



(19) 대한민국특허청(KR)

(12) 등록특허공보(B1)

(45) 공고일자 2016년01월14일

(11) 등록번호 10-1585146

(24) 등록일자 2016년01월07일

(51) 국제특허분류(Int. Cl.)  
**G06F 15/16** (2006.01) **H04L 9/32** (2006.01)  
 (21) 출원번호 10-2010-0134840  
 (22) 출원일자 2010년12월24일  
 심사청구일자 2013년08월06일  
 (65) 공개번호 10-2012-0072907  
 (43) 공개일자 2012년07월04일  
 (56) 선행기술조사문헌  
 KR1020020050160 A\*  
 KR1020110086114 A  
 KR1020100107479 A\*  
 JP2002318790 A  
 \*는 심사관에 의하여 인용된 문헌

(73) 특허권자  
**주식회사 케이티**  
 경기도 성남시 분당구 불정로 90(정자동)  
 (72) 발명자  
**김미점**  
 경기도 안양시 동안구 부림로 113, 현대아이스페이스 606호 (관양동)  
**박창식**  
 서울특별시 구로구 오리로 1197, 119동 205호 (오류동, 금강수목원아파트)  
 (뒷면에 계속)  
 (74) 대리인  
**리엔텍특허법인**

전체 청구항 수 : 총 27 항

심사관 : 경연정

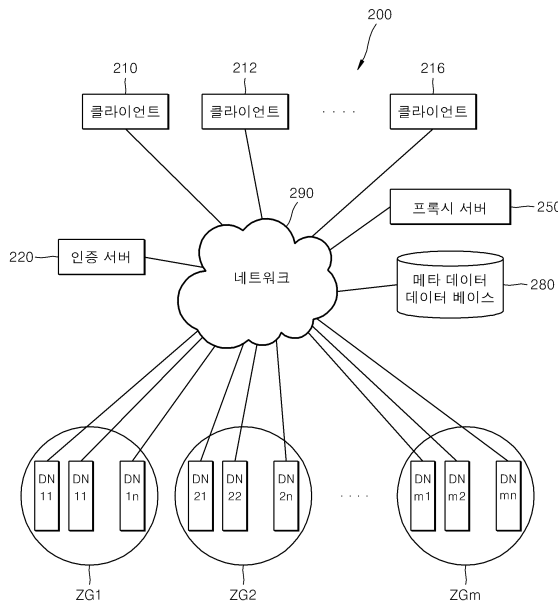
(54) 발명의 명칭 **오브젝트를 복수 개의 데이터 노드들의 위치에 기반하여 분산 저장하는 분산 저장 시스템 및 그 위치 기반 분산 저장 방법 및 컴퓨터에 의하여 독출 가능한 저장 매체**

**(57) 요약**

분산 저장 시스템 및 분산 저장 시스템에서 오브젝트를 분산 저장하기 위한 방법이 개시된다. 분산 저장 방법은 클라이언트를 인증하는 단계, 데이터 노드들을 각 데이터 노드들의 위치 정보를 기반으로 지역별로 그룹핑하여 지역 그룹을 생성하는 단계, 저장할 대상 오브젝트가 존재할 경우, 대상 오브젝트가 저장될 대상 데이터 노드가

(뒷면에 계속)

**대표도** - 도2



속하는 지역 그룹을 결정하는 지역 그룹 결정 단계, 동일한 지역 그룹에 속하는 데이터 노드들 중 하나의 데이터 노드를 대상 데이터 노드로서 결정하는 대상 데이터 노드 결정 단계, 대상 데이터 노드에 대상 오브젝트를 저장하는 저장 단계, 및 대상 오브젝트의 고유 정보 및 대상 오브젝트가 저장된 데이터 노드의 고유 정보를 이용하여 메타데이터를 갱신하는 메타데이터 갱신 단계를 포함한다. 본 발명에 의하여, 데이터 노드들을 그룹핑하기 때문에 실제 복제본들이 안전하게 지역적으로 분산 저장되어 하나의 지역에 문제가 발생할 경우에도 신뢰성이 향상된다. 또한, 클라이언트의 위치와 물리적으로 가까운 데이터 노드에 오브젝트를 배치함으로써 오브젝트의 저장과 조회 지연을 최소화할 수 있다.

(72) 발명자

**이어형**

경기 용인시 수지구 현암로125번길 11, 723동 102호 (죽전동, 새터마을죽전힐스테이트)

**황진경**

경기도 과천시 관문로 143, 삼성레미안아파트 1108동 705호 (중앙동)

**명세서**

**청구범위**

**청구항 1**

복수 개의 클라이언트들로부터 네트워크를 통해 전송되는 오브젝트를 복수 개의 데이터 노드들에 분산 저장하는 분산 저장 시스템(distribution storage system)에 있어서,

상기 클라이언트를 인증하기 위한 인증 서버;

지역(zone)별로 그룹핑되며 각각 적어도 하나의 오브젝트들을 저장하는 복수 개의 데이터 노드들로서, 동일한 오브젝트는 동일한 지역 그룹(zone group)에 속한 데이터 노드들 중 오직 하나에만 저장되도록 구성되는 복수 개의 데이터 노드들;

상기 오브젝트의 고유 정보 및 상기 오브젝트가 저장된 데이터 노드의 고유 정보를 포함하는 메타데이터를 저장하는 메타데이터 데이터베이스; 및

인증된 클라이언트의 질의에 응답하여, 상기 메타데이터를 참조하여 상기 클라이언트의 관리 동작의 객체인 대상 오브젝트(target object)와 관련되는 대상 데이터 노드(target data node)의 고유 정보의 목록을 상기 클라이언트에게 제공하는 프락시 서버를 포함하며,

상기 클라이언트는, 상기 대상 데이터 노드의 고유 정보의 목록을 참조하여 상기 대상 오브젝트의 관리 동작(management process)을 수행하도록 구성되고,

상기 메타데이터는, 상기 데이터 노드들의 사용량, 각 지역 그룹에 속한 데이터 노드들의 목록, 대상 오브젝트에 대한 지역 그룹별 우선 순위, 및 동일한 지역 그룹에 속한 데이터 노드들 간의 우선 순위 중 적어도 하나를 더 포함하는 것을 특징으로 하는 분산 저장 시스템.

**청구항 2**

제1항에 있어서,

상기 관리 동작은 상기 대상 오브젝트를 상기 데이터 노드에 저장하는 저장 동작(storage process)이며,

상기 저장 동작을 수행하기 위하여, 상기 클라이언트는,

상기 대상 오브젝트를 저장할 대상 데이터 노드들이 속하는 지역 그룹들 및 각 지역 그룹 당 하나의 대상 데이터 노드를 선택하고,

선택된 지역 그룹들에 대한 상기 클라이언트의 선호도(preference)에 의하여 결정되는 우선 순위(priority)를 상기 지역 그룹에게 부여하며,

가장 높은 우선 순위를 가지는 지역 그룹에 속하는 대상 데이터 노드에 상기 대상 오브젝트를 저장함으로써, 상기 우선 순위에 따라 순차적으로 더 낮은 우선 순위를 가지는 지역 그룹에 속하는 대상 데이터 노드들에 상기 대상 오브젝트가 복제되는 복제 동작(replication process)이 수행되도록 야기하도록 더욱 구성되고, 상기 프락시 서버는,

상기 대상 오브젝트가 저장된 대상 데이터 노드들의 목록을 상기 클라이언트에게 질의(query)하고,

상기 클라이언트에 의하여 상기 대상 데이터 노드들의 목록이 회신되면, 회신된 목록을 이용하여 상기 메타데이터 데이터베이스를 갱신하도록 구성되는 것을 특징으로 하는 분산 저장 시스템.

