

	(19) 대한민국특허청(KR) (12) 공개특허공보(A)	(11) 공개번호 10-2014-0056838 (43) 공개일자 2014년05월12일
(51) 국제특허분류(Int. Cl.) G06F 15/16 (2006.01) G06F 9/06 (2006.01)		(71) 출원인 서강대학교산학협력단 서울특별시 마포구 백범로 35 (신수동, 서강대학교)
(21) 출원번호 10-2012-0122978		(72) 발명자 박성용 서울 송파구 중대로 24, 229동 1103호 (문정동, 올림픽훼밀리타운)
(22) 출원일자 2012년11월01일 심사청구일자 2012년11월01일		이권용 서울특별시 마포구 신수동 167번지 메트로디오빌 2704호 김준영 서울 강북구 오현로9길 72, 401호 (미아동, 태양빌라)
		(74) 대리인 특허법인충현

전체 청구항 수 : 총 11 항

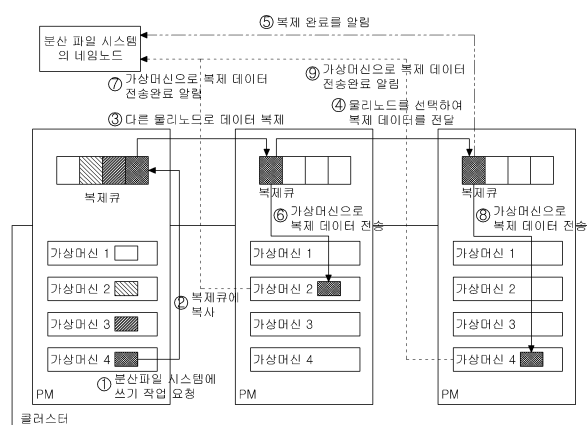
(54) 발명의 명칭 **가상 클라우드 환경에서 클러스터의 분산 파일 시스템 및 데이터 복제 방법**

(57) 요약

본 발명은 가상 클라우드 환경에서 클러스터의 분산 파일 시스템 및 데이터 복제 방법에 관한 것으로, 보다 구체적으로는 적어도 하나의 가상머신을 포함하는 클러스터의 각 가상머신의 물리노드에 존재하며, 분산 파일 시스템의 데이터 복제를 수행하는 노드로부터 복제 데이터를 수신한 후, 상기 가상머신이 탑재된 물리노드의 자원에 대한 가용율에 기초하여 수신한 상기 복제 데이터를 상기 가상머신으로 전송하는 드라이버 도메인;을 포함한다.

이러한 구성에 의해, 본 발명의 가상 클라우드 환경에서 클러스터의 분산 파일 시스템 및 데이터 복제 방법은 드라이버 도메인이 가상머신을 대신하여 복제 데이터를 수신한 후, 상기 가상머신이 탑재된 물리노드의 시스템에 대한 자원 가용율에 따라 수신한 복제 데이터를 가상머신으로 전달함으로써, 물리적 수준의 내고장성 정도를 일정하게 유지할 수 있고, 복제 오버헤드의 발생을 감소시킴에 따라, 분산 파일 시스템의 성능을 향상시킬 수 있는 효과가 있다.

대표도 - 도4



이 발명을 지원한 국가연구개발사업

과제고유번호 1415123192

부처명 지식경제부

연구사업명 대학 IT 연구센터 육성지원사업

연구과제명 소셜 미디어 서비스를 위한 클라우드 플랫폼 및 응용서비스 기술 개발

기 여 율 1/1

주관기관 건국대학교 산학협력단

연구기간 2012.01.01 ~ 2012.12.31

특허청구의 범위

청구항 1

가상 클라우드 환경에서 클러스터의 분산 파일 시스템에 있어서,

적어도 하나의 가상머신을 포함하는 클러스터의 물리노드에 존재하며, 분산 파일 시스템의 데이터 복제를 수행하는 노드로부터 복제 데이터를 수신한 후, 상기 가상머신이 탑재된 물리노드의 자원에 대한 가용율에 기초하여 수신한 상기 복제 데이터를 상기 가상머신으로 전송하는 드라이버 도메인;

을 포함하는 것을 특징으로 하는 가상 클라우드 환경에서 클러스터의 분산 파일 시스템.

청구항 2

제1항에 있어서,

상기 드라이버 도메인은

상기 가상머신이 탑재된 물리노드의 CPU 또는 메모리 중 적어도 하나의 자원에 대한 가용율이 기설정된 임계치보다 높은 경우에, 상기 분산 파일 시스템으로부터 수신한 복제 데이터를 자신이 속한 가상머신으로 전송하는 것을 특징으로 하는 가상 클라우드 환경에서 클러스터의 분산 파일 시스템.

청구항 3

제1항에 있어서,

상기 드라이버 도메인은

상기 분산 파일 시스템이 복제 대상 가상머신으로 전송한 복제 데이터를 가로채어 대신 수신하거나, 상기 분산 파일 시스템으로부터 복제 데이터를 직접 수신하는 것을 특징으로 하는 가상 클라우드 환경에서 클러스터의 분산 파일 시스템.

청구항 4

제3항에 있어서,

상기 드라이버 도메인은

상기 물리노드에 탑재되는 가상머신으로 전달되는 네트워크 패킷을 감시하는 패킷감시모듈;

복제대상 가상머신으로 전달되는 네트워크 패킷을 가로채어 복제 데이터를 대신 수신하는 데이터대행수신모듈;

수신한 상기 복제 데이터를 복제 큐(Queue)에 저장하도록 제어하는 저장제어모듈; 및

상기 네트워크 패킷에 대한 응답 패킷을 상기 분산 파일 시스템으로 전송하는 응답패킷전송모듈;

을 포함하는 것을 특징으로 하는 가상 클라우드 환경에서 클러스터의 분산 파일 시스템.

