테이너로 축소될 수 있다. 더욱이, 노드 중 적어도 하나의 디스크는 배타적으로 별개의 컨테이너이지만 노드의 둘 이상의 다른 디스크는 스페닝 및/또는 부분적 스페닝된 컨테이너를 유지한다는 점이 고려된다. 더욱이, 컨테이너가 (저장 및 액세스를 위해) 유지되는 위치(예를 들어, 물리적 디스크)가 클라이언트, 사용자, 하이퍼바이저 및/또는 호스트 파일 시스템에 의해 선택할 수 있다는 점이 고려된다. 하나 이상의 디스크에 대한 컨테이너의 할당(예를 들어, 수량, 타입, 크기)을 선택하는 능력은 아래에서 논의되는 다양한 최적화를 허용한다. 컨테이너, 볼륨 및/또는 VHD의 어떤 조합은 본 발명의 실시예에서 구현될 수 있다는 점이 고려된다.

[0028] 도 4로 복귀하면, 서로 다른 부분/객체를 논의할 때 채택되는 관점에 따라서, 이러한 부분/객체는 서로 다른 이름에 의해 지칭될 수 있다. 예를 들면, VHD 2(414)는 호스트 레벨에서는 별개의 파일로 보일 수 있지만, VM의 관점에서는 하드 디스크(즉, VHD)로 보일 수 있다. 볼륨 1(408)은 호스트 레벨에서 다수의 스페닝된 파일로 지칭될 수 있지만, VM 레벨에서 VHD로 지칭될 수 있다. 따라서, 일부 실시예에서, 컨테이너와 같이 일반적인 방식으로 저장소의 부분을 지칭하는 것이 도움이 될 수 있다.

[0029] 도 5는 본 발명의 실시예에 따른 예시적인 호스트(500)를 도시한다. 예시적인 실시예에서의 호스트는 도 1의 컴퓨팅 장치(100)와 같은 컴퓨팅 장치이다. 용어 '호스트'는 호스트가 VM을 포함할 수 있는 하나 이상의 '게스트(guest)'를 호스팅할 때에 사용된다. 예를 들면, 호스트(500)는 가상 머신(502)의 동작을 지원할 수 있다. 더욱이, 호스트(500)는 가상 머신 N(504)과 같은 추가적인 가상 머신을 지원할 수 있다. VM과 같은 게스트는 호스트의 자원을 이용한다. 예를 들면, VM은 호스트의 물리적 공간에 실시되고, 호스트의 자원을 확장하지만, 외부의 클라이언트에 대한 호스트와 무관하게 보이는 가상화된 컴퓨팅 시스템일 수 있다. 따라서, 예시적인 실시예에서, VM은 물리적 머신과 같은 프로그램을 실행하는 컴퓨팅 장치의 소프트웨어 구현이다.

[0030] VM(502)은 VM 파일 시스템(506), 디스크 드라이버(508) 및 디스크 A(510)로 구성된다. VM 파일 시스템(506)은 가상 환경에서 운영하는 운영 체제이다. 디스크 드라이버(508)는 디스크 A(510)와 같은 VHD와 인터페이싱한다. 그러나, 디스크 A(510)는 단지 디스크 1(520) 및 디스크 Y(522)와 같은 호스트(500)의 CSM 자원에 실제로 의존하는 가상 디스크이다. VM(502)는 디스크 A(510)가 상주(resident) CSM임을 인식하지만, 실제로는 호스트의 하나 이상의 CSM에 의해 유지되는 VHD이다.

[0031] 하이퍼바이저 계층(512)은 기본(underlying) 호스트의 자원에 대한 VM 액세스를 지원하는 호스트 파일 시스템(516) 위의 인터페이스 계층이다. 하이퍼바이저 계층은 호스트의 일반적인 자원에 액세스하는 다수의 가상 머신을 지원할 수 있다. VHD 드라이버(514)는 아래에서 상세히 논의될 것이다.

[0032] 호스트 파일 시스템(516)은 호스트(500)에 대한 OS 계층이다. 호스트 파일 시스템(516)은 호스트(500)의 물리적 구성 요소, 연결부 및 프로세스를 제어한다. 도 6에 대해 논의되는 예시적인 실시예에서, 호스트 파일 시스템은 VHD 필터로 구성될 수 있다.

