



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2017-0075009
 (43) 공개일자 2017년06월30일

(51) 국제특허분류(Int. Cl.)
 G06F 17/30 (2006.01)
 (52) CPC특허분류
 G06F 17/30091 (2013.01)
 G06F 17/30194 (2013.01)
 (21) 출원번호 10-2017-7014507
 (22) 출원일자(국제) 2014년10월31일
 심사청구일자 2017년05월26일
 (85) 번역문제출일자 2017년05월26일
 (86) 국제출원번호 PCT/CN2014/090018
 (87) 국제공개번호 WO 2016/065611
 국제공개일자 2016년05월06일

(71) 출원인
 후아웨이 테크놀로지 컴퍼니 리미티드
 중국 518129 광둥성 셴젠 룡강 디스트릭트 반티안 후아웨이 어드미니스트레이션 빌딩
 (72) 발명자
 수 준
 중국 518129 광둥 셴젠 룡강 디스트릭트 반티안 후아웨이 어드미니스트레이션 빌딩
 왕 위안강
 중국 518129 광둥성 셴젠 룡강 디스트릭트 반티안 후아웨이 어드미니스트레이션 빌딩
 주 광위
 중국 518129 광둥 셴젠 룡강 디스트릭트 반티안 후아웨이 어드미니스트레이션 빌딩
 (74) 대리인
 유미특허법인

전체 청구항 수 : 총 18 항

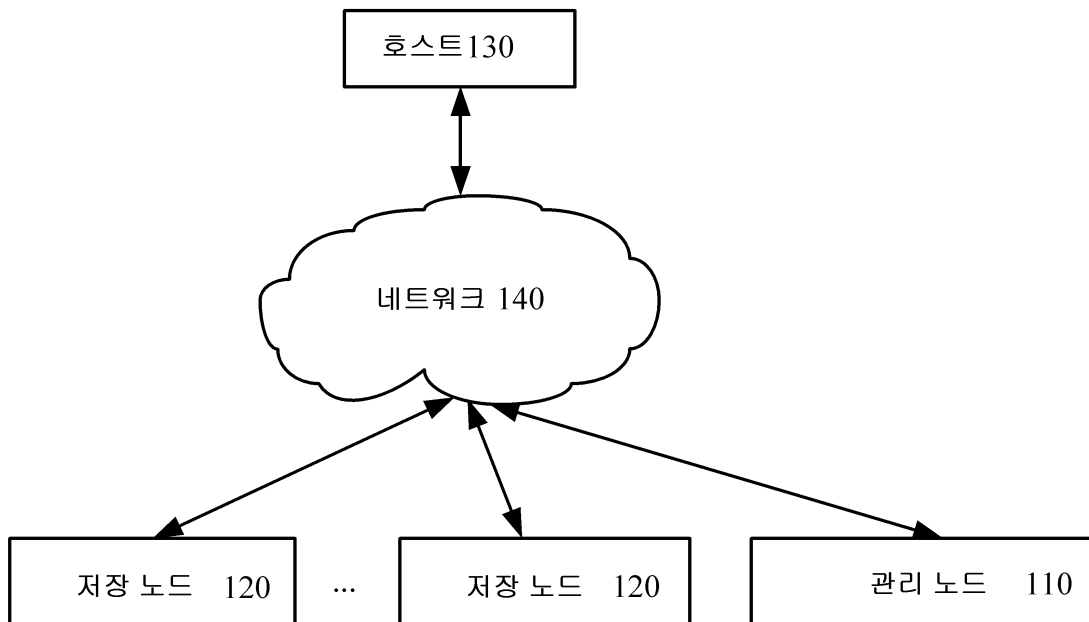
(54) 발명의 명칭 파일 액세스 방법, 시스템, 및 호스트

(57) 요약

본 발명은 파일 액세스 방법, 시스템 및 호스트를 개시한다. 이러한 방법은, 호스트가, 타깃 파일의 제1 가상 공간에 관한 정보를 획득하는 단계 - 상기 제1 가상 공간은 관리 노드에 의해, 글로벌 가상 어드레스 공간 내에서 타깃 파일에 할당됨 - ; 제1 가상 공간에 관한 정보에 따라 호스트의 로컬 가상 어드레스 공간 내에서 제2 가상

(뒷면에 계속)

대표도 - 도1



공간을 타깃 파일에 할당하는 단계; 제1 가상 공간과 제2 가상 공간 사이의 대응 관계를 구축하는 단계; 제1 가상 공간과 제2 가상 공간 사이의 대응 관계에 따라 제2 가상 공간을 액세스하는 제2 액세스 요청을 제1 가상 공간을 액세스하는 제1 액세스 요청으로 변환하는 단계 - 제1 액세스 요청의 제1 가상 공간의 어드레스는 제1 저장 노드의 장치 정보를 포함함 -; 및 네트워크 장치가 제1 액세스 요청을 제1 저장 노드로 라우팅할 수 있도록, 제1 액세스 요청을 네트워크 장치에 송신하는 단계를 포함한다. 본 발명의 실시예에 따르면, 파일 액세스 효율은 개선될 수 있다.

명세서

청구범위

청구항 1

파일 액세스 방법으로서,

상기 파일 액세스 방법은 분산 저장 시스템(distributed storage system)에 적용되고, 상기 분산 저장 시스템은 관리 노드, 복수의 저장 노드, 및 네트워크 장치를 포함하고, 상기 분산 저장 시스템 내의 각 저장 노드의 로컬 가상 어드레스 공간(local virtual address space)은 상기 분산 저장 시스템의 글로벌 가상 어드레스 공간(global virtual address space)의 일부에 대응하고,

상기 파일 액세스 방법은,

호스트(host)가, 액세스할 타깃 파일의 제1 가상 공간에 관한 정보를 획득하는 단계 - 상기 제1 가상 공간은 상기 관리 노드에 의해 상기 글로벌 가상 어드레스 공간 내에서 상기 타깃 파일에 할당된 공간임 -;

상기 호스트가, 상기 제1 가상 공간에 관한 정보에 따라 상기 호스트의 로컬 가상 어드레스 공간 내에서 제2 가상 공간을 상기 타깃 파일에 할당하는 단계;

상기 호스트가, 상기 제1 가상 공간과 상기 제2 가상 공간 사이의 대응 관계를 구축하는 단계;

상기 호스트가, 상기 제1 가상 공간과 제2 가상 공간 사이의 대응 관계에 따라, 상기 제2 가상 공간을 액세스하는 제2 액세스 요청을 상기 제1 가상 공간을 액세스하는 제1 액세스 요청으로 변환하는 단계 - 상기 제1 액세스 요청은 상기 제1 가상 공간의 어드레스를 포함하고, 상기 제1 가상 공간의 어드레스는 제1 저장 노드의 장치 정보를 포함하며, 상기 제1 저장 노드는 상기 복수의 저장 노드 내에 있고 상기 타깃 파일을 저장하도록 구성된 저장 노드임 -; 및

상기 호스트가, 상기 네트워크 장치가 상기 제1 가상 공간의 어드레스 내에 있는 제1 저장 노드의 장치 정보에 따라 상기 제1 액세스 요청을 상기 제1 저장 노드에 라우팅(route)할 수 있도록, 상기 제1 액세스 요청을 상기 네트워크 장치로 송신하는 단계

를 포함하는 파일 액세스 방법.

청구항 2

제1항에 있어서,

상기 호스트가, 액세스할 타깃 파일의 제1 가상 공간에 관한 정보를 획득하는 단계는,

상기 호스트가, 상기 타깃 파일의 메타 데이터를 획득하고, 상기 타깃 파일의 메타 데이터로부터 상기 타깃 파일의 상기 제1 가상 공간에 관한 정보를 획득하는 단계

를 포함하는, 파일 액세스 방법.

청구항 3

제1항 또는 제2항에 있어서,

상기 제1 가상 공간에 관한 정보는 상기 제1 가상 공간의 시작 어드레스 및 크기를 포함하는, 파일 액세스 방법.

청구항 4

제3항에 있어서,

상기 호스트가, 상기 제1 가상 공간에 관한 정보에 따라 상기 호스트의 로컬 가상 어드레스 공간 내에서 제2 가상 공간을 상기 타깃 파일에 할당하는 단계는

상기 호스트가, 상기 호스트의 로컬 가상 어드레스 공간 내에서, 상기 제1 가상 공간과 동일한 크기를 가지는

상기 제2 가상 공간을 상기 타깃 파일에 할당하는 단계를 포함하는, 파일 액세스 방법.

청구항 5

제4항에 있어서,

상기 호스트가, 상기 제1 가상 공간과 상기 제2 가상 공간 사이의 대응 관계를 구축하는 단계는,

상기 호스트가, 상기 제1 가상 공간과 상기 제2 가상 공간 사이의 대응 관계인 이하:

$LVA = GVA - GVA_start) + LVA_start$ 를 구축하는 단계

를 포함하고,

상기 GVA_start 는 상기 제1 가상 공간의 시작 어드레스이고 LVA_start 는 상기 제2 가상 공간의 시작 어드레스이며 GVA 는 상기 제1 가상 공간의 어드레스이고 LVA 는 상기 제1 가상 공간의 어드레스에 대응하는 상기 제2 가상 공간의 어드레스인, 파일 액세스 방법.

청구항 6

제1항 내지 제5항 중 어느 한 항에 있어서,

상기 제1 가상 공간의 어드레스의 비트 수는 $2N$ 이고 N 은 상기 제1 저장 노드의 로컬 가상 어드레스 공간의 어드레스의 비트 수이며, 상기 제1 가상 공간의 어드레스의 N 개의 최상위 비트(most significant bit)는 상기 제1 저장 노드의 장치 정보를 포함하고, 상기 제1 가상 공간의 어드레스의 N 개의 최하위 비트(least significant bit)는 상기 제1 저장 노드에 있고 상기 타깃 파일을 저장하는 데 사용되는 상기 로컬 가상 어드레스 공간의 어드레스인, 파일 액세스 방법.