청구항 5

제4항에 있어서,

상기 데이터대행수신모듈은

상기 분산 파일 시스템에 상기 드라이버 도메인이 최종 목적지로 설정되어, 복제 데이터를 직접 수신하는 것을

더 포함하는 것을 특징으로 하는 가상 클라우드 환경에서 클러스터의 분산 파일 시스템.

청구항 6

가상 클라우드 환경에서 클러스터 내 분산 파일 시스템의 데이터 복제 방법 에 있어서,

상기 클러스터가 분산 파일 시스템으로 데이터 복제를 위한 쓰기 작업을 요청하는 단계;

상기 클러스터 내 포함된 가상머신에 탑재된 물리노드의 드라이버 도메인이 복제 데이터를 수신하여 복제 큐 (Queue)에 저장하는 단계;

상기 드라이버 도메인이 상기 가상머신이 탑재된 물리노드의 자원에 대한 가용율에 따라 상기 가상머신으로 상기 복제 데이터를 전송하는 단계; 및

상기 드라이버 도메인이 상기 가상머신으로의 복제 데이터 전송을 완료하면, 분산 파일 시스템에 이를 알리는 단계;

를 포함하는 것을 특징으로 하는 가상 클라우드 환경에서 클러스터 내 분산 파일 시스템의 데이터 복제 방법.

청구항 7

제6항에 있어서,

상기 드라이버 도메인이 복제 데이터를 수신하여 복제 큐에 저장하는 단계는

상기 분산 파일 시스템이 복제 대상 가상머신으로 전송한 복제 데이터를 가로채어 대신 수신하는 과정;

상기 분산 파일 시스템으로부터 복제 데이터를 직접 수신하는 과정;

중 적어도 하나의 과정을 수행하는 것을 특징으로 하는 가상 클라우드 환경에서 클러스터 내 분산 파일 시스템의 데이터 복제 방법.

청구항 8

제7항에 있어서,

상기 드라이버 도메인이 복제 데이터를 가로채어 대신 수신하는 과정은

상기 물리노드에 탑재되는 가상머신으로 전달되는 네트워크 패킷을 감시하는 과정;

복제 대상 가상머신으로 전달되는 복제 데이터의 네트워크 패킷을 가로채어 복제 데이터를 수신하는 과정;

수신한 상기 복제 데이터를 복제 큐에 저장하는 과정; 및

상기 네트워크 패킷에 대한 응답 패킷을 상기 분산 파일 시스템으로 전송하는 과정;

을 포함하는 것을 특징으로 하는 가상 클라우드 환경에서 클러스터 내 분산 파일 시스템의 데이터 복제 방법.

청구항 9

제7항에 있어서,

상기 분산 파일 시스템으로부터 복제 데이터를 직접 수신하는 과정은

상기 분산 파일 시스템에 상기 드라이버 도메인이 최종 목적지로 설정되어, 상기 분산 파일 시스템으로부터 복제 데이터를 직접 수신하는 것을 특징으로 하는 가상 클라우드 환경에서 클러스터 내 분산 파일 시스템의 데이터 복제 방법.

청구항 10

제6항에 있어서,

상기 드라이버 도메인이 가상머신으로 복제 데이터를 전송하는 단계는

상기 가상머신이 탑재된 물리노드의 CPU 또는 메모리 중 적어도 하나의 자원에 대한 가용율을 확인하는 과정;

상기 가용율을 기설정된 임계치와 비교하는 과정; 및

상기 가용율이 기설정된 임계치보다 높으면, 상기 분산 파일 시스템으로부터 수신한 복제 데이터를 상기 가상머신으로 전송하는 과정;

을 포함하는 것을 특징으로 하는 가상 클라우드 환경에서 클러스터 내 분산 파일 시스템의 데이터 복제 방법.

청구항 11

제6항 내지 제10항 중 어느 한 항에 따른 방법을 컴퓨터로 실행하기 위한 프로그램이 기록된 컴퓨터 판독가능 기록매체.

명세서

기술분야

- [0001] 본 발명은 가상 클라우드 환경에서 클러스터의 분산 파일 시스템 및 데이터 복제 방법에 관한 것으로, 특히 분산 파일 시스템의 쓰기 작업이 많은 클라우드 기반 클러스터의 성능을 향상시킬 수 있는 가상 클라우드 환경에서 클러스터의 분산 파일 시스템 및 데이터 복제 방법에 관한 것이다.