[0033] 호스트 디스크 드라이버(518)는 호스트 파일 시스템(516) 및 디스크 1(520)과 같은 하나 이상의 CSM의 인터페이스를 허용하는 호스트(500)를 위한 디스크 드라이버이다. 예시적인 실시예에서, 호스트 디스크 드라이버는 더욱 높은 계층(예를 들어, 호스트 파일 시스템(516))이 하나 이상의 디스크와 상호 작용하도록 하는 컴퓨팅 프로그램이다.

[0034] 디스크 1(520)은 호스트(500)의 예시적인 CSM이다. 예시적인 실시예에서, 디스크 1(520)은 호스트(500) 내에서 4개의 디스크 중 하나이다. 예시적인 실시예에서, 디스크 1(520)은 하드 드라이브, VHD A(524)로서 VM(502)에 의해 인식되는 파일을 유지한다. 마찬가지로, 디스크 Y(522)는 또한 VHD X(526)로서 VM에 의해 인식되는 파일을 포함한다. 도 5가 별개의 VHD(예를 들어, 컨테이너)를 도시하지만, 하나 이상의 스페닝된(또는 부분적으로 스페닝된) 컨테이너는 또한 하이브리드 또는 배타적인 방식으로 디스크 1(520) 및/또는 디스크 Y(522)에 통합될

수 있다는 점이 고려된다.