**청구항 3**

제2항에 있어서,

상기 관리 동작은 선택된 지역 그룹들을 변경하기 위한 지역 그룹 변경 동작(zone group change process)이며,

상기 지역 그룹 변경 동작을 수행하기 위하여, 상기 클라이언트는,

상기 프락시 서버에게 지역 그룹 변경 요청(zone group change request)을 송신한 후 상기 지역 그룹 변경 요청이 수락되면, 상기 대상 오브젝트를 저장할 새로운 대상 데이터 노드들이 속하는 지역 그룹들 및 각 지역 그룹 당 하나의 대상 데이터 노드를 선택하고,

선택된 지역 그룹들에 대한 상기 클라이언트의 선호도에 의하여 결정되는 우선 순위를 상기 지역 그룹에게 부여하며,

가장 높은 우선 순위를 가지는 지역 그룹에 속하는 대상 데이터 노드에 상기 대상 오브젝트를 저장함으로써, 상기 우선 순위에 따라 순차적으로 더 낮은 우선 순위를 가지는 지역 그룹에 속하는 대상 데이터 노드들에 상기 대상 오브젝트가 복제되는 복제 동작이 수행되도록 야기하도록 더욱 구성되고, 상기 프락시 서버는,

저장된 오브젝트와 관련된 대상 데이터 노드들의 목록을 상기 메타데이터 데이터베이스에서 삭제하고,

상기 클라이언트가 선택한 새로운 지역 그룹에 속하는 대상 데이터 노드들의 목록을 상기 클라이언트에게 질의하며,

상기 클라이언트에 의하여 새로운 지역 그룹에 속하는 대상 데이터 노드들의 목록이 회신되면, 회신된 목록을 이용하여 상기 메타데이터 데이터베이스를 갱신하도록 더욱 구성되는 것을 특징으로 하는 분산 저장 시스템.

**청구항 4**

제1항에 있어서,

상기 관리 동작은 상기 대상 오브젝트를 상기 데이터 노드에 저장하는 저장 동작이며,

상기 저장 동작을 수행하기 위하여, 상기 클라이언트는,

상기 대상 오브젝트를 저장할 대상 데이터 노드들이 속하는 지역 그룹들을 선택하고,

선택된 지역 그룹들에 대한 상기 클라이언트의 선호도에 의하여 결정되는 우선 순위를 상기 지역 그룹에게 부여하며,

상기 지역 그룹 당 하나의 대상 데이터 노드의 목록을 상기 프락시 서버로부터 수신하고,

가장 높은 우선 순위를 가지는 지역 그룹에 속하는 대상 데이터 노드에 상기 대상 오브젝트를 저장함으로써, 상기 우선 순위에 따라 순차적으로 더 낮은 우선 순위를 가지는 지역 그룹에 속하는 대상 데이터 노드들에 상기 대상 오브젝트가 복제되는 복제 동작이 수행되도록 야기하도록 더욱 구성되고, 상기 프락시 서버는,

상기 클라이언트가 선택한 지역 그룹들 각각에 포함되는 데이터 노드들 중 하나의 데이터 노드를 대상 데이터 노드로서 결정하여 상기 클라이언트로 전송하고,

상기 대상 데이터 노드들의 목록을 이용하여 상기 메타데이터 데이터베이스를 갱신하도록 구성되는 것을 특징으로 하는 분산 저장 시스템.

**청구항 5**

제4항에 있어서, 상기 프락시 서버는,

동일한 지역 그룹에 포함되는 데이터 노드들의 가용 저장 용량 및 오브젝트 저장 내역을 고려하여 동일한 지역 그룹에 포함되는 데이터 노드들에게 우선 순위를 부여하고,

동일한 지역 그룹에 포함되는 데이터 노드들 중에서 가장 높은 우선 순위를 가지는 데이터 노드를 상기 대상 데이터 노드로서 결정하도록 더욱 구성되는 것을 특징으로 하는 분산 저장 시스템.

**청구항 6**

제5항에 있어서,

상기 관리 동작은 선택된 지역 그룹들을 변경하기 위한 지역 그룹 변경 동작이며,

상기 지역 그룹 변경 동작을 수행하기 위하여, 상기 클라이언트는,

상기 프락시 서버에게 지역 그룹 변경 요청을 송신한 후 상기 지역 그룹 변경 요청이 수락되면, 상기 대상 오브

젝트를 저장할 새로운 대상 데이터 노드들이 속하는 지역 그룹을 선택하고,  
 선택된 지역 그룹들에 대한 상기 클라이언트의 선호도에 의하여 결정되는 우선 순위를 상기 지역 그룹에게 부여하며,  
 상기 지역 그룹 당 하나의 대상 데이터 노드의 목록을 상기 프락시 서버로부터 수신하고,  
 가장 높은 우선 순위를 가지는 지역 그룹에 속하는 대상 데이터 노드에 상기 대상 오브젝트를 저장함으로써, 상기 우선 순위에 따라 순차적으로 더 낮은 우선 순위를 가지는 지역 그룹에 속하는 대상 데이터 노드들에 상기 대상 오브젝트가 복제되는 복제 동작이 수행되도록 야기하도록 더욱 구성되고, 상기 프락시 서버는,  
 저장된 오브젝트와 관련된 대상 데이터 노드들의 목록을 상기 메타데이터 데이터베이스에서 삭제하고,  
 상기 클라이언트가 선택한 지역 그룹들 각각에 포함되는 데이터 노드들 중 하나의 데이터 노드를 대상 데이터 노드로서 결정하여 상기 클라이언트로 전송하며,  
 상기 대상 데이터 노드들의 목록을 이용하여 상기 메타데이터 데이터베이스를 갱신하도록 더욱 구성되는 것을 특징으로 하는 분산 저장 시스템.

**청구항 7**

제6항에 있어서, 상기 프락시 서버는,  
 동일한 지역 그룹에 포함되는 데이터 노드들의 가용 저장 용량 및 오브젝트 저장 내역을 고려하여 동일한 지역 그룹에 포함되는 데이터 노드들에게 우선 순위를 부여하고,  
 동일한 지역 그룹에 포함되는 데이터 노드들 중에서 가장 높은 우선 순위를 가지는 데이터 노드를 상기 대상 데이터 노드로서 결정하도록 더욱 구성되는 것을 특징으로 하는 분산 저장 시스템.