배경기술

- [0002] 각종 분야의 전산화로 인하여 데이터량이 급속도로 증가함에 따라, 데이터 병렬 또는 분산 처리를 위한 클러스터 컴퓨팅이 널리 사용되고 있으며, 이러한 클러스터 컴퓨팅 시스템은 다수의 물리 머신을 연결하여 이용하고 있다.
- [0003] 최근에는, 가상머신 기반의 클라우드 서비스들이 출현함에 따라, 기존의 클러스터 시스템들이 점차 클라우드 상으로 이동하고 있다.
- [0004] 이처럼, 클라우드 상에서 동작하는 클러스터 시스템은 관리의 유용성, 신뢰성, 가용성, 그리고 클러스터 구성의 용이성 등의 장점을 갖는다.
- [0005] 이러한 클러스터 시스템의 기반 기술 중 하나인 분산 파일 시스템은 데이터에 대한 빠른 접근과 내고장성(Fault tolerance)을 위해 데이터를 적절한 노드에 복제해둔다. 이러한 분산 파일 시스템의 데이터 복제는 네트워크를 통해 클러스터 상의 다른 노드로 복제 데이터를 전송하는 방식으로 수행된다. 클라우드 환경에서의 데이터 복제는 복제 시작 노드인 하나의 가상머신에서 복제 대상 노드인 다른 가상머신으로의 복제 데이터 전송으로 이루어지고, 가상화된 노드의 드라이버 도메인들을 거쳐 수행된다. 하지만 이처럼, 모든 복제 데이터가 드라이버 도메인을 거치게 되므로, 분산 파일 시스템의 쓰기 작업은 성능 저하의 문제점이 발생했다.
- [0006] 상술한 바와 같이, 가상 클라우드 환경에서 데이터 복제를 수행하는 클러스터 및 클러스터의 데이터 복제방법을 살펴보면 다음과 같다.
- [0007] 선행기술 1은 한국등록특허공보 제10-108551호(2011.01.10)로서, 클라우드 방식의 파일 복사 및 광역 통신망을 통한 디스크 복제 시스템 및 그 방법에 관한 것이다. 이러한 선행기술 1은 상기 서버가 서버 관리자로부터 이미지 소스 드라이브를 입력받고, 그 입력받은 이미지 소스 드라이브를 클라우드 디스크로 지정하고, IVP 서버 모듈에 디스크 관리자여권을 넘기도록 운영체제(OS)에 요청하는 인터럽트를 생성하는 클라우드 캐스트(Cloud Cast) 서버 모듈; 및 상기 인터럽트 요청을 받은 운영체제(OS)에 의해 램(RAM)에 로딩되고, 상기 운영체제(OS)로부터 디스크 관리 제어권을 넘겨받아 상기 클라우드 디스크를 포함하는 서버의 파일 시스템을 복제하고, 적어도 하나 이상의 클라이언트로부터 클라우드 디스크의 마운트 요청을 입력받고, 광역 통신망(WAN)을 통하여 상기

마운트를 요청한 클라이언트에 상기 복제된 파일 시스템을 전송하고, 상기 클라이언트로부터 파일 제공 요청을 대기하고 파일 제공요청이 입력되면 그 파일을 상기 클라이언트에 전송하는 IVP(Internet Virtual Partition) 서버 모듈;을 포함하고, 상기 적어도 하나 이상의 클라이언트는, 사용자로부터 서버의 클라우드 디스크를 마운트 할 드라이브 명과 마운트 요청을 입력받고 상기 IVP 서버 모듈에 클라우드 디스크의 마운트를 요청하고, 광역 통신망(WAN)을 통하여 서버의 복제된 파일 시스템을 전송받아 메모리에 로드하여, 상기 전송받은 복제된 파일 시스템의 클라우드 디스크의 드라이브 명을 상기 마운트할 드라이브 명으로 변경하여 가상 디스크를 생성하고, 자신 클라이언트의 원본 파일시스템에 상기 생성된 가상 디스크를 포함하여 자신 클라이언트의 사본 파일 시스템을 생성하는 클라우드 캐스트 클라이언트 모듈을 포함한다.

[0008] 또한, 선행기술 2는 한국등록특허공보 제10-1008557호(2011.01.10)로서, 클라우드 컴퓨팅을 이용한 디스크 순간 복구 시스템 및 그 방법에 관한 것이다.

[0009] 이러한 선행기술 2는 서버 관리자로부터 클라이언트와 공유할 클라우드 디스크로 지정될 이미지 소스 드라이브를 입력받는 클라우드 캐스트(Cloud Cast) 서버 모듈 및 적어도 하나 이상의 클라이언트로부터 상기 클라우드 디스크의 마운트 요청을 입력받고, 상기 클라우드 디스크를 포함하는 서버의 파일 시스템을 복제하여 복제된 파일 시스템을 상기 클라우드 디스크의 마운트를 요청한 클라이언트에 전송하고, 상기 클라이언트로부터 파일 제공 요청을 대기하고 파일 제공 요청이 입력되면 그 파일을 상기 클라이언트에 전송하고, 상기 클라이언트로부터 상기 클라우드 디스크의 재마운트 요청을 대기하고, 재마운트 요청이 입력되면 상기 복제된 파일 시스템을 다시 상기 클라이언트에 전송하는 IVP(Internet Virtual Partition) 서버 모듈을 포함하는 서버; 및 클라우드 캐스트 클라이언트 모듈을 포함하는 적어도 하나 이상의 클라이언트;를 포함하고, 상기 클라우드 캐스트 클라이언트 모듈은, 사용자로부터 서버의 클라우드 디스크를 마운트 할 드라이브 명과 마운트 요청을 입력받거나 또는, 사용자 또는 자신의 클라이언트의 운영체제(OS)로부터 재마운트 요청을 입력받고, 에이전트 프로그램(Agent Program)을 통하여 서버에 접속하여 IVP 서버 모듈에 상기 클라우드 디스크의 마운트 또는 재마운트를 요청하며, 상기 IVP 서버 모듈로부터 복제된 파일 시스템을 전송받아 메모리에 로드하여, 상기 전송받은 복제된 파일 시스템의 클라우드 디스크의 드라이브 명을 상기 마운트할 드라이브 명으로 변경하여 가상 디스크를 생성하고, 자신 클라이언트의 원본 파일시스템에 상기 생성된 가상 디스크를 포함하여 자신 클라이언트의 사본 파일 시스템을 생성한다.