- [0035] 본 명세서에서 설명된 도 5 및 다른 배치는 예로서만 설명된다고 이해되어야 한다. 다른 배치 및 요소(예를 들어, 머신, 인터페이스, 기능, 순서 및 기능의 그룹화, 계층, 레벨, 드라이버, Oss, VM, CSM 등)는 도시된 것 이외에 또는 대신에 사용될 수 있으며, 일부 요소는 모두 생략될 수 있다. 본 명세서에서 설명된 추가의 많은 요소는 별개 또는 분산된 구성 요소로서 또는 다른 구성 요소와 함께 어떤 적절한 조합 및 위치에서 구현 될 수 있는 기능적 엔티티이다. 하나 이상의 엔티티에 의해 수행되는 것으로 본 명세서에 설명된 다양한 기능은 하드웨어, 펌웨어 및/또는 소프트웨어에 의해 실행될 수 있다. 예를 들면, 다양한 기능은 메모리에 저장된 명령어를 실행하는 프로세서에 의해 수행될 수 있다.
- [0036] 구성 요소는 제한없이 하나 이상의 근거리 통신망(LAN) 및/또는 광역 통신망(WAN)을 포함할 수 있는 네트워크를 통해 서로 통신할 수 있다. 이러한 네트워킹 환경은 사무실, 전사적 컴퓨터 네트워크, 인트라넷 및 인터넷에서 아주 흔하다. 많은 호스트, 서버, VM 및 CSM은 본 발명의 범위 내에서 시스템 내에 채용될 수 있다는 것이 이해되어야 한다. 추가적으로, 도시되지 않은 다른 구성 요소는 또한 이러한 시스템에 포함될 수 있다.
- [0037] 도 5를 참조하면, 도 5는 호스트(500)의 CSM에 유지된 데이터에 액세스하기 위해 I/O 요청을 전달하는 VM(502)을 도시한다. I/O 요청은 VM(502)으로부터 전달되고 하이퍼바이저 계층(512)에서 수신된다. I/O 요청은 VHD 드라이버(514)에 의해 하이퍼바이저 계층(512)에서 인터셉트(intercept)된다.
- [0038] VHD 드라이버(514)는 저장 옵션을 식별하기 위해 I/O 요청을 인터셉트하는 프로그램이다. 저장 옵션은 요청된 데이터가 저장되는 위치(예를 들어, 무슨 컨테이너, 무슨 호스트의 CSM, 무슨 디렉토리 등)를 결정하는 것을 포함할 수 있다. 저장 옵션은 데이터가 저장되어야 하는 위치(예를 들어, 무슨 컨테이너, 무슨 호스트의 CSM, 무슨 디렉토리 등)를 결정하는 것을 포함할 수 있다. 저장 옵션은 데이터가 저장되는 방법(예를 들어, 스페닝된, 부분적으로 스페닝된, 별개의)을 결정하는 것을 포함할 수 있다. 저장 옵션은 데이터가 마이그레이션(migration)해야 하는지(예를 들어, 특정 CSM을 히트(hit)하는 I/O의 최적화, 컨테이너의 확장, 컨테이너의 수축 등)를 판단하는 것을 포함할 수 있다.
- [0039] VHD 드라이버(514)의 기능을 달성하기 위해, 메타 파일은 VM에 유지될 수 있다. 예를 들면, 메타 파일은 데이터 저장 위치, 옵션, VHD 타입, 및 데이터 액세스 사이즈, 빈도 등과 관련된 계량의 인덱스를 유지할 수 있다. 메타 파일은 VHD 드라이버(514)에 의해 특정 데이터가 저장되는 위치, 특정 데이터가 저장되는 방법, 및 최적화가 수행될 수 있는지를 판단하는 데 사용될 수 있다. 최적화는 예를 들어 하나 이상의 CSM에 대한 I/O를 줄이기 위해 하나의 CSM에서 하나 이상의 다른 CSM으로 컨테이너를 이동하는 것을 포함할 수 있다. 다른 최적화는 특정 CSM에 상주하는 컨테이너의 수를 줄임으로써 결합 영향을 감소시키는 것을 포함할 수 있다. 다른 최적화는 동일한 크기의 컨테이너가 더욱 적은 CSM에 유지될 수 있을 때 컨테이너에 의해 스페닝된 CSM의 수를 줄이는 것을 포함할 수 있다. 다른 최적화가 고려된다.
- [0040] 메타 파일은 쉽고 효율적인 액세스를 위해 호스트의 특정 위치에 유지될 수 있다. 예를 들면, 메타 파일은 특정 CSM(예를 들어, 디스크 0의 제 1 섹터)의 특정 컨테이너에 유지될 수 있다. 대안으로, 메타 파일은 I/O 요청의 헤더 또는 다른 부분에서 식별되는 위치에 유지될 수 있음으로써, I/O 요청을 해석할 때 관련된 메타 파일의 위치가 결정될 수 있다.
- [0041] 예시적인 실시예에서, VHD 드라이버(514)는 런타임 환경에서 저장 구조의 동적 조정을 허용한다. 예를 들면, 하이퍼바이저 계층(512)에서 I/O 요청을 인터셉트하는 프로세스는 호스트(500)가 디스크 1(520) 및 디스크 Y(522)에 데이터를 저장하는 방법을 동적으로 조정하도록 허용한다. 이러한 동적 조정은 별개의 컨테이너가 스페닝된 컨테이너와 같은 CSM에 상주하도록 허용한다. 동적 조정은 별개의 컨테이너(예를 들어, 스페닝되지 않은 VHD)가 런타임 환경에서 스페닝된 컨테이너로 이동하도록 허용한다. 이점으로, VM(502)은 동시에 데이터를 처리하도록 허용되지만, 컨테이너는 스페닝된 상태에서 스페닝되지 않은 상태로 이동한다(또는 그 반대로 된다). 더욱이, 저장의 런타임 할당을 허용함으로써, CSM은 전체 호스트를 재구성하지 않고 호스트에서 최적화될 수 있다. 결과적으로, VHD 드라이버(514)의 활용은 I/O 패킷이 VM(502)에서 적절한 CSM으로 전송되도록 한다.
- [0042] 도 5의 이러한 예시적인 실시예에서, I/O 요청은 VHD 드라이버(514)에 의해 제공되는 방향에 따라 호스트 파일 시스템(516) 및 호스트 디스크 드라이버(518)를 통해 디스크 1(520)의 VHD A(524)로 전송된다. I/O 요청이 VHD X(526)에서 유지되는 데이터와 관련되었을 경우, VHD 드라이버(514)는 I/O 요청을 인터셉트하고, I/O 요청을 디스크 Y(522)로 다시 전송할 수 있다.
- [0043] 도 6은 본 발명의 실시예에 따른 호스트(600)를 제공한다. 예시적인 실시예에서, 호스트(600)의 다양한 계층,

드라이버, 머신, CSM 등은 도 5의 호스트(500)에 대해 이전에 논의된 유사하게 명명된 항목에 비교할 수 있다. 예를 들면, 호스트(600)는 VM(602), VM 파일 시스템(606), 디스크 드라이버(608), VM 디스크(610), 가상 머신 N(604), 하이퍼바이저 계층(612), 호스트 파일 시스템(616), VHD 필터(617), 호스트 디스크 드라이버(618), 디스크 1(620), 디스크 Y(622), VHD A(624) 및 VHD X(626)로 구성된다. 결과로서, 도 5의 유사하게 명명된 항목에 대해 논의된 상세 사항은 도 6에 대해 반복되지 않을 것이다. 그러나, 도 5의 호스트(500)와는 달리, 호스트 파일 시스템(616)은 VHD 필터(617)로 구성된다.