**청구항 8**

제1항에 있어서,  
 상기 관리 동작은 상기 대상 오브젝트를 상기 데이터 노드에 저장하는 저장 동작이며,  
 상기 분산 저장 시스템은, 상기 데이터 노드들 및 상기 클라이언트의 위치 관계에 기반하여 상기 대상 오브젝트를 저장할 대상 데이터 노드가 속하는 지역 그룹을 선택하고 선택된 지역 그룹 및 상기 클라이언트 간의 거리에 기반하여 상기 지역 그룹별 우선 순위를 결정하는 위치 인식 서버(location-aware server)를 더 포함하고,  
 상기 저장 동작을 수행하기 위하여, 상기 프락시 서버는,  
 상기 위치 인식 서버가 선택한 지역 그룹 당 하나의 대상 데이터 노드를 결정하고,  
 결정된 대상 데이터 노드들의 목록을 이용하여 상기 메타데이터 데이터베이스를 갱신하며,  
 상기 대상 데이터 노드들의 목록 및 상기 지역 그룹별 우선 순위를 상기 클라이언트에게 전송하도록 구성되고,

상기 클라이언트는, 가장 높은 지역 그룹별 우선 순위를 가지는 지역 그룹에 속하는 대상 데이터 노드에 상기 대상 오브젝트를 저장함으로써, 상기 우선 순위에 따라 순차적으로 더 낮은 우선 순위를 가지는 지역 그룹에 속하는 대상 데이터 노드들에 상기 대상 오브젝트가 복제되는 복제 동작이 수행되도록 야기하도록 더욱 구성되는 것을 특징으로 하는 분산 저장 시스템.

**청구항 9**

제8항에 있어서, 상기 프락시 서버는,  
 동일한 지역 그룹에 포함되는 데이터 노드들의 가용 저장 용량 및 오브젝트 저장 내역을 고려하여 동일한 지역 그룹에 포함되는 데이터 노드들에게 우선 순위를 부여하고,  
 가장 높은 우선 순위를 가지는 데이터 노드를 상기 대상 데이터 노드로서 결정하도록 더욱 구성되는 것을 특징으로 하는 분산 저장 시스템.

**청구항 10**

제2항 내지 제9항 중 어느 한 항에 있어서,  
 상기 복제 동작을 수행하기 위하여, 상기 클라이언트는,  
 가장 높은 우선 순위보다 낮은 우선 순위를 가지는 모든 지역 그룹들에 속하는 대상 데이터 노드에 상기 대상 오브젝트를 저장하도록 더욱 구성되는 것을 특징으로 하는 분산 저장 시스템.

**청구항 11**

제2항 내지 제9항 중 어느 한 항에 있어서,  
 상기 복제 동작을 수행하기 위하여, 상기 대상 데이터 노드들은,  
 자신이 속한 지역 그룹의 우선 순위보다 한 단계 높은 우선 순위를 가지는 지역 그룹에 속하는 대상 데이터 노드로부터 상기 대상 오브젝트를 수신하여 저장하도록 구성되는 것을 특징으로 하는 분산 저장 시스템.

**청구항 12**

제2항 내지 제9항 중 어느 한 항에 있어서,  
 상기 분산 저장 시스템은 가장 높은 우선 순위를 가지는 지역 그룹에 속하는 대상 데이터 노드에 저장된 대상 오브젝트를 독출하고, 독출된 대상 오브젝트를 가장 높은 우선 순위보다 낮은 우선 순위를 가지는 모든 지역 그룹들에 속하는 대상 데이터 노드에 저장하는 복제 서버(replication server)를 더 포함하며,  
 상기 복제 동작은 상기 복제 서버에 의하여 수행되는 것을 특징으로 하는 분산 저장 시스템.

**청구항 13**

제1항에 있어서,  
 상기 오브젝트의 고유 정보는 상기 오브젝트의 ID, 크기, 데이터 타입, 및 작성자 중 적어도 하나를 포함하고,  
 상기 데이터 노드의 고유 정보는 상기 데이터 노드의 ID, IP(Internet Protocol) 주소, 및 물리적 위치 중 적어도 하나를 포함하는 것을 특징으로 하는 분산 저장 시스템.

**청구항 14**

삭제

**청구항 15**

복수 개의 클라이언트들로부터 네트워크를 통해 전송되는 오브젝트를 복수 개의 데이터 노드들에 분산 저장하는 분산 저장 시스템에서 오브젝트를 분산 저장하기 위한 방법에 있어서,

상기 클라이언트를 인증하는 단계;

상기 데이터 노드들을 각 데이터 노드들의 위치 정보를 기반으로 지역별로 그룹핑하여 지역 그룹을 생성하는 단계;

저장할 대상 오브젝트가 존재할 경우, 상기 대상 오브젝트가 저장될 대상 데이터 노드가 속하는 지역 그룹을 결정하는 지역 그룹 결정 단계;

동일한 지역 그룹에 속하는 데이터 노드들 중 하나의 데이터 노드를 상기 대상 데이터 노드로서 결정하는 대상 데이터 노드 결정 단계;

상기 대상 데이터 노드에 상기 대상 오브젝트를 저장하는 저장 단계; 및

상기 대상 오브젝트의 고유 정보 및 상기 대상 오브젝트가 저장된 데이터 노드의 고유 정보를 이용하여 메타데이터를 갱신하는 메타데이터 갱신 단계를 포함하고, 상기 메타데이터는, 상기 데이터 노드들의 사용량, 각 지역 그룹에 속한 데이터 노드들의 목록, 대상 오브젝트에 대한 지역 그룹별 우선 순위, 및 동일한 지역 그룹에 속한 데이터 노드들 간의 우선 순위 중 적어도 하나를 더 포함하는 것을 특징으로 하는 분산 저장 시스템에서 오브젝

트를 분산 저장하기 위한 방법.

**청구항 16**

제15항에 있어서,

상기 지역 그룹 결정 단계는, 상기 클라이언트가 상기 대상 오브젝트를 저장할 대상 데이터 노드들이 속하는 지역 그룹들을 선택하고, 선택된 지역 그룹들에 대한 상기 클라이언트의 선호도에 의하여 결정되는 우선 순위를 상기 지역 그룹에게 부여하는 단계를 포함하고,

상기 대상 데이터 노드 결정 단계는, 상기 클라이언트가 동일한 지역 그룹에 속하는 데이터 노드들 중 하나의 데이터 노드를 상기 대상 데이터 노드로서 결정하는 단계를 포함하며,

상기 저장 단계는,

상기 클라이언트가 가장 높은 우선 순위를 가지는 지역 그룹에 속하는 대상 데이터 노드에 상기 대상 오브젝트를 저장하는 단계; 및

상기 우선 순위에 따라 순차적으로 더 낮은 우선 순위를 가지는 지역 그룹에 속하는 대상 데이터 노드들에 상기 대상 오브젝트를 복제하는 복제 단계를 포함하고,

상기 메타데이터 갱신 단계는, 상기 클라이언트에 의하여 상기 대상 데이터 노드들의 목록이 회신되면, 프락시 서버가 회신된 목록을 이용하여 상기 메타데이터 데이터베이스를 갱신하는 단계를 포함하는 것을 특징으로 하는 분산 저장 시스템에서 오브젝트를 분산 저장하기 위한 방법.