발명의 내용

해결하려는 과제

[0010] 상기와 같은 종래 기술의 문제점을 해결하기 위해, 본 발명은 클러스터의 물리노드 내 드라이버 도메인이 복제 데이터를 가상머신 대신 수신한 후, 상기 가상머신이 탑재된 물리노드의 자원 가용율에 따라 수신한 복제 데이터를 가상머신으로 전송하여 데이터 복제로 인해 성능 병목 현상이 발생하던 분산 파일 시스템의 성능을 향상시키는 가상 클라우드 환경에서 클러스터의 분산 파일 시스템 및 데이터 복제 방법을 제공하고자 한다.

과제의 해결 수단

[0011] 위와 같은 과제를 해결하기 위한 본 발명의 한 실시 예에 따른 가상 클라우드 환경 내 데이터 복제를 수행하는 클러스터의 분산 파일 시스템은 적어도 하나의 가상머신을 포함하는 클러스터의 물리노드에 존재하며, 분산 파일 시스템의 데이터 복제를 수행하는 노드로부터 복제 데이터를 수신한 후, 상기 가상머신이 탑재된 물리노드의 자원에 대한 가용율에 기초하여 수신한 상기 복제 데이터를 상기 가상머신으로 전송하는 드라이버 도메인;을 포함하는 것을 특징으로 한다.

[0012] 보다 바람직하게는 상기 가상머신이 탑재된 물리노드의 CPU 또는 메모리 중 적어도 하나의 자원에 대한 가용율이 기설정된 임계치 보다 높은 경우에, 상기 분산 파일 시스템의 복제 시작 가상머신으로부터 수신한 복제 데이터를 복제 대상 가상머신으로 전송하는 드라이버 도메인을 포함할 수 있다.

[0013] 보다 바람직하게는 상기 분산 파일 시스템의 복제 시작 가상머신이 복제 대상 가상머신으로 전송한 복제 데이터를 가로채어 수신하거나, 상기 분산 파일 시스템의 복제 시작 가상머신으로부터 복제 데이터를 직접 수신하는 드라이버 도메인을 포함할 수 있다.

- [0014] 보다 바람직하게는 상기 물리노드에 탑재된 가상머신으로 전달되는 네트워크 패킷을 감시하는 패킷감시모듈; 복제대상 가상머신으로 전달되는 복제 데이터의 네트워크 패킷을 가로채어 대신 수신하는 데이터대행수신모듈; 수신한 상기 복제 데이터를 복제 큐(Queue)에 저장하도록 제어하는 저장제어모듈; 및 상기 네트워크 패킷에 대한 응답 패킷을 상기 분산 파일 시스템으로 전송하는 응답패킷전송모듈;을 포함하는 드라이버 도메인을 포함할 수 있다.
- [0015] 특히, 상기 분산 파일 시스템에 상기 드라이버 도메인이 최종 목적지로 설정되어, 복제 데이터를 직접 수신하는 데이터대행수신모듈을 더 포함할 수 있다.
- [0016] 위와 같은 과제를 해결하기 위한 본 발명의 한 실시 예에 따른 가상 클라우드 환경에서 클러스터의 분산 파일 시스템이 수행하는 데이터 복제 방법은 상기 클러스터가 분산 파일 시스템으로 데이터 복제를 위한 쓰기작업을 요청하는 단계; 상기 클러스터 내 포함된 가상머신 중 복제 대상 가상머신을 선택하는 단계; 상기 클러스터 내 포함된 가상머신이 탑재된 물리노드의 드라이버 도메인이 복제 데이터를 수신하여 복제 큐(Queue)에 저장하는 단계; 상기 드라이버 도메인이 상기 가상머신이 탑재된 물리노드의 자원에 대한 가용율에 따라 상기 가상머신으로 상기 복제 데이터를 전송하는 단계; 및 상기 드라이버 도메인이 상기 가상머신으로의 복제 데이터 전송을 완료하면, 분산 파일 시스템에 이를 알리는 단계;를 포함하는 것을 특징으로 한다.
- [0017] 보다 바람직하게는 상기 분산 파일 시스템이 복제 대상 가상머신으로 전송한 복제 데이터를 가로채어 수신하는 과정; 상기 분산 파일 시스템으로부터 복제 데이터를 직접 수신하는 과정; 중 적어도 하나의 과정을 수행하는 상기 드라이버 도메인이 복제 데이터를 수신하여 복제 큐에 저장하는 단계를 포함할 수 있다.
- [0018] 보다 바람직하게는 상기 물리노드에 탑재되는 가상머신으로 전달되는 네트워크 패킷을 감시하는 과정; 복제 대상 가상머신으로 전달되는 복제 데이터의 네트워크 패킷을 가로채어 대신 수신하는 과정; 수신한 상기 복제 데이터를 복제 큐에 저장하는 과정; 및 상기 네트워크 패킷에 대한 응답 패킷을 상기 분산 파일 시스템으로 전송하는 과정;을 포함하는 상기 드라이버 도메인이 복제 데이터를 가로채어 수신하는 과정을 포함할 수 있다.
- [0019] 특히, 상기 분산 파일 시스템에 상기 드라이버 도메인이 최종 목적지로 설정되어, 상기 분산 파일 시스템으로부터 복제 데이터를 직접 수신하는 상기 분산 파일 시스템으로부터 복제 데이터를 직접 수신하는 과정을 포함할 수 있다.
- [0020] 보다 바람직하게는 상기 가상머신이 탑재된 물리노드의 CPU 또는 메모리 중 적어도 하나의 자원에 대한 가용율을 확인하는 과정; 상기 가용율을 기설정된 임계치와 비교하는 과정; 및 상기 가용율이 기설정된 임계치보다 높으면, 상기 분산 파일 시스템으로부터 수신한 복제 데이터를 상기 가상머신으로 전송하는 과정;을 포함하는 상기 드라이버 도메인이 가상머신으로 복제 데이터를 전송하는 단계를 포함할 수 있다.