[0044] 예시적인 실시예에서, VHD 필터(617)는 도 5의 VHD 드라이버(514)에 대하여 상술된 것과 유사한 기능을 수행한다. 그러나, 하이퍼바이저 계층에서 I/O 패킷을 인터셉트하는 대신에, VHD 필터는 대신에 호스트 파일 시스템(616)에서 I/O 요청을 인터셉트한다. 결과적으로, VHD 필터(617)는 도 5와 함께 논의된 바와 같이 VM으로부터 직접 대신에 하이퍼바이저 계층(612)으로부터 전달되는 I/O 패킷을 인터셉트한다.

[0045] VHD 필터(617)은 또한 인터셉트된 I/O 패킷과 관련된 정보를 식별하기 위해 메타 파일에 액세스할 수 있다. 예를 들면, 메타데이터 파일로부터 식별된 바와 같이 인터셉트된 I/O 패킷과 관련된 정보는 인터셉트된 I/O에 상응한 데이터가 발견되는 위치, CSM, 컨테이너 등을 포함할 수 있다. 추가적인 정보는 최적화 결정을 하는 데 도움을 주기 위해 데이터, CSM, 컨테이너 또는 VM에 대한 액세스의 빈도를 포함할 수 있다.

[0046] 패킷이 VHD 필터(617)에 의해 인터셉트되면, 패킷은 VHD 필터(617)에 의해 이러한 예에서 디스크 1(620)과 같은 적절한 CSM으로 전송된다. VHD 필터(617)는 VM에 대한 데이터가 유지되는 위치를 식별하고, 데이터가 저장되거나 액세스될 수 있는 하나 이상의 컨테이너를 식별하고, 컨테이너가 스페닝되거나 스페닝되지 않는 지를 판단하며, 하나 이상의 CSM이 (예를 들어, 내결함성, I/O 분리, 액세스 시간 등에 따라) 데이터의 적어도 부분을 유지하기 위해 더욱 적절한 지를 판단할 수 있다는 점이 고려된다.

[0047] 예시적인 실시예에서는, I/O 패킷을 인터셉트하기 위해 하이퍼바이저에서 동작하는 VHD 드라이버를 구현하는 것이 유리하다. 예를 들면, VM에서 발신하거나 VM으로 예정된 I/O 패킷만이 하이퍼바이저 계층을 통과한다. 따라서, I/O 패킷만이 검사되거나 인터셉트된다. 이것은 호스트 파일 시스템에서 I/O 패킷을 인터셉트하고, VM과 다른 구성 요소/애플리케이션으로부터(또는 으) I/O 패킷을 인터셉트할 수 있는 VHD 필터와는 반대로 된다. 다른 예시적인 실시예에서는, I/O 패킷을 인터셉트하기 위해 VHD 필터를 구현하는 것이 유리하다. 예를 들면, VHD 필터는 정기적으로 업데이트되고 VHD 필터가 더욱 쉽게 업데이트되도록 하는 OS 내에 통합될 수 있다. 추가적인 이점이 고려된다.

[0048] 도 7은 본 발명의 실시예에 따른 둘 이상의 CSM에 걸쳐서 스페닝된 가상 하드 드라이브를 유지함과 동시에 둘 이상의 CSM 중 제 1 CSM에서 스페닝되지 않은 가상 하드 드라이브를 유지하기 위한 방법(700)의 블록도를 도시한다. 블록(702)은 2개의 CSM에 걸쳐서 스페닝된 VHD의 데이터를 저장하는 단계를 나타낸다. 예를 들면, 도 4의 블록 1(408)은 또한 도 4의 디스크 A(402) 및 디스크 B(404)를 스페닝한다. 블록 1(408)은 결과적으로 두 개의 디스크를 스페닝하는 VHD 1(412)을 유지하는 컨테이너이다. 스페닝된 VHD에 저장된 데이터는 두 개(또는 그 이상)의 CSM의 자원을 이용하는 VM의 데이터일 수 있다.

[0049] 하이퍼바이저 계층의 VHD 드라이버는 둘 이상의 CSM에 데이터를 저장하고 액세스하는 것을 담당할 수 있다. 대안으로(또는 조합하여), 호스트 파일 시스템 내의 VHD 필터는 둘 이상의 CSM에 데이터를 저장하고 액세스하는 것을 담당할 수 있다. 예시적인 실시예에서, 둘 이상의 CSM은 공통 머신(예를 들어, 서버, 블레이드, 컴퓨팅 장치)에 배치된다. 또한, 예시적인 실시예에서, 두 개 이상의 CSM은 (예를 들어, 모두 스페닝된, 모두 별개의) 전통적으로 전용 구조 형식으로 데이터를 유지하지만, VHD 드라이버 또는 VHD 필터 중 어느 하나의 결과로서 하이브리드 구조 형식을 유지할 수 있다.