청구항 7

시스템으로서,

호스트 및 분산 저장 시스템을 포함하고, 상기 분산 저장 시스템은 관리 노드, 복수의 저장 노드, 및 네트워크 장치를 포함하며, 상기 분산 저장 시스템 내에서 각 저장 노드의 로컬 가상 어드레스 공간은 상기 분산 저장 시스템의 글로벌 가상 어드레스 공간의 일부와 대응하고,

상기 호스트는,

액세스할 타깃 파일의 제1 가상 공간에 관한 정보를 획득하고 - 상기 제1 가상 공간은 상기 관리 노드에 의해 상기 글로벌 가상 어드레스 공간 내에서 상기 타깃 파일에 할당된 공간임 -,

상기 제1 가상 공간에 관한 정보에 따라, 상기 호스트의 로컬 가상 어드레스 공간 내에서, 제2 가상 공간을 상기 타깃 파일에 할당하고,

상기 제1 가상 공간과 상기 제2 가상 공간 사이의 대응 관계를 구축하고,

상기 제1 가상 공간과 제2 가상 공간 사이의 대응 관계에 따라, 상기 제2 가상 공간을 액세스하는 제2 액세스 요청을 상기 제1 가상 공간을 액세스하는 제1 액세스 요청으로 변환하고 - 상기 제1 액세스 요청은 상기 제1 가상 공간의 어드레스를 포함하고, 상기 제1 가상 공간의 어드레스는 제1 저장 노드의 장치 정보를 포함하며, 상기 제1 저장 노드는 상기 복수의 저장 노드 내에 있고 상기 타깃 파일을 저장하도록 구성된 저장 노드임 -,

상기 네트워크 장치가, 상기 제1 가상 공간의 어드레스 내에 있는 제1 저장 노드의 장치 정보에 따라 상기 제1 액세스 요청을 상기 제1 저장 노드에 라우팅할 수 있도록, 상기 제1 액세스 요청을 상기 네트워크 장치로 송신하도록 구성된, 시스템.

청구항 8

제7항에 있어서,

상기 호스트는 구체적으로,

상기 타깃 파일의 메타 데이터를 획득하고, 상기 타깃 파일의 메타 데이터로부터 상기 타깃 파일의 상기 제1 가

상 공간에 관한 정보를 획득하도록 구성된, 시스템.

청구항 9

제7항 또는 제8항에 있어서,

상기 제1 가상 공간에 관한 정보는 상기 제1 가상 공간의 시작 어드레스 및 크기를 포함하는, 시스템.

청구항 10

제9항에 있어서,

상기 호스트는 구체적으로,

상기 호스트의 로컬 가상 어드레스 공간 내에서, 상기 제1 가상 공간과 동일한 크기를 가지는 상기 제2 가상 공간을 상기 타깃 파일에 할당하도록 구성된, 시스템.

청구항 11

제10항에 있어서,

상기 호스트는 구체적으로,

상기 제1 가상 공간과 상기 제2 가상 공간 사이의 대응 관계인 이하:

$LVA = GVA - GVA_start) + LVA_start$ 를 구축하도록 구성되고,

상기 GVA_start 는 상기 제1 가상 공간의 시작 어드레스이고 LVA_start 는 상기 제2 가상 공간의 시작 어드레스이며 GVA 는 상기 제1 가상 공간의 어드레스이고 LVA 는 상기 제1 가상 공간의 어드레스에 대응하는 상기 제2 가상 공간의 어드레스인, 시스템.

청구항 12

제7항 내지 제11항 중 어느 한 항에 있어서,

상기 제1 가상 공간의 어드레스의 비트 수는 $2N$ 이고, N 은 상기 제1 저장 노드의 로컬 가상 어드레스 공간의 어드레스의 비트 수이며, 상기 제1 가상 공간의 어드레스의 N 개의 최상위 비트는 상기 제1 저장 노드의 장치 정보를 포함하고, 상기 제1 가상 공간의 어드레스의 N 개의 최하위 비트는 상기 제1 저장 노드에 있고 상기 타깃 파일을 저장하는 데 사용되는 상기 로컬 가상 어드레스 공간의 어드레스인, 시스템.

청구항 13

호스트로서,

액세스할 타깃 파일의 제1 가상 공간에 관한 정보를 획득하고 - 상기 제1 가상 공간은, 분산 저장 시스템의 글로벌 가상 어드레스 공간 내에서 상기 분산 저장 시스템 내의 관리 노드에 의해 상기 타깃 파일에 할당된 공간이고, 상기 분산 저장 시스템 내에서 각 저장 노드의 로컬 가상 어드레스 공간은 상기 글로벌 가상 어드레스 공간의 일부에 대응함 -, 상기 호스트의 로컬 가상 어드레스 공간 내에서, 상기 제1 가상 공간에 관한 정보에 따라 제2 가상 공간을 상기 타깃 파일에 할당하며, 상기 제1 가상 공간과 제2 가상 공간 사이의 대응 관계를 구축하고, 상기 제1 가상 공간과 제2 가상 공간 사이의 대응 관계에 따라, 상기 제2 가상 공간을 액세스하는 제2 액세스 요청을 상기 제1 가상 공간을 액세스하는 제1 액세스 요청으로 변환하도록 구성된 처리 모듈 - 상기 제1 액세스 요청은 상기 제1 가상 공간의 어드레스를 포함하고, 상기 제1 가상 공간의 어드레스는 제1 저장 노드의 장치 정보를 포함하며, 상기 제1 저장 노드는 상기 복수의 저장 노드 내에 있고 상기 타깃 파일을 저장하도록 구성된 저장 노드 임; 및

상기 네트워크 장치가, 상기 제1 가상 공간의 어드레스 내에 있는 제1 저장 노드의 장치 정보에 따라 상기 제1 액세스 요청을 상기 제1 저장 노드에 라우팅할 수 있도록, 상기 제1 액세스 요청을 상기 분산 저장 시스템 내의 네트워크 장치에 송신하도록 구성된 송신 모듈

포함하는 호스트.

청구항 14

제13항에 있어서,

상기 처리 모듈은 구체적으로,

상기 타깃 파일의 메타 데이터를 획득하고, 상기 타깃 파일의 메타 데이터로부터 상기 타깃 파일의 상기 제1 가상 공간에 관한 정보를 획득하도록 구성된, 호스트.

청구항 15

제13항 또는 제14항에 있어서,

상기 제1 가상 공간에 관한 정보는 상기 제1 가상 공간의 시작 어드레스 및 크기를 포함하는, 호스트.

청구항 16

제15항에 있어서,

상기 처리 모듈은 구체적으로,

상기 호스트의 로컬 가상 어드레스 공간 내에서, 상기 제1 가상 공간과 동일한 크기를 가지는 상기 제2 가상 공간을 상기 타깃 파일에 할당하도록 구성된, 호스트.

청구항 17

제16항에 있어서,

상기 처리 모듈은 구체적으로,

상기 제1 가상 공간과 상기 제2 가상 공간 사이의 대응 관계인 이하:

$LVA = GVA - GVA_start) + LVA_start$ 를 구축하도록 구성되고,

상기 GVA_start는 상기 제1 가상 공간의 시작 어드레스이고 LVA_start는 상기 제2 가상 공간의 시작 어드레스이며 GVA는 상기 제1 가상 공간의 어드레스이고 LVA는 상기 제1 가상 공간의 어드레스에 대응하는 상기 제2 가상 공간의 어드레스인, 호스트.

청구항 18

제13항 내지 제17항 중 어느 한 항에 있어서,

상기 제1 가상 공간의 어드레스의 비트 수는 $2N$ 이고, N 은 상기 제1 저장 노드의 로컬 가상 어드레스 공간의 어드레스의 비트 수이며, 상기 제1 가상 공간의 어드레스의 N 개인, 호스트의 최상위 비트는 상기 제1 저장 노드의 장치 정보를 포함하고, 상기 제1 가상 공간의 어드레스의 N 개의 최하위 비트는 상기 제1 저장 노드에 있고 상기 타깃 파일을 저장하는 데 사용되는 상기 로컬 가상 어드레스 공간의 어드레스인, 호스트.

발명의 설명

기술 분야

[0001] 본 발명은 저장 기술 분야에 관련된 것으로서, 보다 상세하게는, 파일 액세스 방법, 시스템 및 호스트에 관한 것이다.

배경 기술

[0002] PCM(Phase Change Memory), ReRAM(Resistive RAM), MRAM(Magnetic Random Access Memory)와 같은 새로운 NVM(non-volatile memory) 매체는 DRAM(Dynamic Random Access Memory)과 유사한 읽기 및 쓰기 특성을 가지며, 액세스 될 메모리 버스 상에 장착될 수 있다. 현재, 비-휘발성 메모리 매체(non-volatile memory medium)를 기반으로 한 스토리지 시스템은 업계에서 연구 중심이다.

[0003] 파일 시스템을 원격에서 액세스하기 위한 기존의 방법은 NFS(Network File System) 또는 CIFS(Common Internet File System) 프로토콜 등을 사용하는 것을 포함한다. 이러한 일반적인 파일 액세스 방식은 기본적으로 RPC(Remote Procedure Call)를 기반으로 하며 TCP/IP(Transmission Control Protocol/Internet Protocol)를

통한 액세스 방식이다. NVM 기반 스토리지 시스템의 경우, 파일 시스템에 원격에서 액세스하기 위한 기존 방법이 여전히 사용되면, 호스트의 커널 TCP/IP 프로토콜 스택(kernel TCP/IP protocol stack)은 여전히 액세스 요청을 처리해야 하므로 액세스 효율성은 상대적으로 낮다.