**청구항 17**

제16항에 있어서,

선택된 지역 그룹들을 변경하기 위한 지역 그룹 변경 단계를 더 포함하며,

상기 지역 그룹 변경 단계는,

상기 클라이언트가 상기 프락시 서버에게 지역 그룹 변경 요청을 송신하는 단계;

상기 프락시 서버가 상기 지역 그룹 변경 요청을 수락하고 저장된 오브젝트와 관련된 대상 데이터 노드들의 목록을 상기 메타데이터 데이터베이스에서 삭제하는 단계;

상기 지역 그룹 변경 요청이 수락되면, 상기 클라이언트가 상기 대상 오브젝트를 저장할 새로운 대상 데이터 노드들이 속하는 지역 그룹을 선택하고 선택된 지역 그룹들에 대한 상기 클라이언트의 선호도에 의하여 결정되는 우선 순위를 상기 지역 그룹에게 부여하는 단계;

상기 클라이언트가 동일한 우선 순위를 가지는 지역 그룹에 속하는 데이터 노드들 중 하나를 상기 대상 데이터 노드로서 결정하는 단계;

상기 클라이언트가 가장 높은 우선 순위를 가지는 지역 그룹에 속하는 대상 데이터 노드에 상기 대상 오브젝트를 저장함으로써, 상기 우선 순위에 따라 순차적으로 더 낮은 우선 순위를 가지는 지역 그룹에 속하는 대상 데이터 노드들에 상기 대상 오브젝트가 복제되는 복제 단계가 수행되도록 야기하는 단계; 및

상기 클라이언트에 의하여 새로운 지역 그룹에 속하는 대상 데이터 노드들의 목록이 회신되면, 상기 프락시 서버가 회신된 목록을 이용하여 상기 메타데이터 데이터베이스를 갱신하는 단계를 포함하는 것을 특징으로 하는 분산 저장 시스템에서 오브젝트를 분산 저장하기 위한 방법.

**청구항 18**

제15항에 있어서,

상기 지역 그룹 결정 단계는, 상기 클라이언트가 상기 대상 오브젝트를 저장할 대상 데이터 노드들이 속하는 지역 그룹들을 선택하고, 선택된 지역 그룹들에 대한 상기 클라이언트의 선호도에 의하여 결정되는 우선 순위를 상기 지역 그룹에게 부여하는 단계를 포함하고,

상기 대상 데이터 노드 결정 단계는, 프락시 서버가 동일한 지역 그룹에 속하는 데이터 노드들 중 하나의 데이

터 노드를 상기 대상 데이터 노드로서 결정하는 단계를 포함하며,

상기 저장 단계는,

상기 클라이언트가 상기 대상 데이터 노드의 목록을 프락시 서버로부터 수신하는 단계;

상기 클라이언트가 가장 높은 우선 순위를 가지는 지역 그룹에 속하는 대상 데이터 노드에 상기 대상 오브젝트를 저장하는 단계; 및

상기 우선 순위에 따라 순차적으로 더 낮은 우선 순위를 가지는 지역 그룹에 속하는 대상 데이터 노드들에 상기 대상 오브젝트를 복제하는 복제 단계를 포함하고,

상기 메타데이터 갱신 단계는,

프락시 서버가 상기 대상 데이터 노드들의 목록을 이용하여 상기 메타데이터 데이터베이스를 갱신하는 단계를 포함하는 것을 특징으로 하는 분산 저장 시스템에서 오브젝트를 분산 저장하기 위한 방법.

**청구항 19**

제18항에 있어서,

상기 대상 데이터 노드 결정 단계는,

상기 프락시 서버가 동일한 지역 그룹에 포함되는 데이터 노드들의 가용 저장 용량 및 오브젝트 저장 내역을 고려하여 동일한 지역 그룹에 포함되는 데이터 노드들에게 우선 순위를 부여하는 단계 및

상기 프락시 서버가 가장 높은 우선 순위를 가지는 데이터 노드를 상기 대상 데이터 노드로서 결정하는 단계를 포함하는 것을 특징으로 하는 분산 저장 시스템에서 오브젝트를 분산 저장하기 위한 방법.

**청구항 20**

제19항에 있어서,

선택된 지역 그룹들을 변경하기 위한 지역 그룹 변경 단계를 더 포함하며,

상기 지역 그룹 변경 단계는,

상기 클라이언트가 상기 프락시 서버에게 지역 그룹 변경 요청을 송신하는 단계;

상기 지역 그룹 변경 요청이 수락되면, 상기 클라이언트가 상기 대상 오브젝트를 저장할 새로운 대상 데이터 노드들이 속하는 지역 그룹을 선택하고 선택된 지역 그룹들에 대한 상기 클라이언트의 선호도에 의하여 결정되는 우선 순위를 상기 지역 그룹에게 부여하는 단계;

저장된 오브젝트와 관련된 대상 데이터 노드들의 목록을 상기 메타데이터 데이터베이스에서 삭제하는 단계;

상기 프락시 서버가 상기 클라이언트가 선택한 지역 그룹들 각각에 포함되는 데이터 노드들 중 하나의 데이터 노드를 대상 데이터 노드로서 결정하여 상기 클라이언트로 전송하는 단계;

상기 클라이언트가 가장 높은 우선 순위를 가지는 지역 그룹에 속하는 대상 데이터 노드에 상기 대상 오브젝트를 저장함으로써, 상기 우선 순위에 따라 순차적으로 더 낮은 우선 순위를 가지는 지역 그룹에 속하는 대상 데이터 노드들에 상기 대상 오브젝트가 복제되는 복제 단계가 수행되도록 야기하는 단계; 및

상기 프락시 서버가 상기 대상 데이터 노드들의 목록을 이용하여 상기 메타데이터 데이터베이스를 갱신하는 단계를 포함하는 것을 특징으로 하는 분산 저장 시스템에서 오브젝트를 분산 저장하기 위한 방법.

**청구항 21**

제20항에 있어서,

상기 프락시 서버가 동일한 지역 그룹에 포함되는 데이터 노드들의 가용 저장 용량 및 오브젝트 저장 내역을 고려하여 동일한 지역 그룹에 포함되는 데이터 노드들에게 우선 순위를 부여하는 단계; 및

상기 프락시 서버가 가장 높은 우선 순위를 가지는 데이터 노드를 상기 대상 데이터 노드로서 결정하는 단계를 더 포함하는 것을 특징으로 하는 분산 저장 시스템에서 오브젝트를 분산 저장하기 위한 방법.