[0050] 블록(704)에서, 스페닝되지 않은 VHD의 데이터는 블록(702)에서 둘 이상의 CSM 중 제 1 CSM에 저장된다. 예시적인 실시예에서, 스페닝된 VHD의 데이터 및 스페닝되지 않은 VHD의 데이터는 적어도 부분적으로 동시에 공통 CSM에 배치된다. 결과로서, 스페닝된 VHD의 데이터 및 스페닝되지 않은 VHD의 데이터는 실질적으로 동시에 공통 VHD로부터 액세스할 수 있다. 용어 "실질적으로 동시에(substantially simultaneously)"는 데이터의 하나의 부분만이 어떤 주어진 시간에 액세스/저장되도록 하지만, 예를 들어 데이터의 판독/기록의 다른 부분이 액세스될 수 있는 직후에 허용할 수 있는 CSM의 제한을 설명할 수 있다. 예시적인 실시예에서, 블록(702) 및 블록(704)은 도 4의 VHD 1(412) 및 VHD 2(414)가 디스크 A(402)에 유지되도록 할 수 있다.

[0051] 블록(706)에서, I/O 요청(예를 들어, 패킷)은 인터셉트된다. 상술한 바와 같이, VHD 드라이버 또는 VHD 필터 중 하나(또는 둘 다)는 I/O 요청을 인터셉트하기 위해 구현될 수 있다.