발명의 내용

- [0004] 본 발명의 실시예는 파일 액세스 효율을 개선하기 위해, 파일 액세스 방법, 시스템, 및 호스트를 제공한다.
- [0005] 제1 측면에 따르면, 파일 액세스 방법이 제공되고, 이러한 파일 액세스 방법은 상기 파일 액세스 방법은 분산 저장 시스템(distributed storage system)에 적용되고, 상기 분산 저장 시스템은 관리 노드, 복수의 저장 노드, 및 네트워크 장치를 포함하고, 상기 분산 저장 시스템 내의 각 저장 노드의 로컬 가상 어드레스 공간(local virtual address space)은 상기 분산 저장 시스템의 글로벌 가상 어드레스 공간(global virtual address space)의 일부에 대응하고, 상기 파일 액세스 방법은, 호스트(host)가, 액세스할 타깃 파일의 제1 가상 공간에 관한 정보를 획득하는 단계 - 상기 제1 가상 공간은 상기 관리 노드에 의해 상기 글로벌 가상 어드레스 공간 내에서 상기 타깃 파일에 할당된 공간임 -; 상기 호스트가, 상기 제1 가상 공간에 관한 정보에 따라 상기 호스트의 로컬 가상 어드레스 공간 내에서 제2 가상 공간을 상기 타깃 파일에 할당하는 단계; 상기 호스트가, 상기 제1 가상 공간과 상기 제2 가상 공간 사이의 대응 관계를 구축하는 단계; 상기 호스트가, 상기 제1 가상 공간과 제2 가상 공간 사이의 대응 관계에 따라, 상기 제2 가상 공간을 액세스하는 제2 액세스 요청을 상기 제1 가상 공간을 액세스하는 제1 액세스 요청으로 변환하는 단계 - 상기 제1 액세스 요청은 상기 제1 가상 공간의 어드레스를 포함하고, 상기 제1 가상 공간의 어드레스는 제1 저장 노드의 장치 정보를 포함하며, 상기 제1 저장 노드는 상기 복수의 저장 노드 내에 있고 상기 타깃 파일을 저장하도록 구성된 저장 노드임 -; 및 상기 호스트가, 상기 네트워크 장치가 상기 제1 가상 공간의 어드레스 내에 있는 제1 저장 노드의 장치 정보에 따라 상기 제1 액세스 요청을 상기 제1 저장 노드에 라우팅(route)할 수 있도록, 상기 제1 액세스 요청을 상기 네트워크 장치로 송신하는 단계를 포함한다.
- [0006] 제1 측면을 참조하면, 가능한 제1 구현 방식에서, 상기 호스트가, 액세스할 타깃 파일의 제1 가상 공간에 관한 정보를 획득하는 단계는, 상기 호스트가, 상기 타깃 파일의 메타 데이터를 획득하고, 상기 타깃 파일의 메타 데이터로부터 상기 타깃 파일의 상기 제1 가상 공간에 관한 정보를 획득하는 단계를 포함한다.
- [0007] 제1 측면 또는 제1 측면의 가능한 제1 구현 방식을 참조하면, 제1 측면의 가능한 제2 구현 방식에서, 상기 제1 가상 공간에 관한 정보는 상기 제1 가상 공간의 시작 어드레스 및 크기를 포함한다.
- [0008] 제1 측면의 가능한 제2 구현 방식을 참조하면, 제1 측면의 가능한 제3 구현 방식에서, 상기 호스트가, 상기 제1 가상 공간에 관한 정보에 따라 상기 호스트의 로컬 가상 어드레스 공간 내에서 제2 가상 공간을 상기 타깃 파일에 할당하는 단계는 상기 호스트가, 상기 호스트의 로컬 가상 어드레스 공간 내에서, 상기 제1 가상 공간과 동일한 크기를 가지는 상기 제2 가상 공간을 상기 타깃 파일에 할당하는 단계를 포함한다.
- [0009] 제1 측면의 가능한 제3 구현 방식을 참조하면, 제1 측면의 가능한 제4 구현 방식에서, 상기 호스트가, 상기 제1 가상 공간과 상기 제2 가상 공간 사이의 대응 관계를 구축하는 단계는, 상기 호스트가, 상기 제1 가상 공간과 상기 제2 가상 공간 사이의 대응 관계인 이하: $LVA = GVA - GVA_start + LVA_start$ 를 구축하는 단계를 포함하고, 상기 GVA_start 는 상기 제1 가상 공간의 시작 어드레스이고 LVA_start 는 상기 제2 가상 공간의 시작 어드레스이며 GVA 는 상기 제1 가상 공간의 어드레스이고 LVA 는 상기 제1 가상 공간의 어드레스에 대응하는 상기 제2 가상 공간의 어드레스이다.
- [0010] 제1 측면 또는 제1 측면의 가능한 제1 내지 제4 구현 방식을 참조하면, 제1 측면의 가능한 제5 구현 방식에서, 상기 제1 가상 공간의 어드레스의 비트 수는 $2N$ 이고 N 은 상기 제1 저장 노드의 로컬 가상 어드레스 공간의 어드레스의 비트 수이며, 상기 제1 가상 공간의 어드레스의 N 개의 최상위 비트(most significant bit)는 상기 제1 저장 노드의 장치 정보를 포함하고, 상기 제1 가상 공간의 어드레스의 N 개의 최하위 비트(least significant bit)는 상기 제1 저장 노드에 있고 상기 타깃 파일을 저장하는 데 사용되는 상기 로컬 가상 어드레스 공간의 어드레스이다.
- [0011] 제2 측면에 따르면, 시스템이 제공되고, 이러한 시스템은 호스트 및 분산 저장 시스템을 포함하고, 상기 분산 저장 시스템은 관리 노드, 복수의 저장 노드, 및 네트워크 장치를 포함하며, 상기 분산 저장 시스템 내에서 각 저장 노드의 로컬 가상 어드레스 공간은 상기 분산 저장 시스템의 글로벌 가상 어드레스 공간의 일부와 대응하고, 상기 호스트는, 액세스할 타깃 파일의 제1 가상 공간에 관한 정보를 획득하고 - 상기 제1 가상 공간은 상기 관리 노드에 의해 상기 글로벌 가상 어드레스 공간 내에서 상기 타깃 파일에 할당된 공간임 -, 상기 제1 가상

공간에 관한 정보에 따라, 상기 호스트의 로컬 가상 어드레스 공간 내에서, 제2 가상 공간을 상기 타깃 파일에 할당하고, 상기 제1 가상 공간과 상기 제2 가상 공간 사이의 대응 관계를 구축하고, 상기 제1 가상 공간과 제2 가상 공간 사이의 대응 관계에 따라, 상기 제2 가상 공간을 액세스하는 제2 액세스 요청을 상기 제1 가상 공간을 액세스하는 제1 액세스 요청으로 변환하고 - 상기 제1 액세스 요청은 상기 제1 가상 공간의 어드레스를 포함하고, 상기 제1 가상 공간의 어드레스는 제1 저장 노드의 장치 정보를 포함하며, 상기 제1 저장 노드는 상기 복수의 저장 노드 내에 있고 상기 타깃 파일을 저장하도록 구성된 저장 노드임 -, 상기 네트워크 장치가, 상기 제1 가상 공간의 어드레스 내에 있는 제1 저장 노드의 장치 정보에 따라 상기 제1 액세스 요청을 상기 제1 저장 노드에 라우팅할 수 있도록, 상기 제1 액세스 요청을 상기 네트워크 장치로 송신하도록 구성된다.

[0012] 제2 측면을 참조하면, 가능한 제1 구현 방식에서, 상기 호스트는 구체적으로, 상기 타깃 파일의 메타 데이터를 획득하고, 상기 타깃 파일의 메타 데이터로부터 상기 타깃 파일의 상기 제1 가상 공간에 관한 정보를 획득하도록 구성된다.

[0013] 제2 측면 또는 제2 측면의 가능한 제1 구현 방식을 참조하면, 제2 측면의 가능한 제2 구현 방식에서, 상기 제1 가상 공간에 관한 정보는 상기 제1 가상 공간의 시작 어드레스 및 크기를 포함한다.

[0014] 제2 측면의 가능한 제2 구현 방식을 참조하면, 제2 측면의 가능한 제3 구현 방식에서, 상기 호스트는 구체적으로, 상기 호스트의 로컬 가상 어드레스 공간 내에서, 상기 제1 가상 공간과 동일한 크기를 가지는 상기 제2 가상 공간을 상기 타깃 파일에 할당하도록 구성된다.

[0015] 제2 측면의 가능한 제3 구현 방식을 참조하면, 제2 측면의 가능한 제4 구현 방식에서, 상기 호스트는 구체적으로, 상기 제1 가상 공간과 상기 제2 가상 공간 사이의 대응 관계인 이하: $LVA = GVA - GVA_start) + LVA_start$ 를 구축하도록 구성되고, 상기 GVA_start 는 상기 제1 가상 공간의 시작 어드레스이고 LVA_start 는 상기 제2 가상 공간의 시작 어드레스이며 GVA 는 상기 제1 가상 공간의 어드레스이고 LVA 는 상기 제1 가상 공간의 어드레스에 대응하는 상기 제2 가상 공간의 어드레스이다.

[0016] 제2 측면 또는 제2 측면의 가능한 제1 내지 제4 구현 방식을 참조하면, 제2 측면의 가능한 제5 구현 방식에서, 상기 제1 가상 공간의 어드레스의 비트 수는 $2N$ 이고, N 은 상기 제1 저장 노드의 로컬 가상 어드레스 공간의 어드레스의 비트 수이며, 상기 제1 가상 공간의 어드레스의 N 개의 최상위 비트는 상기 제1 저장 노드의 장치 정보를 포함하고, 상기 제1 가상 공간의 어드레스의 N 개의 최하위 비트는 상기 제1 저장 노드에 있고 상기 타깃 파일을 저장하는 데 사용되는 상기 로컬 가상 어드레스 공간의 어드레스이다.