**청구항 22**

제15항에 있어서,

상기 지역 그룹 결정 단계는,

위치 인식 서버가 상기 데이터 노드들 및 상기 클라이언트의 위치 관계에 기반하여 상기 대상 오브젝트를 저장할 대상 데이터 노드가 속하는 지역 그룹을 선택하는 단계 및

위치 인식 서버가 선택된 지역 그룹 및 상기 클라이언트 간의 거리에 기반하여 상기 지역 그룹별 우선 순위를 결정하는 단계를 포함하고,

상기 대상 데이터 노드 결정 단계는,

프락시 서버가 상기 위치 인식 서버가 선택한 지역 그룹 당 하나의 대상 데이터 노드를 결정하는 단계; 및

프락시 서버가 상기 대상 데이터 노드들의 목록 및 상기 지역 그룹별 우선 순위를 상기 클라이언트에게 전송하는 단계를 포함하며,

상기 메타데이터 갱신 단계는, 프락시 서버가 결정된 대상 데이터 노드들의 목록을 이용하여 상기 메타데이터 데이터베이스를 갱신하는 단계를 포함하고,

상기 저장 단계는,

상기 클라이언트가 가장 높은 우선 순위를 가지는 지역 그룹에 속하는 대상 데이터 노드에 상기 대상 오브젝트를 저장하는 단계; 및

상기 우선 순위에 따라 순차적으로 더 낮은 우선 순위를 가지는 지역 그룹에 속하는 대상 데이터 노드들에 상기 대상 오브젝트를 복제하는 복제 단계를 포함하는 것을 특징으로 하는 분산 저장 시스템에서 오브젝트를 분산 저장하기 위한 방법.

**청구항 23**

제22항에 있어서,

상기 대상 데이터 노드 결정 단계는,

상기 프락시 서버가 동일한 지역 그룹에 포함되는 데이터 노드들의 가용 저장 용량 및 오브젝트 저장 내역을 고려하여 동일한 지역 그룹에 포함되는 데이터 노드들에게 우선 순위를 부여하는 단계; 및

상기 프락시 서버가 가장 높은 우선 순위를 가지는 데이터 노드를 상기 대상 데이터 노드로서 결정하는 단계를 포함하는 것을 특징으로 하는 분산 저장 시스템에서 오브젝트를 분산 저장하기 위한 방법.

**청구항 24**

제16항 내지 제23항 중 어느 한 항에 있어서,

상기 복제 단계는,

상기 클라이언트가 가장 높은 우선 순위보다 낮은 우선 순위를 가지는 모든 지역 그룹들에 속하는 대상 데이터 노드에 상기 대상 오브젝트를 저장하는 단계를 포함하는 것을 특징으로 하는 분산 저장 시스템에서 오브젝트를 분산 저장하기 위한 방법.

**청구항 25**

제16항 내지 제23항 중 어느 한 항에 있어서,

상기 복제 단계는,

상기 대상 데이터 노드들이, 자신이 속한 지역 그룹의 우선 순위보다 한 단계 높은 우선 순위를 가지는 지역 그룹에 속하는 대상 데이터 노드로부터 상기 대상 오브젝트를 수신하여 저장하는 단계를 포함하는 것을 특징으로 하는 분산 저장 시스템에서 오브젝트를 분산 저장하기 위한 방법.

**청구항 26**

제16항 내지 제23항 중 어느 한 항에 있어서,

상기 복제 단계는,

복제 서버가 가장 높은 우선 순위를 가지는 지역 그룹에 속하는 대상 데이터 노드에 저장된 대상 오브젝트를 독출하는 단계; 및

상기 복제 서버가 독출된 대상 오브젝트를 가장 높은 우선 순위보다 낮은 우선 순위를 가지는 모든 지역 그룹들에 속하는 대상 데이터 노드에 저장하는 단계를 포함하는 것을 특징으로 하는 분산 저장 시스템에서 오브젝트를 분산 저장하기 위한 방법.

**청구항 27**

제15항에 있어서,

상기 오브젝트의 고유 정보는 상기 오브젝트의 ID, 크기, 데이터 타입, 및 작성자 중 적어도 하나를 포함하고,

상기 데이터 노드의 고유 정보는 상기 데이터 노드의 ID, IP(Internet Protocol) 주소, 및 물리적 위치 중 적어도 하나를 포함하는 것을 특징으로 하는 분산 저장 시스템에서 오브젝트를 분산 저장하기 위한 방법.

**청구항 28**

삭제

**청구항 29**

제15항 내지 제23항 중 어느 한 항에 따르는 방법을 구현하기 위하여 컴퓨터에 의하여 실행될 수 있는 컴퓨터 프로그램 명령어들을 저장하는 컴퓨터에 의하여 독출가능한 저장 매체.

**발명의 설명**

**기술 분야**

[0001] 본 발명은 복수 개의 클라이언트들로부터 네트워크를 통해 전송되는 오브젝트를 복수 개의 데이터 노드들에 분산 저장하는 분산 저장 시스템에 관한 것으로서, 특히, 분산 저장 시스템에서 오브젝트의 업로드 및 다운로드시 지연을 줄이기 위해 클라이언트 및 데이터 노드의 물리적인 위치를 고려하여 최적의 데이터 노드를 선택하여 오브젝트를 배치하는 방법에 관한 것이다.

**배경 기술**

[0002] 클라우드 컴퓨팅은 인터넷 망을 통해 여러 가지 IT(Information Technology) 자원들을 분배하여 서비스하는 개념으로 가장 일반적인 서비스 분류로는 첫째, 하드웨어 인프라를 서비스로 제공하는 IaaS(Infrastructure As A Service), 둘째, 응용 개발 및 실행 플랫폼을 서비스로 제공하는 PaaS(Platform As A Service), 그리고 마지막으로 어플리케이션을 서비스로 제공하는 SaaS(Software As A Service)로 나눈다.

[0003] IaaS 에는 많은 서비스 카테고리가 있는데, 대표적으로 연산 자원을 가상 머신 형태로 제공하는 연산 서비스와 저장 서비스가 있다. 이러한 분산 저장 시스템은 클라우드 스토리지 서비스를 제공하는데, 이를 이용하면 저사양의 하드웨어를 활용하여 공용의 저장소 풀(storage pool)을 만들어 탄력적이고 유연한 사용량을 적시에 만족하는 특징을 가진다. 이를 위해 널리 이용되는 단순하고 강력한 오브젝트 기반 스토리지 기법에서는 물리적인 저장 공간 관리 기능을 저장 장치 자체에서 직접 수행하게 한다. 그러므로, 저장 장치의 성능이 향상되고 손쉽게 저장 장치의 용량을 확장할 수 있다. 또한, 플랫폼과 독립적으로 데이터를 안전하게 공유할 수 있는 특징을 가진다.

[0004] 도 1은 종래 기술에 의한 분산 저장 시스템을 개념적으로 나타내는 도면이다.

[0005] 도 1에 도시되는 오브젝트 스토리지 시스템은 클라이언트의 인증을 처리하는 인증 서버, 클라이언트의 요구사항을 처리하는 프록시 서버(proxy server)(또는 마스터 서버), 오브젝트들의 물리적인 위치를 포함하는 메타데이터를 저장하는 메타데이터 데이터베이스, 실제 오브젝트 저장 및 관리를 담당하는 데이터 노드, 그리고 데이터

의 복제 등을 관리하는 복제 서버(replicator server) 등을 포함한다. 클라이언트는 초기에 인증 서버를 통해 인증을 받으며, 인증이 완료된 후에는 프락시 서버에게 원하는 오브젝트를 관리하는 데이터 노드의 정보를 요청한다. 클라이언트의 요청에 대해 프락시 서버는 메타데이터를 참조하여 해당하는 데이터 노드의 아이디 및 주소를 클라이언트에게 전달한다. 이후 클라이언트는 프락시 서버를 거치지 않고 직접 데이터 노드와 통신하여 오브젝트의 저장 및 조회 등의 서비스를 요청한다. 이 경우 제어 채널과 데이터 채널이 분리되어 지연이나 데이터 트래픽이 감소하는 효과를 기대할 수 있다.