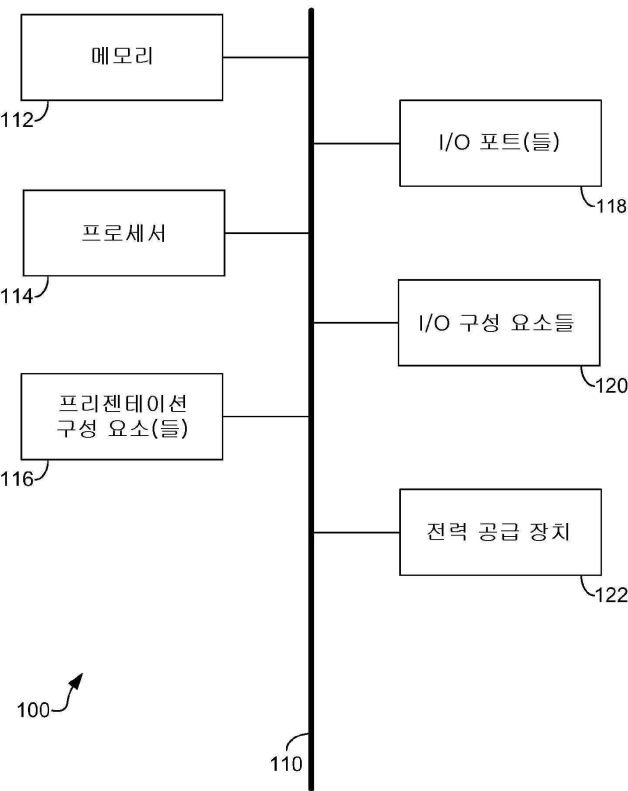
- [0052] 블록(708)에서, 인터셉트된 I/O 요청은 적절한 CSM으로 전송된다. 예를 들면, VHD 드라이버 또는 VHD 필터는 특히 I/O 요청 내용이 전송될 수 있는 CSM의 컨테이너, VHD를 식별하기 위해 I/O 요청을 다시 제시할 수 있다. 마찬가지로, VHD 드라이버 및/또는 VHD 필터는 I/O 요청의 내용을 적절한 CSM, 컨테이너, VHD 등에 전송하기 위해 인터셉트된 I/O 요청을 변경할 수 있다. VHD 드라이버 또는 VHD 필터는 원하는 VHD 수신지를 결정하기 위해 I/O 요청을 검사할 수 있다. 검사한다는 것은 나중에 결정할 시에 사용하기 위한 패킷과 관련된 정보를 식별하는 프로세스이다. 예를 들면, 검사 중에 식별된 정보는 메타 파일 또는 메타 파일 내의 정보를 찾는 데 사용될 수 있다.
- [0053] 블록(704)에서 어떤 데이터가 저장된 스페닝되지 않은 VHD는 스페닝되지 않은 VHD가 둘 이상의 CSM을 스페닝하도록 하기 위해 동적으로 변경될 수 있다는 점이 고려된다. 런타임에서 컨테이너의 저장 구조를 동적으로 변경하는 능력은 정적 구성에 의해 이전에 인식된 것보다 더 큰 정도의 유연성을 허용한다. 마찬가지로, 패닝된 VHD는 별개의 CSM(예를 들어, 하드 디스크의 특정 플래튼, 특정 하드 디스크 등)에서 별개의 컨테이너(예를 들어, 별개의 비스페닝 볼륨, 별개의 파일 등)가 되도록 런타임 동안(또는 어떤 점에서) 동적으로 변경될 수 있다는 점이 고려된다
- [0054] 도 8은 본 발명의 실시예에 따른 둘 이상의 CSM에 걸쳐서 스페닝된 가상 하드 드라이브를 유지함과 동시에 둘 이상의 CSM 중 제 1 CSM에서 스페닝되지 않은 가상 하드 드라이브를 유지하기 위한 방법(800)의 블록도를 도시한다. 블록(802)은 스페닝된 VHD의 데이터에 대해 수신되는 제 1 요청을 제공한다. 예시적인 실시예에서, 제 1 요청은 하이퍼바이저 계층 또는 호스트 파일 시스템 계층에서 수신된다. 요청은 하이퍼바이저 계층 및 호스트 파일 시스템 계층 위에 상주하는 VM으로부터의 I/O 요청일 수 있다.
- [0055] 블록(804)에서, 수신된 제 1 요청은 둘 이상의 CSM을 스페닝하는 스페닝된 VHD로 전송된다. 예를 들면, 제 1 요청은 VHD에 저장된 데이터에 액세스하려고 시도하는 VM 머신으로부터의 I/O 요청일 수 있다. VM은 액세스하려고 하는 VHD가 둘 이상의 CSM에 걸쳐 스페닝되는 것을 인식하지 못할 수 있다. 더욱이, VM은 VHD가 유지되는 위치조차도 인식하지 못할 수 있다. 결과적으로, VHD 드라이버 또는 VHD 필터는 이러한 속성(예를 들어, 요청된 데이터가 저장되는 위치 및 데이터를 액세스하는 방법)을 식별할 수 있다. 블록(806)에서, 스페닝되지 않은 VHD의 데이터에 대한 제 2 요청이 수신된다. 제 2 요청은 VHD 드라이버 또는 I/O 패킷을 인터셉트하는 VHD 필터에 의해 수신될 수 있다.
- [0056] 블록(808)에서, 제 2 요청은 둘 이상의 CSM 중 제 1 CSM으로 전송된다. VHD 드라이버 또는 VHD 필터는 제 2 요청을 전송하거나 제 2 요청이 제 1 CSM으로 전송되도록 한다. 예시적인 실시예에서, 제 1 요청 및 제 2 요청이 제각기 전송되는 스페닝된 VHD 및 스페닝되지 않은 VHD는 적어도 부분적으로 공통 CSM에 유지된다.
- [0057] 블록(810)에서, 제 3 VHD의 데이터에 대한 제 3 요청이 수신된다. 예를 들면, VHD 드라이버 또는 VHD 필터는 제 3 요청을 수신하기 위해 I/O 패킷을 인터셉트할 수 있다. 그 후, 수신 엔티티는 제 3 VHD를 포함하는 컨테이너(예를 들어, 볼륨)가 불충분한 지를 판단할 수 있다. 불충분한 컨테이너는 추가적인 데이터가 저장되도록 하기 위한 충분한 공간을 갖지 않은 컨테이너일 수 있다. 불충분한 컨테이너는 단위 시간 당 I/O의 결정된 임계값을 충족하지 않는 컨테이너일 수 있다. 불충분한 컨테이너는 내결함성의 임계 레벨을 제공하지 않는 컨테이너일 수 있다. 이러한 결정은 VHD 드라이버, VHD 필터, OS, 하이퍼바이저 및/또는 어떤 조합에 의해 수행될 수 있다.
- [0058] 블록(814)에서, 제 3 VHD를 유지하는 컨테이너가 불충분하다는 판단에 응답하여, 제 3 VHD는 제 1 CSM 및 공통 노드 중 적어도 하나의 추가적인 CSM에 걸쳐 스페닝된다. 결과적으로, 예시적인 실시예에서, 제 1 CSM은 스페닝된 VHD 중 적어도 부분, 스페닝되지 않은 VHD, 및 적어도 하나의 다른 CSM을 스페닝하는 동적으로 변경된 VHD 중 적어도 부분을 유지한다.
- [0059] 본 명세서에서 제공되는 추가적인 이점은 하나 이상의 CSM의 장애에 의해 영향을 받는 VM 및/또는 VHD를 식별하는 능력을 포함한다. 예를 들면, 기존의 RAID-0 타입 구성에서는, 노드에서 임의의 CSM을 이용하는 모든 VM이 심지어 CSM 중 하나에 장애가 발생할 경우에 영향을 받는 것으로 알려져 있다. 유사하게, 정적 JPOD 구조에서, CSM에 장애가 발생하게 될 경우에 어느 VM이 장애가 발생된 CSM을 이용할 수 있었는지가 알려져 있다. 그러나, 스페닝되고, 부분적으로 스페닝되며, 스페닝되지 않은 컨테이너가 존재하도록 허용하는 동적으로 조정 가능한 구조에서는, 어떤 VM이 장애가 있는 CSM에 의해 영향을 받는다는 것을 식별하는 것이 어려울 수 있다. 메타 파일은 영향을 받는 CSM에 데이터를 유지할 수 있는 이러한 VM을 식별하기 위해 OS, 드라이버, 하이퍼바이저 또는 임의의 다른 구성 요소/프로세스에 의해 액세스될 수 있다.