[0017] 제3 측면에 따르면, 호스트가 제공되고, 이러한 호스트는, 액세스할 타깃 파일의 제1 가상 공간에 관한 정보를 획득하고 - 상기 제1 가상 공간은, 분산 저장 시스템의 글로벌 가상 어드레스 공간 내에서 상기 분산 저장 시스템 내의 관리 노드에 의해 상기 타깃 파일에 할당된 공간이고, 상기 분산 저장 시스템 내에서 각 저장 노드의 로컬 가상 어드레스 공간은 상기 글로벌 가상 어드레스 공간의 일부에 대응함 -, 상기 호스트의 로컬 가상 어드레스 공간 내에서, 상기 제1 가상 공간에 관한 정보에 따라 제2 가상 공간을 상기 타깃 파일에 할당하며, 상기 제1 가상 공간과 제2 가상 공간 사이의 대응 관계를 구축하고, 상기 제1 가상 공간과 제2 가상 공간 사이의 대응 관계에 따라, 상기 제2 가상 공간을 액세스하는 제2 액세스 요청을 상기 제1 가상 공간을 액세스하는 제1 액세스 요청으로 변환하도록 구성된 처리 모듈 - 상기 제1 액세스 요청은 상기 제1 가상 공간의 어드레스를 포함하고, 상기 제1 가상 공간의 어드레스는 제1 저장 노드의 장치 정보를 포함하며, 상기 제1 저장 노드는 상기 복수의 저장 노드 내에 있고 상기 타깃 파일을 저장하도록 구성된 저장 노드 임; 및 상기 네트워크 장치가, 상기 제1 가상 공간의 어드레스 내에 있는 제1 저장 노드의 장치 정보에 따라 상기 제1 액세스 요청을 상기 제1 저장 노드에 라우팅할 수 있도록, 상기 제1 액세스 요청을 상기 분산 저장 시스템 내의 네트워크 장치에 송신하도록 구성된 송신 모듈 포함한다.

[0018] 제3 측면을 참조하면, 가능한 제1 구현 방식에서, 상기 타깃 파일의 메타 데이터를 획득하고, 상기 타깃 파일의 메타 데이터로부터 상기 타깃 파일의 상기 제1 가상 공간에 관한 정보를 획득하도록 구성된다.

[0019] 제3 측면 또는 제3 측면의 가능한 제1 구현 방식을 참조하면, 제3 측면의 가능한 제2 구현 방식에서, 상기 제1 가상 공간에 관한 정보는 상기 제1 가상 공간의 시작 어드레스 및 크기를 포함한다.

[0020] 제3 측면의 가능한 제2 구현 방식을 참조하면, 제3 측면의 가능한 제3 구현 방식에서, 상기 처리 모듈은 구체적으로, 상기 호스트의 로컬 가상 어드레스 공간 내에서, 상기 제1 가상 공간과 동일한 크기를 가지는 상기 제2 가상 공간을 상기 타깃 파일에 할당하도록 구성된다.

[0021] 제3 측면의 가능한 제3 구현 방식을 참조하면, 제3 측면의 가능한 제4 구현 방식에서, 상기 처리 모듈은 구체적

으로, 상기 제1 가상 공간과 상기 제2 가상 공간 사이의 대응 관계인 이하: $LVA = GVA - GVA_start$) + LVA_start 를 구축하도록 구성되고, 상기 GVA_start 는 상기 제1 가상 공간의 시작 어드레스이고 LVA_start 는 상기 제2 가상 공간의 시작 어드레스이며 GVA 는 상기 제1 가상 공간의 어드레스이고 LVA 는 상기 제1 가상 공간의 어드레스에 대응하는 상기 제2 가상 공간의 어드레스이다.

[0022] 제3 측면 또는 제3 측면의 가능한 제1 내지 제4 구현 방식을 참조하면, 제3 측면의 가능한 제5 구현 방식에서, 상기 제1 가상 공간의 어드레스의 비트 수는 $2N$ 이고, N 은 상기 제1 저장 노드의 로컬 가상 어드레스 공간의 어드레스의 비트 수이며, 상기 제1 가상 공간의 어드레스의 N 개인, 호스트의 최상위 비트는 상기 제1 저장 노드의 장치 정보를 포함하고, 상기 제1 가상 공간의 어드레스의 N 개의 최하위 비트는 상기 제1 저장 노드에 있고 상기 타깃 파일을 저장하는 데 사용되는 상기 로컬 가상 어드레스 공간의 어드레스이다.

[0023] 제4 측면에 따르면, 프로그램 코드를 저장하는 컴퓨터 판독 가능한 기억 매체를 포함하는 컴퓨터 프로그램 제품이 제공되고, 상기 프로그램 코드에 포함되는 명령은 제1 측면의 방법을 실행하는데 사용된다.

[0024] 전술한 기술적 해결 수단에 기초하여, 본 발명의 본 실시예에서의 파일 액세스 방법, 시스템, 및 호스트에 의하면, 분산 저장 시스템 내의 파일이 액세스 될 때, 호스트의 로컬 가상 어드레스 공간 내의 제2 가상 공간이 그 파일에 할당되고, 제2 가상 공간을 액세스하는 제2 액세스 요청은 분산 저장 시스템 내에 있고 파일의 제1 가상 공간에 액세스하는 제1 액세스 요청으로 변환되고, 제1 액세스 요청은 네트워크에 송신된다. 이러한 방식으로, 분산 저장 시스템 내의 파일이 위치하는 물리적 저장 공간은 호스트의 커널을 사용하지 않고 직접 액세스 될 수 있으므로, 지연이 감소되고 파일 액세스 효율이 개선 될 수 있다.

도면의 간단한 설명

[0025] 본 발명의 실시예에서의 기술적 해결 수단을 보다 명확하게 설명하기 위해, 이하에서 실시예를 설명하기 위해 요청되는 첨부 도면을 간단히 설명한다. 명백하게, 다음의 설명에서의 첨부 도면은 단지 본 발명의 일부 실시예를 나타낸다.

- 도 1은 본 발명의 일 실시예에 따른 애플리케이션 시나리오의 개략도이다.
- 도 2는 본 발명의 일 실시예에 따른 파일 액세스 방법의 개략적 흐름도이다.
- 도 3은 본 발명의 일 실시예에 따른 파일 공간의 개략도이다.
- 도 4는 본 발명의 일 실시예에 따른 시스템의 개략적 블록도이다.
- 도 5는 본 발명의 일 실시예에 따른 호스트의 개략적 블록도이다.
- 도 6은 본 발명의 일 실시예에 따른 호스트의 개략적 구조도이다.

발명을 실시하기 위한 구체적인 내용

[0026] 이하, 본 발명의 실시예의 첨부 도면을 참조하여, 본 발명의 실시예의 기술적 해결 수단을 명확하고 완전하게 설명한다. 명백하게, 설명된 실시예는 본 발명의 모든 실시예가 아닌 일부이다.

[0027] 본 발명의 실시예에서의 기술적 해결 수단은 다양한 저장 시스템에 적용될 수 있으며, 특히 비휘발성 저장 매체에 기초한 분산 저장 시스템에 적용될 수 있음을 이해해야 한다.

[0028] 또한, 본 발명의 실시예에서, "가상 어드레스 공간"이라는 용어는 "가상 공간", 즉 "가상 어드레스 공간"과 "가상 공간"이 같은 의미로 표현될 수 있다. 마찬가지로 "로컬 물리적 저장 공간"과 "로컬 물리적 공간"은 동일한 의미를 가진다.

[0029] 도 1은 본 발명의 일 실시예에 따른 애플리케이션 시나리오의 개략도이다. 도 1에서, 네트워크(140)를 사용하여 호스트(130), 저장 노드(120), 및 관리 노드(110)는 서로 연결되어 있다. 저장 노드(102), 관리 노드(110), 및 네트워크(140)는 분산 저장 시스템을 형성한다.

[0030] 호스트(130)는 컴퓨터 일 수도 있고 다른 사용자 장비 일 수도 있으며, 호스트(130)는 네트워크를 이용하여 분산 저장 시스템과 통신할 수 있다. 호스트(130)는 서버, 데스크톱 컴퓨터 및 운영 시스템과 같은 현재 기술의 임의의 알려진 컴퓨팅 장치를 포함할 수 있고, 운영 체제 및 다른 애플리케이션 프로그램은 호스트(130)에 설치된다. 호스트(130)는 다수 존재할 수 있다. 호스트(130)는 분산 저장 시스템에 있는 파일에 액세스할 수 있다. 예를 들어, 호스트(130)는 파일 인터페이스 모듈을 사용하여 호스트 내의 애플리케이션 프로그램에 대해 파일

액세스 인터페이스를 제공할 수 있다.