발명의 효과

- [0021] 본 발명의 가상 클라우드 환경에서 클러스터의 분산 파일 시스템 및 데이터 복제 방법은 드라이버 도메인이 복제 대상 가상머신을 대신하여 복제 데이터를 수신한 후, 상기 가상머신이 탑재된 물리노드의 시스템에 대한 자원 가용율에 따라 수신한 복제 데이터를 가상머신으로 전달함으로써, 물리적 수준의 내고장성 정도를 일정하게 유지할 수 있고, 복제 오버헤드의 발생을 감소시킴에 따라, 분산 파일 시스템의 성능을 향상시킬 수 있는 효과가 있다.

도면의 간단한 설명

- [0022] 도 1은 데이터 복제 매커니즘을 나타낸 도면이다.
- 도 2는 가상머신으로 구성된 클러스터의 복제 데이터 수에 따른 분산 파일 시스템의 쓰기 성능을 나타낸 그래프이다.
- 도 3은 가상머신 및 여러 드라이버 도메인으로 구성된 클러스터의 복제 데이터 수에 따른 분산 파일 시스템의 쓰기 성능을 나타낸 그래프이다.
- 도 4는 본 발명의 일 실시 예에 따른 가상 클라우드 환경에서 데이터 복제를 수행하는 클러스터 및 클러스터의

데이터 복제방법을 나타낸 도면이다.

발명을 실시하기 위한 구체적인 내용

- [0023] 이하, 본 발명을 바람직한 실시 예와 첨부한 도면을 참고로 하여 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자가 용이하게 실시할 수 있도록 상세히 설명한다. 그러나 본 발명은 여러 가지 상이한 형태로 구현될 수 있으며, 여기에서 설명하는 실시 예에 한정되는 것은 아니다.
- [0024] 클라우드 기반의 클러스터 시스템에서 사용되던 분산 파일 시스템의 기존 데이터 복제 기법을 사용하는 경우에는 분산 파일 시스템의 쓰기 작업으로 인해 가상 머신 간 데이터 복제 발생 시, 복제 데이터가 물리 노드에 도착하여 물리적 수준의 내고장성이 보장되더라도 복제 대상 가상 머신까지 복제 데이터가 도달해야 다음 작업을 수행한다.
- [0025] 도 1은 데이터 복제 매커니즘을 나타낸 도면이다.
- [0026] 도 1에 도시된 바와 같이, HDFS 클라이언트가 특정 파일을 만들어서(1~2) 쓰기를 수행하는 경우(3~4), 복제 데이터는 여러 데이터 노드에 걸쳐 파이프라인을 통해 전송되고 저장된다(5~6). 이러한 데이터 복제 과정이 끝난 후에야 비로소 해당 파일의 쓰기 작업이 완료(7)되고 분산 파일 시스템에서의 다음 작업을 수행하게 된다.
- [0027] 분산 파일 시스템에서 쓰기 작업으로 인한 데이터 복제가 필요한 경우에는, 파이프라인을 통해 복제 데이터가 병렬로 전송되지만, 전송 과정 중 오버헤드가 발생하게 된다. 또한 클라우드 기반의 클러스터 시스템에서는 클라우드를 구성하기 위한 가상화 기술의 오버헤드가 더해짐에 따라, 쓰기 작업이 빈번하게 일어나는 경우, 그 성능이 현저히 느려지는 현상이 발생한다. 이러한 현상은 복제 데이터가 이미 물리 노드에 도착하여 물리적 수준의 내고장성을 충족시켰음에도 불구하고 가상 머신까지 전달되기 위해 필요한 시간과 자원이 추가적으로 발생하기 때문이다.
- [0028] 도 2는 가상머신으로 구성된 클러스터의 복제 데이터 수에 따른 분산 파일 시스템의 쓰기 성능을 나타낸 그래프이다.
- [0029] 도 2에 도시된 바와 같이, 분산 파일 시스템에서 각 데이터에 대한 복제 데이터의 개수가 증가할수록 해당 데이터의 쓰기 작업에 걸리는 시간이 증가함을 알 수 있다.
- [0030] 도 3은 가상머신 및 여러 드라이버 도메인으로 구성된 클러스터의 복제 데이터 수에 따른 분산 파일 시스템의 쓰기 성능을 나타낸 그래프이다.
- [0031] 복제 대상 가상 머신까지 복제 데이터를 전달하지 않더라도 복제 데이터가 복제 대상 가상 머신이 탑재된 물리 노드에 도착하였을 때 이미 물리적 수준의 내고장성이 보장되므로 도 3에서 도시된 시스템의 성능 향상을 얻을 수 있게 된다.
- [0032] 종래의 가상 머신 간의 데이터 복제에서의 분산 파일 시스템은 물리 노드에 복제 데이터가 도착하여 물리적인 수준에서의 내고장성을 유지하더라도 복제 대상 가상 머신까지 복제 데이터가 도달해야만, 다음 수행해야 하는 작업단계로 넘어갈 수가 있어 쓰기 작업이 느려지는 문제점이 있었다.
- [0033] 이에 따라, 본 발명은 가상 머신들 사이에서 이루어지던 데이터 복제를 복제 대상 가상 머신이 아닌 가상 머신이 탑재된 물리 노드에 전송하는 것으로 물리적인 수준의 내고장성을 보장하고, 이후 물리 노드가 충분한 CPU나 메모리 등의 시스템 자원을 가질 때 복제 대상 가상 머신으로 복제 데이터를 전달하도록 한다.
- [0034] 또한 데이터 복제를 위한 대상 가상 머신을 선택할 때 물리 노드의 자원 가용성에 기반함에 따라, 클라우드 기반의 클러스터 시스템에서 동작하는 분산 파일 시스템의 보다 빠른 쓰기 작업이 이루어지도록 한다.
- [0035] 이하, 도 4를 참조하여 본 발명의 일 실시 예에 따른 가상 클라우드 환경에서 데이터 복제를 수행하는 클러스터에 대하여 자세히 살펴보도록 한다.
- [0036] 도 4는 본 발명의 일 실시 예에 따른 가상 클라우드 환경에서 클러스터의 분산 파일 시스템 및 데이터 복제 방법을 나타낸 도면이다.
- [0037] 도 4에 도시된 바와 같이, 본 발명에 따른 가상 클라우드 환경에서 데이터 복제를 수행하기 위해, 클러스터를 구성하는 적어도 하나의 가상머신은 각각의 물리노드에 드라이버 도메인을 포함한다. 이러한 드라이버 도메인은 분산 파일 시스템 또는 데이터 복제를 수행하는 노드로부터 복제 데이터를 수신한 후, 상기 가상머신이 탑재된 물리노드의 자원 즉, CPU 또는 메모리 등에 대한 가용율에 기초하여 수신한 상기 복제 데이터를 상기 가상머신