[0060] 도시된 다양한 구성 요소뿐만 아니라 도시되지 않은 구성 요소의 많은 서로 다른 배치가 본 발명의 사상 및 범위를 벗어나지 않으면서 가능하다. 본 발명의 실시예는 제한적이라기보다는 예시적인 의도로 설명되었다. 대안적인 실시예가 본 발명의 범위에서 벗어나지 않으면서 당업자에게 자명하게 될 것이다. 당업자는 본 발명의 범위에서 벗어나지 않으면서 상술한 개선 사항을 구현하는 대안적인 수단을 개발할 수 있다.

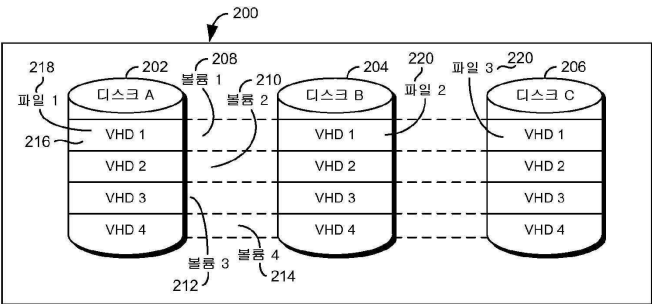
[0061] 특정 특징 및 부조합(subcombination)은 유용하고, 다른 특징 및 부조합에 관계없이 사용될 수 있으며 청구의 범위 내에서 고려된다는 것이 이해될 것이다. 다양한 도면에 열거된 모든 단계가 설명된 특정 순서로 수행될 필요가 없다.