- [0031] 저장 노드(120)는 파일 데이터를 저장하고 파일 액세스 요청을 처리하도록 구성된다. 저장 노드(120)는 적어도 한 유형의 비휘발성 메모리, 예를 들어 PCM, ReRAM, 및 MRAM과 같은 바이트 단위 어드레싱 및 액세스 기능을 갖는 새로운 NVM을 포함할 수 있다. 본 발명의 실시예에서, 파일은 NVM 메모리에 저장되고, 저장 노드(120)는 메모리 버스를 사용하여 새로운 NVM에 액세스 할 수 있다.
- [0032] 관리 노드(110)는 파일의 메타 데이터를 관리하고, 파일의 메타 데이터에 대하여 호스트의, 검색 요청, 업데이트 요청 등을 처리하도록 구성된다. 파일의 메타 데이터는 파일의 시작 주소, 할당된 공간의 크기, 및 파일의 실제크기와 같은 정보를 포함할 수 있다. 관리 노드(110)는 또한 분산 저장 시스템의 글로벌 가상 어드레스 공간을 관리하도록 구성 될 수있다.
- [0033] 관리 노드(110)와 저장 노드(120)는 결합될 수 있다. 즉, 노드는 관리 노드와 저장 노드 모두의 기능을 가질 수 있다.
- [0034] 본 발명의 실시예에서, 관리 노드는 독립적으로 배치될 수 있고(예를 들어, 도 1의 관리 노드(110)), 또는 관리 노드와 저장 노드 기능을 모두 가지는 노드 일 수 있다. 설명을 쉽게 하기 위해 다음은 설명을 위해 관리 노드를 예로서 사용한다.
- [0035] 네트워크(140)의 물리 계층 및 링크 계층은 이더넷, PCIE 스위치 네트워크, InfiniBand 네트워크, 광 네트워크 등을 기반으로 할 수 있다.
- [0036] 네트워크(140)는 액세스 요청 및 액세스 응답을 라우팅하도록 구성되고, 구체적으로 네트워크(140) 내의 네트워크 장치(예를 들어, 라우터 또는 스위치)에 의해 라우팅이 수행될 수 있다. 설명의 편의를 위해, 네트워크 장치를 설명하기 위한 예로서 사용한다.
- [0037] 도 2는 본 발명의 일 실시예에 따른 파일 액세스 방법(200)의 개략적 흐름도이다.
- [0038] 이러한 방법(200)은 분산 저장 시스템에 적용되고, 분산 저장 시스템은, 관리 노드, 복수의 저장 노드, 및 네트워크 장치를 포함한다. 분산 저장 시스템 내에서 각 저장 노드의 로컬 가상 어드레스 공간은 분산 저장 시스템의 글로벌 가상 어드레스 공간의 일부에 대응한다. 방법(200)은 호스트에 의해 실행된다.
- [0039] S210. 호스트는 액세스할 타깃 파일의 제1 가상 공간에 관한 정보를 획득하고, 제1 가상 공간은 관리 노드에 의해 글로벌 가상 어드레스 공간 내에서 타깃 파일에 할당된 공간이다.
- [0040] 분산 저장 시스템의 글로벌 가상 어드레스 공간은 분산 저장 시스템 내의 파일 관리를 위해 가상화된 공간이다. 분산 저장 시스템 내에서 서로 다른 저장 노드의 로컬 가상 어드레스 공간은 분산 저장 시스템의 글로벌 가상 어드레스 공간의 상이한 부분에 대응한다. 예를 들어, 글로벌 가상 어드레스 공간의 제1 부분의 공간은 분산 저장 시스템 내의 제1 저장 노드의 로컬 가상 어드레스 공간에 매핑되고, 글로벌 가상 어드레스 공간의 제2 부분의 공간은 분산 저장 시스템 내의 제2 저장 노드에 매핑된다. 제1 부분의 공간과 제2 부분의 공간은 다르다. 본 발명의 본 실시예에서, 관리 노드는, 분산 저장 시스템의 글로벌 가상 어드레스 공간 내에서, 연속 공간의 섹션(제1 가상 공간으로 대표됨)을 타깃 파일에 할당한다. 분산 저장 시스템 내에서 각 저장 노드의 로컬 가상 어드레스 공간은 분산 저장 시스템의 글로벌 가상 어드레스 공간의 일부에 대응하기 때문에, 제1 저장 공간은, 글로벌 가상 어드레스 공간과 로컬 가상 어드레스 공간 사이의 대응관계에 따라 저장 노드의 로컬 가상 어드레스 공간에 매핑될 수 있다.
- [0041] 선택적으로, 본 발명의 실시예로서, 제1 가상 공간에 관한 정보는 제1 가상 공간의 시작 어드레스 및 크기를 포함할 수 있다.
- [0042] 글로벌 가상 어드레스 공간의 어드레스는 글로벌 가상 어드레스(global virtual address)이다. 글로벌 가상 어드레스의 비트 수는 로컬 가상 어드레스의 비트 수보다 많을 수 있다. 제1 가상 공간의 어드레스는 제1 저장 노드의 장치정보(예를 들어, ID, 또는 ID 와 노드 타입 모두)를 포함한다. 제1 저장 노드는 복수의 저장 노드 내에 있고 타깃 파일을 저장하도록 구성된 저장 노드를 나타낸다. 선택적으로, 제1 가상 공간의 어드레스의 비트 수(즉, 글로벌 가상 어드레스의 비트 $2N$ 이고, N 은 분산 저장 시스템 내의 저장 노드의 로컬 가상 어드레스 공간의 어드레스의 비트 수이다(즉, 로컬 가상 어드레스의 비트 수). 제1 가상 공간의 어드레스의 N 개의 최상위 비트는 제1 저장 노드의 장치 정보를 포함하고, 제1 가상 공간의 어드레스의 N 개의 최하위 비트는 제1 저장 노드에 있고 타깃 파일을 저장하는 데 사용되는 로컬 가상 어드레스 공간의 어드레스이다.

- [0043] 로컬 가상 어드레스의 비트 수는 운영 체제(Operating System, OS)의 비트 수이다. 예를 들어, 64-bit OS에 대해, 로컬 가상 어드레스의 비트 수는 64이고, 글로벌 가상 어드레스의 비트 수는 128일 수 있다. 이런 경우, 글로벌 가상 어드레스의 64개의 최상위 비트는 저장 노드의 장치 정보를 포함할 수 있고, 글로벌 가상 어드레스의 64개의 최하위 비트는 로컬 가상 어드레스에 대응할 수 있다. 저장 노드의 장치 정보는 64개의 최상위 비트 중 일부 비트만 사용하여 운반될 수 있고(예를 들어, 8개의 최상위 비트는 제1 저장 노드의 ID 이다.), 남은 비트는 다른 목적으로 예약(reserve)되거나 유지된다.
- [0044] 제1 가상 공간에 관한 정보는 제1 가상 공간의 시작 글로벌 가상 어드레스 및 제1 가상 공간의 크기를 포함할 수 있다.
- [0045] 선택적으로, 본 발명의 실시예로서, 호스트는 타깃 파일의 메타 데이터를 획득하고, 타깃 파일의 메타 데이터로부터 타깃 파일의 제1 가상 공간에 관한 정보를 획득할 수 있다.
- [0046] 호스트가 타깃 파일에 액세스하는 경우는, 타깃 파일을 열고 액세스하여 생성되거나, 또는 생성된 타깃 파일을 열고 액세스하는 것일 수 있다. 호스트는 타깃 파일을 열고, 타깃 파일의 메타 데이터로부터 제1 가상 공간의 시작 어드레스 및 키를 획득할 때, 관리 노드로부터 타깃 파일의 메타 데이터를 획득할 수 있다.
- [0047] S220. 호스트는 호스트의 로컬 가상 어드레스 공간 내에서, 제1 가상 공간에 관한 정보에 따라 제2 가상 공간을 타깃 파일에 할당한다.
- [0048] 본 발명의 실시예에서, 선택적으로, OS의 가상 어드레스 공간의 미사용 부분이 로컬 가상 어드레스 공간으로서 사용될 수 있다. 예를 들어, 64비트 OS의 가상 어드레스 공간 크기는 2^{64} 이며, 사용자 모드 가상 어드레스 공간(user-mode virtual address space) 및 커널 모드 가상 어드레스 공간(kernel-mode virtual address space) 외에도 예약된 가상 주소 공간이 있으며, OS 내의 예약된 가상 주소 공간은 로컬 가상 어드레스 공간으로 사용될 수 있다. 운영 체제의 가상 주소 공간은 요청 사항에 따라 재분할되어 로컬 가상 주소 공간으로 사용될 가상 주소 공간의 섹션을 얻을 수 있다.
- [0049] 호스트는, 로컬 가상 어드레스 공간 내에서 가상 어드레스 공간의 섹션(제2 가상 공간으로서 제시됨)을 액세스할 타깃 파일에 할당한다. 선택적으로, 호스트는 호스트의 로컬 가상 주소 공간 내에서 제1 가상 공간과 동일한 크기를 갖는 제2 가상 공간을 타깃 파일에 할당한다.
- [0050] S230. 호스트는 제1 가상 공간과 제2 가상 공간 사이의 대응 관계를 구축한다.
- [0051] 제2 가상 공간을 타깃 파일에 할당한 후에, 호스트는 글로벌 가상 어드레스 공간 내의 제1 가상 공간과 호스트의 가상 어드레스 공간 내의 제2 가상 공간 사이의 대응 관계를 구축한다.
- [0052] 선택적으로, 본 발명의 실시예로서, 호스트는 제1 가상 공간과 제2 가상 공간 사이의 대응 관계: $LVA = (GVA - GVA_start) + LVA_start$ (수학식 1)
- [0053] 를 구축하고, GVA_start 는 제1 가상 공간의 시작 어드레스이고, LVA_start 는 제2 가상 공간의 시작 어드레스이며, GVA 는 제1 가상 공간의 어드레스이고, LVA 는 제1 가상 공간의 어드레스에 대응하는 제2 가상 공간의 어드레스이다.
- [0054] 수학식 1을 사용하여, 제1 가상 공간의 어드레스 GVA 및 제2 가상 공간 LVA 의 어드레스 사이의 대응 관계가 구축될 수 있다.
- [0055] 다른 일대일 대응 관계가 제1 가상 공간의 어드레스 GVA 및 제2 가상 공간 LVA 의 어드레스 사이에 더 사용될 수 있으나, 이것이 수학식 1에서 나타난 대응관계를 제한하지 않는다. 예를 들어, 제1 가상 공간의 어드레스가 제2 가상 공간의 어드레스와 직접 대응하는 일대일 대응이 더 사용될 수 있다.
- [0056] S240. 호스트는 제1 가상 공간과 제2 가상 공간 사이의 대응 관계에 따라, 제2 가상 공간을 액세스하는 제2 액세스 요청을 제1 가상 공간을 액세스하는 제1 액세스 요청으로 변환하고, 제1 액세스 요청은 제1 가상 공간의 어드레스를 포함하고, 제1 가상 공간의 어드레스는 제1 저장 노드의 장치 정보를 포함한다.
- [0057] 호스트는, 제2 가상 공간을 타깃 파일에 할당하고, 제1 가상 공간과 제2 가상 공간 사이의 대응 관계를 구축한다. 제2 가상 공간을 액세스하여 호스트 내의 애플리케이션은 제2 가상 공간을 액세스한다. 즉, 호스트 내의 애플리케이션이 타깃 파일에 액세스할 때, 제2 가상 공간을 액세스하는 제2 액세스 요청이 생성된다. 예를 들어, 호스트는 제2 가상 공간의 시작 어드레스에서 파일 핸들 포인팅(file handle pointing)을 애플리케이션에 송신

하고, 애플리케이션은 파일 핸들을 사용하여 제2 가상 공간에 액세스하며, 호스트는 제2 액세스 요청을 생성한다.