[0006] 오브젝트 저장소는 데이터의 안전성과 높은 가용성을 위해 데이터를 복제(replication)하며 이런 복제본을 레플리카(replica)라고 한다. 널리 이용되는 분산 저장 시스템은 일반적으로 2개 내지 3개의 복제본을 가지지만, 오브젝트의 중요도에 따라 더 많은 복제본을 가지기도 한다. 오브젝트의 복제본은 서로 동기를 유지해야 하며 이는 보통 별도의 복제 서버 (replicator server)에서 처리한다.

[0007] 그런데, 종래 기술에 의한 분산 저장 시스템의 경우, 클라이언트 및 데이터 노드의 물리적 위치와 무관하게 오브젝트의 저장 및 복제 동작이 수행된다. 그 결과, 물리적으로 원거리에 위치한 클라이언트 및 데이터 노드 간에 오브젝트를 업로드 및 다운로드하는 경우가 발생하는데, 이는 오브젝트의 입출력 시간을 길게 만든다. 또한, 특정 지역에 위치한 복수 개의 데이터 노드에만 오브젝트를 저장 및 복제하는 경우도 발생한다. 따라서, 해당 지역의 네트워크에 문제가 발생할 경우 관심 대상인 오브젝트를 저장하는 모든 데이터 노드가 동작하지 않게 되기 때문에 오브젝트를 이용할 수 없는 심각한 문제가 대두된다.

[0008] 그러므로, 클라이언트 및 데이터 노드의 위치 관계를 고려하여 클라이언트로부터 수신되는 오브젝트를 저장하기에 적합한 데이터 노드를 결정하여 저장하는 것은 물론, 특정 지역의 네트워크 오류 등의 사건이 발생할 경우에도 오브젝트를 안정적으로 획득할 수 있는 기술이 절실히 요구된다. 또한, 오브젝트의 업로드 및 다운로드 지연을 최소화할 수 있는 기술이 절실히 요구된다.

**발명의 내용**

**해결하려는 과제**

[0009] 본 발명의 목적은 클라우드 스토리지 서비스를 위한 오브젝트 스토리지 시스템에서 클라이언트 및 데이터 노드의 위치 정보에 기반하여 효과적으로 복제본을 저장하고 관리할 수 있는 오브젝트의 분산 저장 방법을 제공하는 것이다.

[0010] 본 발명의 다른 목적은 클라이언트 및 데이터 노드의 위치 정보를 이용하여 오브젝트를 선호도가 높은 데이터 노드에 우선적으로 저장함으로써 저장된 오브젝트 조회 및 저장 지연을 최소화할 수 있는 분산 저장 시스템을 제공하는 것이다.

**과제의 해결 수단**

[0011] 상기와 같은 목적들을 달성하기 위한 본 발명의 일면은, 복수 개의 클라이언트들로부터 네트워크를 통해 전송되는 오브젝트를 복수 개의 데이터 노드들에 분산 저장하는 분산 저장 시스템(distribution storage system)에 관한 것이다. 본 발명에 의한 분산 저장 시스템은 클라이언트를 인증하기 위한 인증 서버, 지역(zone)별로 그룹핑되며 각각 적어도 하나의 오브젝트들을 저장하는 복수 개의 데이터 노드들로서, 동일한 오브젝트는 동일한 지역 그룹(zone group)에 속한 데이터 노드들 중 오직 하나에만 저장되도록 구성되는 복수 개의 데이터 노드들, 오브젝트의 고유 정보 및 오브젝트가 저장된 데이터 노드의 고유 정보를 포함하는 메타데이터를 저장하는 메타데이터 데이터베이스, 및 인증된 클라이언트의 질의에 응답하여, 메타데이터를 참조하여 클라이언트의 관리 동작의 객체인 대상 오브젝트(target object)와 관련되는 대상 데이터 노드(target data node)의 고유 정보의 목록을 클라이언트에게 제공하는 프락시 서버를 포함하며, 클라이언트는, 대상 데이터 노드의 고유 정보의 목록을 참조하여 대상 오브젝트의 관리 동작(management process)을 수행하도록 구성된다. 특히, 관리 동작은 대상 오브젝트를 데이터 노드에 저장하는 저장 동작(storage process)이며, 저장 동작을 수행하기 위하여, 클라이언트는 대상 오브젝트를 저장할 대상 데이터 노드들이 속하는 지역 그룹들 및 각 지역 그룹 당 하나의 대상 데이터 노드를 선택하고, 선택된 지역 그룹들에 대한 클라이언트의 선호도(preference)에 의하여 결정되는 우선 순위(priority)를 지역 그룹에게 부여하며, 가장 높은 우선 순위를 가지는 지역 그룹에 속하는 대상 데이터 노드에 대상 오브젝트를 저장함으로써, 우선 순위에 따라 순차적으로 더 낮은 우선 순위를 가지는 지역 그룹에 속하는 대상 데이터 노드들에 대상 오브젝트가 복제되는 복제 동작(replication process)이 수행되도록 야기하도록 더욱 구성되고, 프락시 서버는, 대상 오브젝트가 저장된 대상 데이터 노드들의 목록을 클라이언트에게 질의