도면

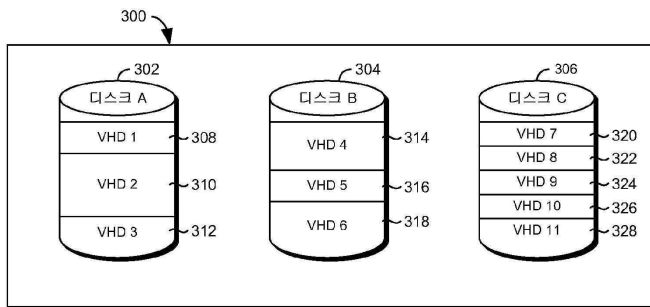
도면1



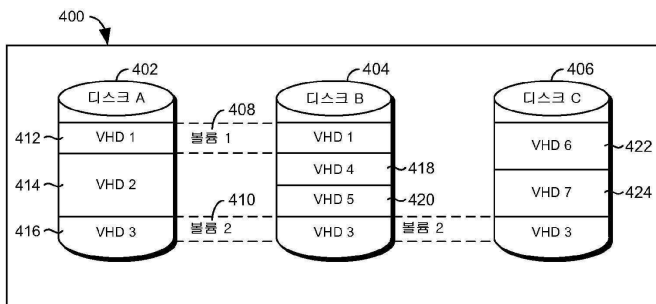
도면2



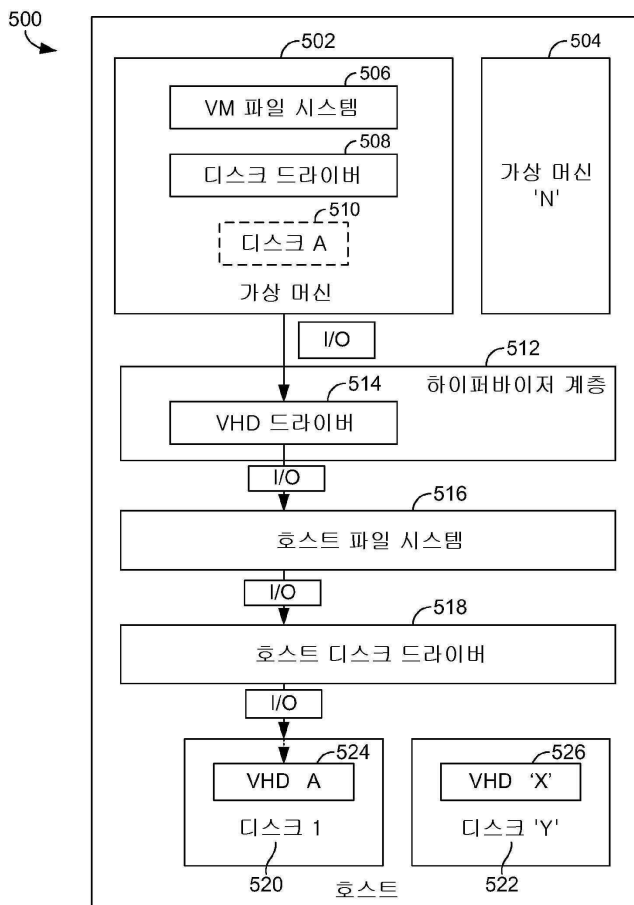
도면3



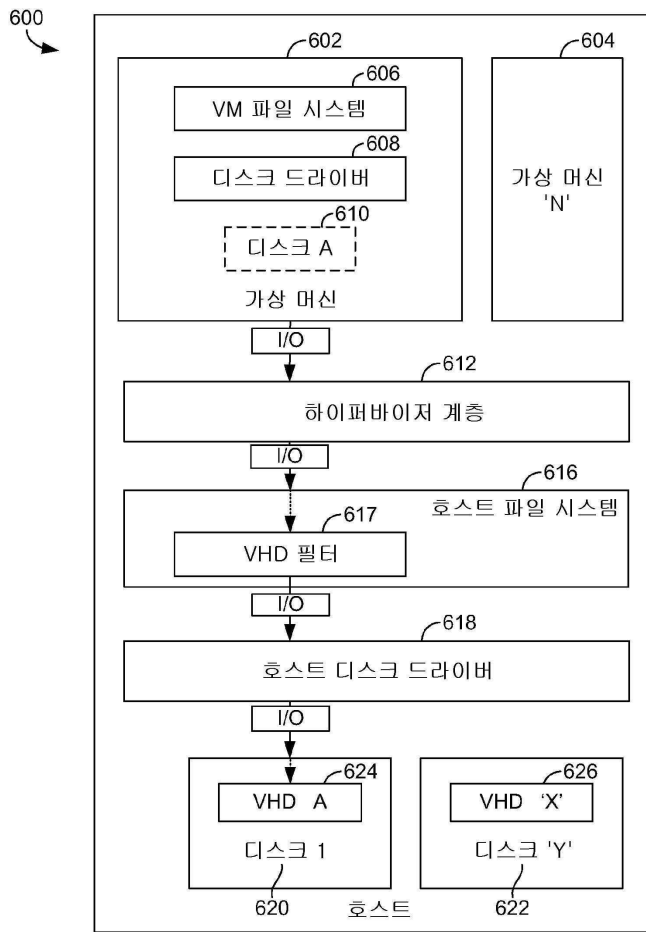
도면4



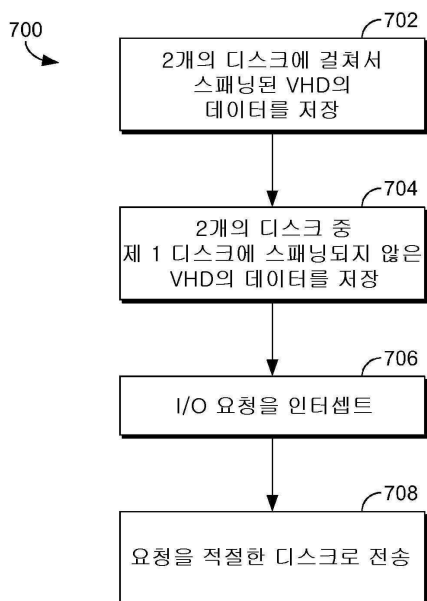
도면5



도면6



도면7



도면8