으로 전송한다. 특히, 상기 드라이버 도메인은 상기 가상머신이 탑재된 물리노드의 CPU 또는 메모리 중 적어도 하나의 자원에 대한 가용율이 기설정된 임계치 보다 높은 경우에, 상기 가상머신이 탑재된 물리노드의 시스템 자원에 여유가 존재한다고 판단하여, 상기 분산 파일 시스템으로부터 수신한 복제 데이터를 상기 가상머신으로 전달한다.

[0038] 이러한 드라이버 도메인은 패킷감시모듈, 데이터대행수신모듈, 저장제어모듈 및 응답패킷전송모듈을 포함한다.

[0039] 패킷감시모듈은 상기 물리노드에 탑재되는 가상머신으로 전달되는 네트워크 패킷을 감시한다.

[0040] 데이터대행수신모듈은 복제 대상 가상머신으로 전달되는 복제 데이터의 네트워크 패킷을 가로채어 대신 수신하거나, 상기 분산 파일 시스템에 상기 드라이버 도메인이 최종 목적으로 설정되어, 복제 데이터를 직접 수신할 수 있다.

[0041] 저장제어모듈은 수신한 상기 복제 데이터를 복제 큐(Queue)에 저장하도록 제어한다.

[0042] 응답패킷전송모듈은 상기 네트워크 패킷에 대한 응답 패킷을 상기 분산 파일 시스템으로 전송한다.

[0043] 이어서, 본 발명의 다른 실시 예에 따른 가상 클라우드 환경에서 클러스터의 분산 파일 시스템 및 데이터 복제 방법에 대하여 보다 자세히 살펴보도록 한다.

[0044] 도 4에 도시된 바와 같이, 클러스터의 클라이언트가 분산 파일 시스템에 쓰기 작업을 요청하여 특정 데이터를 작성하면(1) 해당 데이터는 클라이언트가 위치한 가상 머신이 탑재된 물리 노드의 드라이버 도메인에 위치하는 복제 큐(Queue)에 1차적으로 복제(2)된다. 그 후 복제 대상 가상 머신을 선택할 때, 다른 물리 노드들 중 복제 데이터를 빠르게 받아줄 수 있는 물리 노드를 선택하여 복제 데이터를 전달하고 복제 데이터를 받은 물리 노드의 드라이버 도메인은 받은 복제 데이터를 복제 큐에 저장(3)한다. 그리고 다음 복제 데이터를 전달할 물리노드를 선택하여 전달(4)한다. 마지막으로 복제 데이터를 받은 물리 노드의 드라이버 도메인은 복제 작업이 완료되었음을 알린다(5).

[0045] 이에 따라, 분산 파일 시스템은 다음 작업으로 넘어갈 수 있게 된다. 결국, 물리적 수준의 내고장성은 기존의 기법과 동일하게 유지하면서 보다 빠른 작업 처리를 수행할 수 있게 된다. 이후 복제 큐에 복제 데이터를 담고 있는 드라이버 도메인은 해당 물리 노드의 시스템 자원이 충분하여 다른 가상 머신들의 성능에 영향을 미치지 않고도 복제 데이터를 복제 대상 가상머신에 전달할 수 있을 때, 상기 복제 데이터를 가상머신으로 전달하고(6 또는 8), 해당 복제 데이터가 복제 대상 가상 머신으로 입력되었음을 분산 파일 시스템에 알린다(7 또는 9).