- [0058] 호스트는, 제1 가상 공간과 제2 가상 공간 사이의 대응 관계에 따라, 제2 액세스 요청을, 제1 가상 공간에 액세스하는 제1 액세스 요청으로 변환한다. 구체적으로, 호스트는 제2 액세스 요청의 제2 가상 공간의 어드레스를 제1 가상 공간의 어드레스로 변환하고 제1 액세스 요청을 생성한다. 제1 가상 공간의 어드레스는 제1 저장 장치 노드의 장치 정보(예를 들어, ID 또는 ID와 노드 유형 모두)를 포함한다. 제1 저장 노드는 복수의 저장 노드 내에 있고 타깃 파일을 저장하도록 구성된 저장 노드이다.
- [0059] S250. 호스트는, 네트워크 장치가 제1 가상 공간의 어드레스 내에 있는 제1 저장 노드의 장치 정보에 따라 제1 액세스 요청을 제1 저장 노드에 라우팅할 수 있도록, 제1 액세스 요청을 네트워크 장치에 송신한다.
- [0060] 호스트는, 변환 방식으로 획득한 제1 액세스 요청을 네트워크 장치에 송신하고, 네트워크 장치는 제1 액세스 요청의 제1 가상 공간의 어드레스에 따라, 제1 액세스 요청을 제1 저장 노드에 라우팅한다. 예를 들어, 제1 가상 공간의 어드레스의 비트 수가 $2N$ 이고, 제1 가상 공간의 어드레스의 N 개의 최상위 비트가 제1 저장 노드의 장치 정보를 포함하는 경우, 네트워크 장치는, 제1 가상 공간의 어드레스의 N 개의 최상위 비트 내에 포함되고, 제1 저장 노드의 장치 정보에 따라 제1 액세스 요청을 제1 저장 노드에 라우팅한다.
- [0061] 네트워크 장치에 의해 송신된 제1 액세스 요청을 수신한 후에, 제1 저장 노드는 제1 액세스 요청에 따라 타깃 파일에 액세스한다. 구체적으로 로컬 가상 어드레스 공간과, 파일을 저장하는 데 사용되는 로컬 물리적 저장 공간 사이의 매핑 관계가 제1 저장 노드에 저장되고, 제1 저장 노드는 액세스 요청의 목적지 어드레스에서의 로컬 가상 어드레스 공간의 어드레스에 따라, 타깃 파일을 저장하는 데 사용되는 로컬 물리적 저장 공간의 어드레스를 결정할 수 있다. 또한, 제1 저장 노드는, 로컬 물리적 저장 공간에 저장된 타깃 파일에 액세스하기 위해, 결정된 로컬 물리적 저장 공간의 어드레스에 따라 로컬 물리적 저장 공간에 액세스할 수 있다. 실제애플리케이션에서, 제1 저장 노드는 추가로, 제1 액세스 요청에 응답하는 액세스 응답을 네트워크 장치에 송신할 수 있다. 네트워크 장치는 제1 저장 노드의 액세스 응답을 호스트에 송신한다.
- [0062] 본 발명의 본 실시예에서의 파일 액세스 방법에 따르면, 호스트의 커널을 사용하지 않고, 직접 액세스 어드레스(directly-accessed address)를 협상하지 않고, 분산 저장 시스템 내의 파일이 배치된 물리적 저장 공간은 직접 액세스 될 수 있다. 따라서, 지연은 감소되고 파일 액세스 효율은 개선될 수 있다.
- [0063] 따라서, 본 발명의 본 실시예에서의 파일 액세스 방법에 따르면, 분산 저장 시스템 내의 파일이 액세스 되는 경우, 호스트의 로컬 가상 어드레스 공간에 제2 가상 공간을 파일에 할당하고, 제2 가상 공간에 액세스하는 액세스 요청을, 분산 저장 시스템에 있고 파일의 제1 가상 공간에 액세스하는 제1 액세스 요구로 변환하고, 제1 액세스 요청은 네트워크로 송신된다. 이런 방식으로, 분산 저장 시스템에서 파일이 위치한 물리적 저장 공간은, 호스트의 커널을 사용하지 않고 직접 액세스 될 수 있으므로, 지연을 줄이고 파일 액세스 효율을 향상시킬 수 있다.
- [0064] 도 3은 본 발명의 일 실시예에 따른 파일 공간의 개략도이다. 도 3에 도시된 바와 같이, 총 M 개의 저장 노드가 분산 저장 시스템에 있고, 분산 저장 시스템 내의 파일은 분산 저장 시스템의 글로벌 가상 공간 내의 제1 가상 공간을 점유하며, 제1 가상 공간은, 제1 저장 노드의 로컬 가상 어드레스 공간의 섹션에 대응하고 로컬 물리적 저장 공간(즉, NVM 페이지와 같은 물리적 페이지)에 매핑된다. 파일을 열 때, 호스트는, 호스트의 로컬 가상 주소 공간 내에서 제2 가상 공간을 파일에 할당하고, 제1 가상 공간과 제2 가상 공간 간의 대응 관계를 구축, 즉 제2 가상 공간이 제1 가상 공간에 대응하도록 한다. 따라서, 파일의 로컬 가상 어드레스 공간(제2 가상 공간)에 액세스할 때, 호스트는 파일의 글로벌 가상 어드레스 공간(제1 가상 공간)에 액세스하는 액세스 요구를 생성하고 네트워크 장치(예를 들어, 글로벌 가상 어드레스 공간에 기초한 스위치)에 송신할 수 있고, 네트워크 장치는, 분산 저장 시스템 내의 파일에 액세스하기 위해, 글로벌 가상 어드레스에 기초하여, 액세스 요청을 제1 저장 노드에 라우팅한다.
- [0065] 본 발명의 실시예에서의 구체적인 예는, 당업자가 본 발명의 실시예를 더 잘 이해할 수 있도록 하기 위한 것일 뿐이며, 본 발명의 실시예의 범위를 한정하는 것은 아니다.
- [0066] 진술 한 프로세스의 시퀀스 번호는 본 발명의 다양한 실시예에서 실행 시퀀스를 의미하지 않는다는 것을 또한 이해해야 한다. 프로세스의 실행 순서는 프로세스의 기능 및 내부 논리에 따라 결정되어야 하며, 본 발명의 실시예의 구현 프로세스에 대한 임의의 제한으로 해석되어서는 안된다.
- [0067] 본 발명의 실시예에 따른 파일 액세스 방법을 상세히 설명하고, 본 발명의 일 실시예에 따른 시스템 및 호스트

를 설명한다.

- [0068] 도 4는 본 발명의 일 실시예에 따른 시스템(400)의 개략적 블록도이다. 도 4에 도시된 바와 같이, 시스템(400)은 호스트(410) 및 분산 저장 시스템(420)을 포함한다.
- [0069] 분산 저장 시스템(420)은 관리 노드(421), 복수의 저장 노드(422), 및 네트워크 장치(423)를 포함한다.
- [0070] 분산 저장 시스템(420) 내의 각 저장 노드의 로컬 가상 어드레스 공간(422)은 분산 저장 시스템의 글로벌 가상 공간의 일부(420)에 대응한다.
- [0071] 호스트(410)는, 액세스할 타깃 파일의 제1 가상 공간에 관한 정보를 획득하고 - 제1 가상 공간은 관리 노드(421)에 의해, 글로벌 가상 어드레스 공간 내에서 타깃 파일에 할당된 공간임 -,
- [0072] 제1 가상 공간에 관한 정보에 따라, 호스트의 로컬 가상 어드레스 공간(410) 내에서, 제2 가상 공간을 타깃 파일에 할당하며,
- [0073] 제1 가상 공간과 제2 가상 공간 사이의 대응 관계를 구축하고,
- [0074] 제1 가상 공간과 제2 가상 공간 사이의 대응 관계에 따라, 2 가상 공간을 액세스하는 제2 액세스 요청을 1 가상 공간을 액세스하는 제1 액세스 요청으로 변환하고 - 제1 액세스 요청은 제1 가상 공간의 어드레스를 포함하고, 제1 가상 공간의 어드레스는 제1 저장 노드의 장치 정보를 포함하며, 제1 저장 노드 is 복수의 저장 노드 내에 있고 타깃 파일을 저장하도록 구성된 저장 노드임 -,
- [0075] 네트워크 장치(423)가 제1 가상 공간의 어드레스 내에 있는 제1 저장 노드의 장치 정보에 따라, 제1 액세스 요청을 제1 저장 노드에 라우팅할 수 있도록, 제1 액세스 요청을 네트워크 장치(423)로 송신하도록 구성된다.
- [0076] 본 발명의 본 실시예에서의 시스템에 따르면, 분산 저장 시스템 내의 파일이 액세스 되는 경우, 제2 가상 공간은 호스트의 로컬 가상 어드레스 공간 내에서 파일에 할당되고, 제2 가상 공간을 액세스하는 제2 액세스 요청은, 분산 저장 시스템 내에 있고, 파일의 제1 가상 공간을 액세스하는 제1 액세스 요청으로 변환되며, 제1 액세스 요청은 네트워크에 송신된다. 이런 방식으로, 분산 저장 시스템에서 파일이 위치한 물리적 저장 공간은, 호스트의 커널을 상용하지 않고 직접 액세스 될 수 있으므로, 지연을 줄이고 파일 액세스 효율을 향상시킬 수 있다.
- [0077] 본 발명의 본 실시예에서, 선택적으로, 호스트(410)는 구체적으로, 타깃 파일의 메타 데이터를 획득하고, 타깃 파일의 메타 데이터로부터 타깃 파일의 제1 가상 공간에 관한 정보를 획득하도록 구성된다.
- [0078] 본 발명의 본 실시예에서, 선택적으로, 제1 가상 공간에 관한 정보는 제1 가상 공간의 시작 어드레스 및 크기를 포함한다.
- [0079] 본 발명의 본 실시예에서, 호스트(410)는 구체적으로 제1 가상 공간과 제2 가상 공간 사이의 대응 관계인 이하:
- [0080] $LVA = (GVA - GVA_start) + LVA_start$ 를 획득하도록 구성되고, GVA_start 는 제1 가상 공간의 시작 어드레스이고, LVA_start 는 제2 가상 공간의 시작 어드레스이며, GVA 는 제1 가상 공간의 어드레스이고, LVA 는 제1 가상 공간의 어드레스에 대응하는 제2 가상 공간의 어드레스이다.
- [0081] 발명의 본 실시예에서, 선택적으로, 제1 가상 공간의 어드레스의 비트 수는 $2N$ 이고, N 은 제1 저장 노드의 로컬 가상 어드레스 공간의 어드레스의 비트 수이며, 제1 가상 공간의 어드레스의 N 개의 최상위 비트는 제1 저장 노드의 장치 정보를 포함하고, 제1 가상 공간의 어드레스의 N 개의 최하위 비트는, 제1 저장 노드에 있고 타깃 파일을 저장하는 데 사용되는 로컬 가상 어드레스 공간의 어드레스이다.
- [0082] 본 발명의 실시예에 따른 시스템(400)의 호스트(410) 및 분산 저장 시스템(420)은 상술한 방법 실시예에서의 호스트 및 분산 저장 시스템에 대응할 수 있으며, 호스트(410)는 전술한 방법의 대응하는 과정을 수행할 수 있다. 간결함을 위해, 상세한 설명은 여기에서 다시 기술하지 않는다.
- [0083] 도 5는 본 발명의 일 실시예에 따른 호스트(500)의 개략적 블록도이다. 도 5에 도시된 바와 같이, 호스트(500)는 처리 모듈(510) 및 송신 모듈(520)을 포함한다.
- [0084] 처리 모듈(510)은 액세스할 타깃 파일의 제1 가상 공간에 관한 정보를 획득하고 - 제1 가상 공간은, 분산 저장 시스템의 글로벌 가상 어드레스 공간 내에서 분산 저장 시스템 내의 관리 노드에 의해 타깃 파일에 할당된 공간이고, 분산 저장 시스템 내에서 각 저장 노드의 로컬 가상 어드레스 공간은 글로벌 가상 어드레스 공간의 일부