(query)하고, 클라이언트에 의하여 대상 데이터 노드들의 목록이 회신되면, 회신된 목록을 이용하여 메타데이터 데이터베이스를 갱신하도록 구성된다. 또는, 관리 동작은 선택된 지역 그룹들을 변경하기 위한 지역 그룹 변경 동작(zone group change process)이며, 지역 그룹 변경 동작을 수행하기 위하여, 클라이언트는, 프락시 서버에게 지역 그룹 변경 요청(zone group change request)을 송신한 후 지역 그룹 변경 요청이 수락되면, 대상 오브젝트를 저장할 새로운 대상 데이터 노드들이 속하는 지역 그룹들 및 각 지역 그룹 당 하나의 대상 데이터 노드를 선택하고, 선택된 지역 그룹들에 대한 클라이언트의 선호도에 의하여 결정되는 우선 순위를 지역 그룹에게 부여하며, 가장 높은 우선 순위를 가지는 지역 그룹에 속하는 대상 데이터 노드에 대상 오브젝트를 저장함으로써, 우선 순위에 따라 순차적으로 더 낮은 우선 순위를 가지는 지역 그룹에 속하는 대상 데이터 노드들에 대상 오브젝트가 복제되는 복제 동작이 수행되도록 야기하도록 더욱 구성되고, 프락시 서버는, 저장된 오브젝트와 관련된 대상 데이터 노드들의 목록을 메타데이터 데이터베이스에서 삭제하고, 클라이언트가 선택한 새로운 지역 그룹에 속하는 대상 데이터 노드들의 목록을 클라이언트에게 질의하며, 클라이언트에 의하여 새로운 지역 그룹에 속하는 대상 데이터 노드들의 목록이 회신되면, 회신된 목록을 이용하여 메타데이터 데이터베이스를 갱신하도록 더욱 구성된다. 또한, 관리 동작은 대상 오브젝트를 데이터 노드에 저장하는 저장 동작이며, 저장 동작을 수행하기 위하여, 클라이언트는, 대상 오브젝트를 저장할 대상 데이터 노드들이 속하는 지역 그룹들을 선택하고, 선택된 지역 그룹들에 대한 클라이언트의 선호도에 의하여 결정되는 우선 순위를 지역 그룹에게 부여하며, 지역 그룹 당 하나의 대상 데이터 노드의 목록을 프락시 서버로부터 수신하고, 가장 높은 우선 순위를 가지는 지역 그룹에 속하는 대상 데이터 노드에 대상 오브젝트를 저장함으로써, 우선 순위에 따라 순차적으로 더 낮은 우선 순위를 가지는 지역 그룹에 속하는 대상 데이터 노드들에 대상 오브젝트가 복제되는 복제 동작이 수행되도록 야기하도록 더욱 구성되고, 프락시 서버는, 클라이언트가 선택한 지역 그룹들 각각에 포함되는 데이터 노드들 중 하나의 데이터 노드를 대상 데이터 노드로서 결정하여 클라이언트로 전송하고, 대상 데이터 노드들의 목록을 이용하여 메타데이터 데이터베이스를 갱신하도록 구성된다. 특히, 프락시 서버는 동일한 지역 그룹에 포함되는 데이터 노드들의 가용 저장 용량 및 오브젝트 저장 내역을 고려하여 동일한 지역 그룹에 포함되는 데이터 노드들에게 우선 순위를 부여하고, 동일한 지역 그룹에 포함되는 데이터 노드들 중에서 가장 높은 우선 순위를 가지는 데이터 노드를 대상 데이터 노드로서 결정하도록 더욱 구성된다. 또는, 관리 동작은 선택된 지역 그룹들을 변경하기 위한 지역 그룹 변경 동작이며, 지역 그룹 변경 동작을 수행하기 위하여, 클라이언트는, 프락시 서버에게 지역 그룹 변경 요청을 송신한 후 지역 그룹 변경 요청이 수락되면, 대상 오브젝트를 저장할 새로운 대상 데이터 노드들이 속하는 지역 그룹을 선택하고, 선택된 지역 그룹들에 대한 클라이언트의 선호도에 의하여 결정되는 우선 순위를 지역 그룹에게 부여하며, 지역 그룹 당 하나의 대상 데이터 노드의 목록을 프락시 서버로부터 수신하고, 가장 높은 우선 순위를 가지는 지역 그룹에 속하는 대상 데이터 노드에 대상 오브젝트를 저장함으로써, 우선 순위에 따라 순차적으로 더 낮은 우선 순위를 가지는 지역 그룹에 속하는 대상 데이터 노드들에 대상 오브젝트가 복제되는 복제 동작이 수행되도록 야기하도록 더욱 구성되고, 프락시 서버는, 저장된 오브젝트와 관련된 대상 데이터 노드들의 목록을 메타데이터 데이터베이스에서 삭제하고, 클라이언트가 선택한 지역 그룹들 각각에 포함되는 데이터 노드들 중 하나의 데이터 노드를 대상 데이터 노드로서 결정하여 클라이언트로 전송하며, 대상 데이터 노드들의 목록을 이용하여 메타데이터 데이터베이스를 갱신하도록 더욱 구성된다. 이 경우에, 프락시 서버는 동일한 지역 그룹에 포함되는 데이터 노드들의 가용 저장 용량 및 오브젝트 저장 내역을 고려하여 동일한 지역 그룹에 포함되는 데이터 노드들에게 우선 순위를 부여하고, 동일한 지역 그룹에 포함되는 데이터 노드들 중에서 가장 높은 우선 순위를 가지는 데이터 노드를 대상 데이터 노드로서 결정하도록 더욱 구성된다. 또한, 관리 동작은 대상 오브젝트를 데이터 노드에 저장하는 저장 동작이며, 분산 저장 시스템은, 데이터 노드들 및 클라이언트의 위치 관계에 기반하여 대상 오브젝트를 저장할 대상 데이터 노드가 속하는 지역 그룹을 선택하고 선택된 지역 그룹 및 클라이언트 간의 거리에 기반하여 지역 그룹별 우선 순위를 결정하는 위치 인식 서버(location-aware server)를 더 포함하고, 저장 동작을 수행하기 위하여, 프락시 서버는, 위치 인식 서버가 선택한 지역 그룹 당 하나의 대상 데이터 노드를 결정하고, 결정된 대상 데이터 노드들의 목록을 이용하여 메타데이터 데이터베이스를 갱신하며, 대상 데이터 노드들의 목록 및 지역 그룹별 우선 순위를 클라이언트에게 전송하도록 구성되고, 클라이언트는, 가장 높은 지역 그룹별 우선 순위를 가지는 지역 그룹에 속하는 대상 데이터 노드에 대상 오브젝트를 저장함으로써, 우선 순위에 따라 순차적으로 더 낮은 우선 순위를 가지는 지역 그룹에 속하는 대상 데이터 노드들에 대상 오브젝트가 복제되는 복제 동작이 수행되도록 야기하도록 더욱 구성된다. 특히, 프락시 서버는 동일한 지역 그룹에 포함되는 데이터 노드들의 가용 저장 용량 및 오브젝트 저장 내역을 고려하여 동일한 지역 그룹에 포함되는 데이터 노드들에게 우선 순위를 부여하고, 가장 높은 우선 순위를 가지는 데이터 노드를 대상 데이터 노드로서 결정하도록 더욱 구성된다. 더 나아가, 복제 동작을 수행하기 위하여, 클라이언트는, 가장 높은 우선 순위보다 낮은 우선 순위를 가지는 모든 지역 그룹들에 속하는 대상 데이터 노드에 대상 오브젝트를 저장하도록 더욱 구성된다. 또는, 복제 동작을 수행하기 위하여, 대상 데이터

노드들은 자신이 속한 지역 그룹의 우선 순위보다 한 단계 높은 우선 순위를 가지는 지역 그룹에 속하는 대상 데이터 노드로부터 대상 오브젝트를 수신하여 저장하도록 구성된다. 더 나아가, 분산 저장 시스템은 가장 높은 우선 순위를 가지는 지역 그룹에 속하는 대상 데이터 노드에 저장된 대상 오브젝트를 독출하고, 독출된 대상 오브젝트를 가장 높은 우선 순위보다 낮은 우선 순위를 가지는 모든 지역 그룹들에 속하는 대상 데이터 노드에 저장하는 복제 서버(replication server)를 더 포함하며, 복제 동작은 복제 서버에 의하여 수행된다. 한편, 오브젝트의 고유 정보는 오브젝트의 ID, 크기, 데이터 타입, 및 작성자 중 적어도 하나를 포함하고, 데이터 노드의 고유 정보는 데이터 노드의 ID, IP(Internet Protocol) 주소, 및 물리적 위치 중 적어도 하나를 포함한다. 특히, 메타데이터는, 데이터 노드들의 사용량, 각 지역 그룹에 속한 데이터 노드들의 목록, 대상 오브젝트에 대한 지역 그룹별 우선 순위, 및 동일한 지역 그룹에 속한 데이터 노드들 간의 우선 순위 중 적어도 하나를 더 포함한다.