[0046] 각 물리 노드의 드라이버 도메인이 가상머신을 대신하여 복제 데이터를 수신하는 방법은 두 가지 방법으로 구현될 수 있다. 하나는 가상 머신들로 구성된 클러스터의 분산 파일 시스템이 드라이버 도메인의 존재를 모르는 경우로서 분산 파일 시스템은 복제 데이터를 복제 대상 가상 머신에게 전달하지만 드라이버 도메인이 이 복제 데이터를 중간에서 가로채는 방법이고, 다른 하나는 분산 파일 시스템이 드라이버 도메인의 존재를 알고 복제 데이터를 복제 대상 가상 머신이 탑재된 물리 노드의 드라이버 도메인에게 전달하는 방법이다.

[0047] 먼저, 드라이버 도메인이 복제 데이터를 가로채어 대신 수신하는 과정에 대하여 살펴보도록 한다. 데이터 복제에 사용되는 포트가 고정되어 있는 경우, 각 노드의 드라이버 도메인이 해당 물리 노드에 탑재된 가상 머신들로 전달되는 네트워크 패킷을 감시하고, 데이터 복제에 사용되는 포트를 통해 전달되는 패킷을 가로채어 복제 큐에 저장한 후, 데이터 복제나 응답 패킷 전달을 대신 수행한다. 이에 따라, 기존의 분산 파일 시스템이 수행하는 데이터 복제를 변경할 필요가 없게 된다.

[0048] 이와 달리, 드라이버 도메인이 분산 파일 시스템으로부터 복제 데이터를 직접 수신하는 과정에 대하여 살펴보도록 한다.

[0049] 분산 파일 시스템이 복제 대상 가상 머신을 선정 한 후, 해당 가상 머신이 탑재된 물리 노드의 드라이버 도메인을 데이터 복제의 목적으로 설정하여, 데이터 복제를 수행한다. 이에 따라, 데이터 복제에 사용되는 포트가 고정되어 있지 않거나, 다른 목적으로 해당 포트를 동시에 사용하는 경우에도 데이터 복제를 수행할 수 있다.

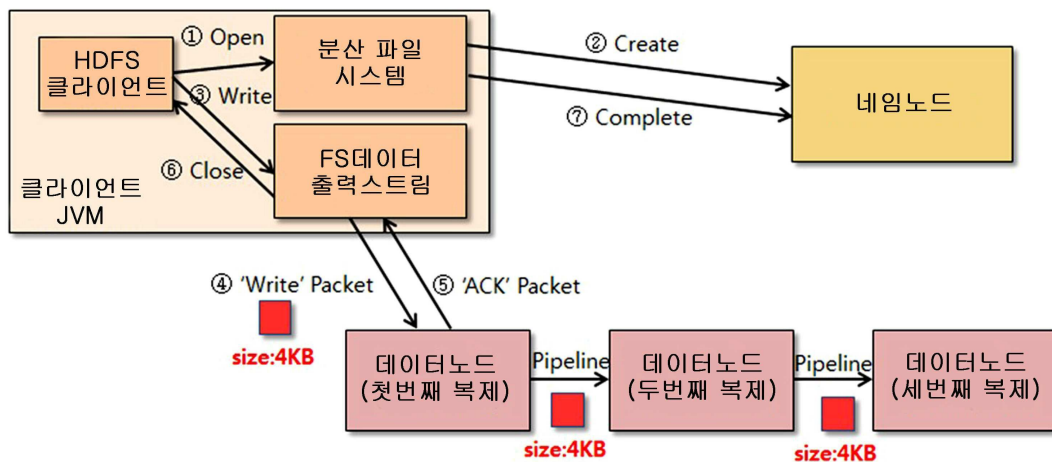
[0050] 이어서, 상기 드라이버 도메인이 상기 가상머신이 탑재된 물리노드의 자원에 대한 가용율에 따라 상기 가상머신으로 상기 복제 데이터를 전송한다. 이러한 복제 데이터 전송 과정은 먼저, 상기 가상머신이 탑재된 물리노드의 CPU 또는 메모리 중 적어도 하나의 자원에 대한 가용율을 확인한다.

[0051] 상기 가용율을 기설정된 임계치와 비교하여, 그 비교결과 상기 가용율이 기설정된 임계치보다 높으면, 상기 분산 파일 시스템으로부터 수신한 복제 데이터를 상기 가상머신으로 전송한다.