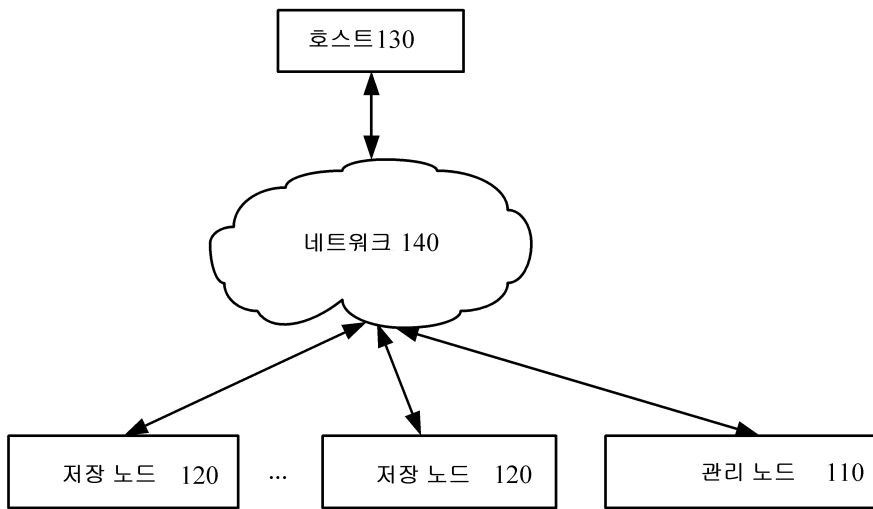
에 대응함 - ,

- [0085] 제1 가상 공간에 관한 정보에 따라, 호스트의 로컬 가상 어드레스 공간 내에서 제2 가상 공간을 타깃 파일에 할당하고,
- [0086] 제1 가상 공간과 제2 가상 공간 사이의 대응 관계를 구축하며,
- [0087] 제1 가상 공간과 제2 가상 공간 사이의 대응 관계에 따라, 제2 가상 공간을 액세스하는 제2 액세스 요청을 제1 가상 공간을 액세스하는 제1 액세스 요청으로 변환하도록 구성되고,
- [0088] 제1 액세스 요청은 제1 가상 공간의 어드레스를 포함하고, 제1 가상 공간의 어드레스는 제1 저장 노드의 장치 정보를 포함하며, 제1 저장 노드는 복수의 저장 노드 내에 있고 타깃 파일을 저장하도록 구성된 저장 노드이다.
- [0089] 송신 모듈(520)은, 네트워크 장치가 제1 가상 공간의 어드레스 내에 있는 제1 저장 노드의 장치 정보에 따라 제1 액세스 요청을 제1 저장 노드에 라우팅할 수 있도록, 제1 액세스 요청을 분산 저장 시스템 내의 네트워크 장치에 송신하도록 구성된다.
- [0090] 이러한 호스트에 따르면, 본 발명의 본 실시예에서, 분산 저장 시스템 내의 파일이 액세스 되는 경우, 제2 가상 공간은 호스트의 로컬 가상 어드레스 공간 내에서 파일에 할당되고, 제2 가상 공간을 액세스하는 제2 액세스 요청은, 분산 저장 시스템 내에 있고, 파일의 제1 가상 공간을 액세스하는 제1 액세스 요청으로 변환되며, 제1 액세스 요청은 네트워크에 송신된다. 이런 방식으로, 분산 저장 시스템에서 파일이 위치한 물리적 저장 공간은, 호스트의 커널을 상용하지 않고 직접 액세스 될 수 있으므로, 지연을 줄이고 파일 액세스 효율을 향상시킬 수 있다.
- [0091] 본 발명의 본 실시예에서, 선택적으로 처리 모듈(510)은 구체적으로 타깃 파일의 메타 데이터를 획득하고, 타깃 파일의 메타 데이터로부터 타깃 파일의 제1 가상 공간에 관한 정보를 획득하도록 구성된다.
- [0092] 발명의 본 실시예에서, 선택적으로, 제1 가상 공간에 관한 정보는 제1 가상 공간의 시작 어드레스 및 크기를 포함한다.
- [0093] 발명의 본 실시예에서, 선택적으로 처리 모듈(510)은 호스트의 로컬 가상 어드레스 공간 내에서 1 가상 공간과 동일한 크기를 가지는 제2 가상 공간을 타깃 파일에 할당하도록 구성된다.
- [0094] 발명의 본 실시예에서, 선택적으로 처리 모듈(510)은 이하:
- [0095] $LVA = (GVA - GVA_start) + LVA_start$ 의 제1 가상 공간과 제2 가상 공간 사이의 대응 관계를 구축하도록 구성되고, GVA_start 는 제1 가상 공간의 시작 어드레스이고, LVA_start 는 제2 가상 공간의 시작 어드레스이며, GVA 는 제1 가상 공간의 어드레스이고, LVA 는 제1 가상 공간의 어드레스에 대응하는 제2 가상 공간의 어드레스이다.
- [0096] 본 발명의 본 실시예에서, 선택적으로, 제1 가상 공간의 어드레스의 비트 수는 $2N$ 이고, N 은 제1 저장 노드의 로컬 가상 어드레스 공간의 어드레스의 비트 수이며, 제1 가상 공간의 어드레스의 N 개의 최상위 비트는 제1 저장 노드의 장치 정보를 포함하고, 제1 가상 공간의 어드레스의 N 개의 최하위 비트는, 제1 저장 노드에 있고 타깃 파일을 저장하는 데 사용되는 로컬 가상 어드레스 공간의 어드레스이다.
- [0097] 본 발명의 실시예에 따른 호스트(500)는 본 발명의 일 실시예에 따른 파일 액세스 방법(200)의 실행 주체에 대응할 수 있으며, 전술한 호스트(500) 내의 모듈의 동작 및/또는 기능은 전술한 방법의 대응하는 절차를 구현하기 위해 각각 사용된다. 간결함을 위해, 상세한 설명은 여기에서 다시 기술하지 않는다.
- [0098] 도 6은 본 발명의 일 실시예에 따른 호스트의 개략적 구조도이다. 호스트는 적어도 하나의 프로세서(602)(예를 들어, CPU), 적어도 하나의 네트워크 인터페이스(605), 또는 다른 통신 인터페이스, 메모리(606), 및 적어도 하나의 통신 버스(603)를 포함한다.
- [0099] 통신 버스(603)는 이 장치들 간의 접속 및 통신을 구현하도록 구성된다.
- [0100] 프로세서(602)는 메모리(606)에 저장된 실행 가능한 모듈, 예를 들어 컴퓨터 프로그램을 실행하도록 구성된다.
- [0101] 메모리(606)는 고속 RAM(Random Access Memory)를 포함하거나, NVM, 예를 들어 적어도 하나의 디스크 메모리를 더 포함할 수 있다.
- [0102] 호스트는 (유선 또는 무선일 수 있는)적어도 하나의 네트워크 인터페이스(605)를 사용하여 네트워크에 대한 통신 접속을 구현한다.