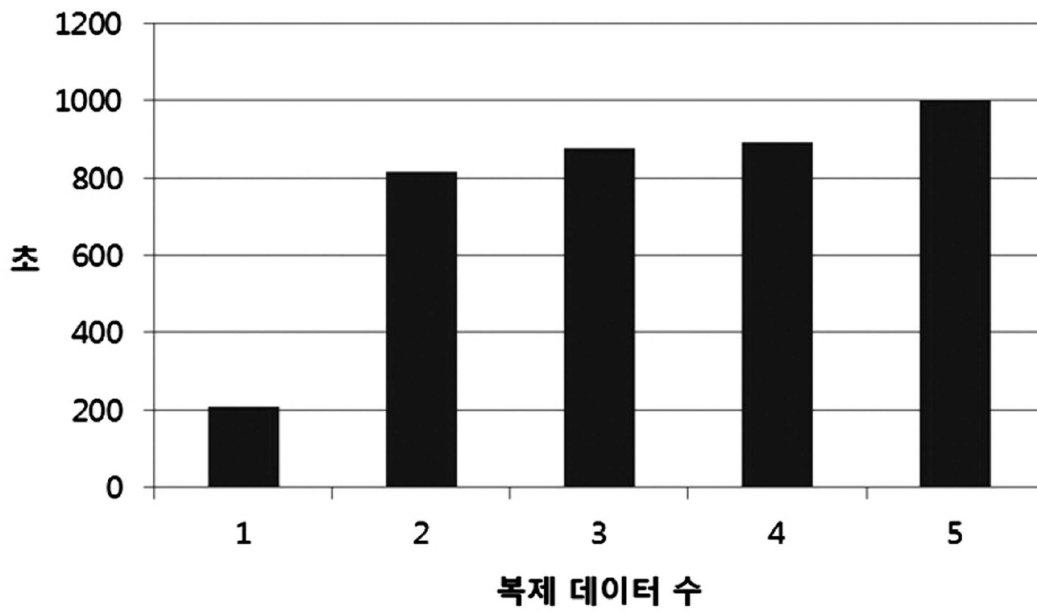
- [0052] 이와 같이, 상기 드라이버 도메인이 상기 가상머신으로의 복제 데이터 전송을 완료하면, 분산 파일 시스템에 이를 알린다.
- [0053] 본 발명을 통해 클라우드 상에서 가상머신들로 구성된 클러스터의 분산 파일 시스템이 내고장성 등을 위해 사용하는 데이터 복제 기법을 클라우드 환경을 고려하도록 설계됨으로써, 성능을 향상시킬 수 있게 되었다. 뿐만 아니라, 클라우드 환경을 고려하여 분산 파일 시스템의 작업을 물리적 수준의 내고장성은 유지하면서 보다 빠르게 수행할 수 있도록 각 노드의 드라이버 도메인에서 복제 데이터를 대신 수신하여, 추후 충분한 시스템 자원이 제공될 때 복제 대상 가상 머신에 복제 데이터를 전달하도록 하였으며 복제 대상 가상 머신을 선정할 때에는 가상 머신의 자원만이 아닌 물리 노드 전체의 시스템 자원을 고려하여 보다 빠르게 복제 데이터를 받을 수 있도록 한다. 실험 결과 가상 머신 인식 데이터 복제기법을 사용하여 클라우드 상에서 구성된 클러스터의 분산 파일 시스템이 처리하는 쓰기 작업 속도가 보다 빨라졌으며 물리적 수준의 내고장성도 동일하게 보장된다.
- [0054] 본 발명의 가상 클라우드 환경에서 클러스터의 분산 파일 시스템 및 데이터 복제 방법은 드라이버 도메인이 가상머신을 대신하여 복제 데이터를 수신한 후, 상기 가상머신의 시스템에 대한 자원 가용율에 따라 수신한 복제 데이터를 가상머신으로 전달함으로써, 물리적 수준의 내고장성 정도를 유지할 수 있고, 복제 오버헤드의 발생을 감소시킴에 따라, 분산 파일 시스템의 성능을 향상시킬 수 있는 효과가 있다.
- [0055] 상기에서는 본 발명의 바람직한 실시 예에 대하여 설명하였지만, 본 발명은 이에 한정되는 것이 아니고 본 발명의 기술 사상 범위 내에서 여러 가지로 변형하여 실시하는 것이 가능하고 이 또한 첨부된 특허청구범위에 속하는 것은 당연하다.

도면

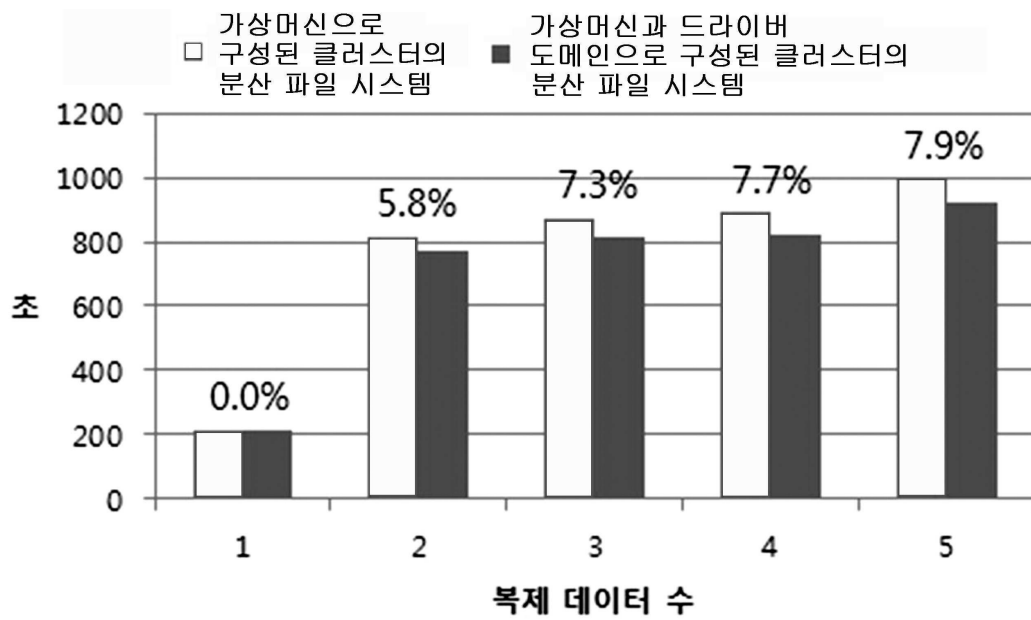
도면1



도면2



도면3



도면4