- [0103] 일부 구현 방식에서, 메모리(606)는 프로그램(6061)을 저장하고, 프로세서(602)는 상술한 방법 실시예에서의 모든 방법을 구현하기 위해 프로그램(6061)을 실행하도록 구성된다.
- [0104] 본 발명의 실시예에서 "및/또는"이라는 용어는 관련 객체를 설명하기 위한 연관 관계만을 기술하고 3개의 관계가 존재할 수 있음을 나타낸다는 것을 이해해야 한다. 예를 들어, A 및/또는 B는, A만 존재하고, A와 B가 모두 존재하며, B만 존재하는 세 가지 경우를 나타낼 수 있다. 또한, 이 명세서에서 문자 "/"는 일반적으로 관련 객체들 간의 "또는" 관계를 나타낸다.
- [0105] 당업자라면, 본 명세서에 개시된 실시예에 설명된 예들을 결합하여, 유닛 및 알고리즘 단계가 전자 하드웨어, 컴퓨터 소프트웨어, 또는 이들의 조합에 의해 구현될 수 있음을 인식할 수 있다. 하드웨어와 소프트웨어 간의 상호 호환성을 명확하게 설명하기 위해, 앞에서는 대체로 기능에 따라 각 예의 구성과 단계를 설명했다. 기능이 하드웨어 또는 소프트웨어로 수행되는지는 기술적 해결 수단의 구체적 애플리케이션 및 설계 제약 조건에 따라 다르다. 당업자는 각 구체적 애플리케이션에 대해 기술된 기능을 구현하기 위해 상이한 방법을 사용할 수 있지만, 이러한 구현이 본 발명의 범위를 벗어나는 것으로 간주되어서는 안된다.
- [0106] 상술한, 시스템, 장치 및 유닛의 상세한 동작 과정에 대한 편리하고 간단한 설명을 위해 당업자라면 본 발명의 범주에 속하는 대응 프로세스를 참조할 수 있음을 명확히 이해할 것이다. 전술한 방법 실시예 및 상세한 설명은 여기서 다시 설명하지 않는다.
- [0107] 본 발명에서 제공된 몇몇 실시예에서, 개시된 시스템, 장치 및 방법은 다른 방식으로 구현될 수 있다는 것을 이해해야 한다. 예를 들어, 설명된 장치 실시예는 단지 예일 뿐이다. 예를 들어, 유닛 부분은 논리적 기능 구분일 뿐이며 실제 구현에서는 다른 부분 일 수 있다. 예를 들어, 복수의 유닛 또는 구성 요소가 결합되거나 다른 시스템에 통합되거나, 일부 특징이 무시되거나 수행되지 않을 수 있다. 또한, 표시되거나 논의된 상호 연결 또는 직접 연결 또는 통신 연결은 일부 인터페이스를 통해 구현될 수 있다. 장치 또는 유닛 간의 간접 연결 또는 통신 연결은 전자적, 기계적 또는 다른 형태로 구현될 수 있다.
- [0108] 분리된 부분으로 설명된 유닛은 물리적으로 분리될 수도 있고 그렇지 않을 수도 있고, 유닛으로 표시된 부분은 물리적 유닛 일 수 있고 아닐 수도 있으며, 한 위치에 배치될 수 있거나 또는 복수의 네트워크 유닛들에 배치될 수 있다. 유닛의 일부 또는 전부는 본 발명의 실시예의 기술적 해결 수단의 목적을 달성하기 위해 실제 필요에 따라 선택될 수 있다.
- [0109] 또한, 본 발명의 실시예에서의 기능 유닛은 하나의 처리 유닛에 통합되어 있거나, 각 유닛이 물리적으로 단독으로 존재할 수 있고, 2 이상의 유닛이 하나의 유닛에 통합될 수 있다. 통합된 유닛은 하드웨어의 형태로 구현될 수 있거나 소프트웨어 기능 유닛의 형태로 구현될 수 있다.
- [0110] 통합 유닛이 소프트웨어 기능 유닛의 형태로 구현되고 독립 제품으로서 판매되거나 사용되는 경우, 통합 유닛은 컴퓨터 판독 가능 저장 매체에 저장될 수 있다. 이러한 이해에 기초하여, 본질적으로, 본 발명의 기술적 해결 수단 또는 종래 기술에 기여하는 부분 또는 기술적 해결 수단의 전부 또는 일부는 소프트웨어 제품의 형태로 구현될 수 있다. 소프트웨어 제품은 저장 매체에 저장되며, 실시예에서 설명된 방법의 단계의 전부 또는 일부를 수행하기 위해 컴퓨터 장치(개인용 컴퓨터, 서버, 또는 네트워크 장치 일 수 있음)를 명령하기 위한 몇 가지 명령을 포함한다. 본 발명의 저장 매체는 USB 플래시 드라이브, 탈착 식 하드 디스크, ROM(Read-Only Memory), RAM(Random Access Memory), 자기 디스크 또는 광디스크 일 수 있다. 전술한 설명은 단지 본 발명의 구체적 실시예에 불과하지만, 본 발명의 보호 범위를 제한하려는 것은 아니다.

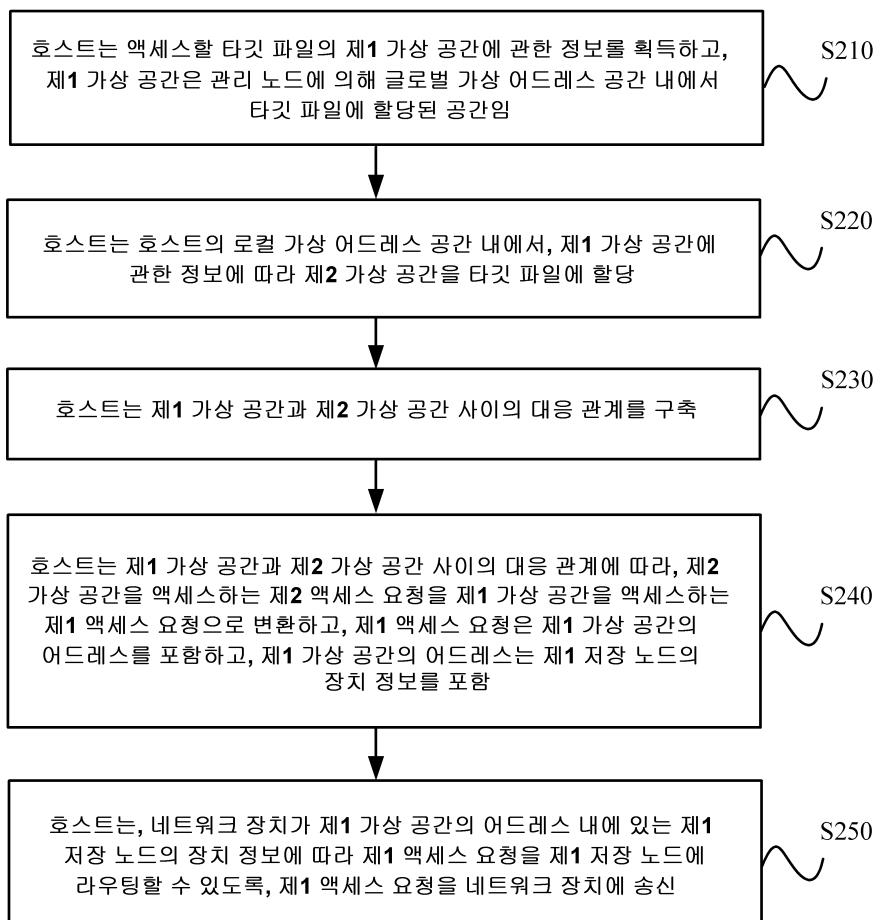
도면

도면1

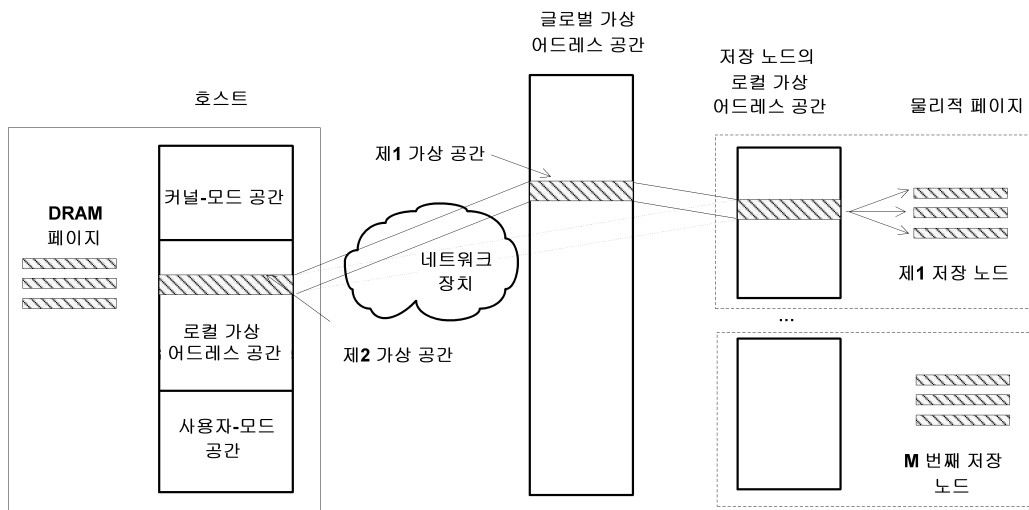


도면2

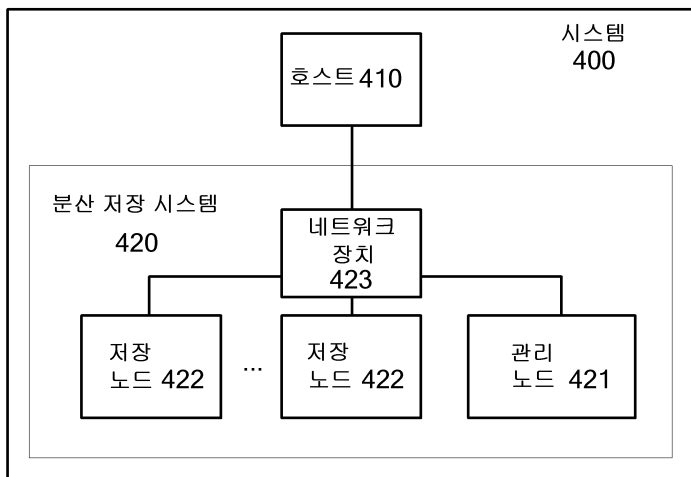
200



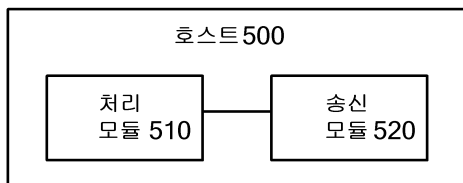
도면3



도면4



도면5



도면6