[0012]

상기와 같은 목적들을 달성하기 위한 본 발명의 다른 면은 복수 개의 클라이언트들로부터 네트워크를 통해 전송되는 오브젝트를 복수 개의 데이터 노드들에 분산 저장하는 분산 저장 시스템에서 오브젝트를 분산 저장하기 위한 방법에 관한 것이다. 본 발명에 의한 오브젝트 분산 저장 방법은 클라이언트를 인증하는 단계, 데이터 노드들을 각 데이터 노드들의 위치 정보를 기반으로 지역별로 그룹화하여 지역 그룹을 생성하는 단계, 저장할 대상 오브젝트가 존재할 경우, 대상 오브젝트가 저장될 대상 데이터 노드가 속하는 지역 그룹을 결정하는 지역 그룹 결정 단계, 동일한 지역 그룹에 속하는 데이터 노드들 중 하나의 데이터 노드를 대상 데이터 노드로서 결정하는 대상 데이터 노드 결정 단계, 대상 데이터 노드에 대상 오브젝트를 저장하는 저장 단계, 및 대상 오브젝트의 고유 정보 및 대상 오브젝트가 저장된 데이터 노드의 고유 정보를 이용하여 메타데이터를 갱신하는 메타데이터 갱신 단계를 포함한다. 특히, 지역 그룹 결정 단계는, 클라이언트가 대상 오브젝트를 저장할 대상 데이터 노드들이 속하는 지역 그룹들을 선택하고, 선택된 지역 그룹들에 대한 클라이언트의 선호도에 의하여 결정되는 우선 순위를 지역 그룹에게 부여하는 단계를 포함하고, 대상 데이터 노드 결정 단계는, 클라이언트가 동일한 지역 그룹에 속하는 데이터 노드들 중 하나의 데이터 노드를 대상 데이터 노드로서 결정하는 단계를 포함하며, 저장 단계는, 클라이언트가 가장 높은 우선 순위를 가지는 지역 그룹에 속하는 대상 데이터 노드에 대상 오브젝트를 저장하는 단계, 및 우선 순위에 따라 순차적으로 더 낮은 우선 순위를 가지는 지역 그룹에 속하는 대상 데이터 노드들에 대상 오브젝트를 복제하는 복제 단계를 포함하고, 메타데이터 갱신 단계는, 클라이언트에 의하여 대상 데이터 노드들의 목록이 회신되면, 프락시 서버가 회신된 목록을 이용하여 메타데이터 데이터베이스를 갱신하는 단계를 포함한다. 특히, 본 발명에 의한 방법은 선택된 지역 그룹들을 변경하기 위한 지역 그룹 변경 단계를 더 포함하며, 지역 그룹 변경 단계는, 클라이언트가 프락시 서버에게 지역 그룹 변경 요청을 송신하는 단계, 프락시 서버가 지역 그룹 변경 요청을 수락하고 저장된 오브젝트와 관련된 대상 데이터 노드들의 목록을 메타데이터 데이터베이스에서 삭제하는 단계, 지역 그룹 변경 요청이 수락되면, 클라이언트가 대상 오브젝트를 저장할 새로운 대상 데이터 노드들이 속하는 지역 그룹을 선택하고 선택된 지역 그룹들에 대한 클라이언트의 선호도에 의하여 결정되는 우선 순위를 지역 그룹에게 부여하는 단계, 클라이언트가 동일한 우선 순위를 가지는 지역 그룹에 속하는 데이터 노드들 중 하나를 대상 데이터 노드로서 결정하는 단계, 클라이언트가 가장 높은 우선 순위를 가지는 지역 그룹에 속하는 대상 데이터 노드에 대상 오브젝트를 저장함으로써, 우선 순위에 따라 순차적으로 더 낮은 우선 순위를 가지는 지역 그룹에 속하는 대상 데이터 노드들에 대상 오브젝트가 복제되는 복제 단계가 수행되도록 야기하는 단계, 및 클라이언트에 의하여 새로운 지역 그룹에 속하는 대상 데이터 노드들의 목록이 회신되면, 프락시 서버가 회신된 목록을 이용하여 메타데이터 데이터베이스를 갱신하는 단계를 포함한다. 또는, 지역 그룹 결정 단계는, 클라이언트가 대상 오브젝트를 저장할 대상 데이터 노드들이 속하는 지역 그룹들을 선택하고, 선택된 지역 그룹들에 대한 클라이언트의 선호도에 의하여 결정되는 우선 순위를 지역 그룹에게 부여하는 단계를 포함하고, 대상 데이터 노드 결정 단계는, 프락시 서버가 동일한 지역 그룹에 속하는 데이터 노드들 중 하나의 데이터 노드를 대상 데이터 노드로서 결정하는 단계를 포함하며, 저장 단계는, 클라이언트가 대상 데이터 노드의 목록을 프락시 서버로부터 수신하는 단계, 클라이언트가 가장 높은 우선 순위를 가지는 지역 그룹에 속하는 대상 데이터 노드에 대상 오브젝트를 저장하는 단계, 및 우선 순위에 따라 순차적으로 더 낮은 우선 순위를 가지는 지역 그룹에 속하는 대상 데이터 노드들에 대상 오브젝트를 복제하는 복제 단계를 포함하고, 메타데이터 갱신 단계는, 프락시 서버가 대상 데이터 노드들의 목록을 이용하여 메타데이터 데이터베이스를 갱신하는 단계를 포함한다. 더 나아가, 대상 데이터 노드 결정 단계는 프락시 서버가 동일한 지역 그룹에 포함되는 데이터 노드들의 가용 저장 용량 및 오브젝트 저장 내역을 고려하여 동일한 지역 그룹에 포함되는 데이터 노드들에게 우선 순위를 부여하는 단계 및 프락시 서버가 가장 높은 우선 순위를 가지는 데이터 노드를 대상 데이터 노드로서 결정하는 단계를 포함한다. 또한, 본 발명에 의한 방법은 선택된 지역 그룹들을 변경하기 위한 지역 그룹 변경 단계를 더 포함하며, 지역 그룹 변경 단계는, 클라이언트가 프락시 서버에게 지역 그룹 변

경 요청을 송신하는 단계, 지역 그룹 변경 요청이 수락되면, 클라이언트가 대상 오브젝트를 저장할 새로운 대상 데이터 노드들이 속하는 지역 그룹을 선택하고 선택된 지역 그룹들에 대한 클라이언트의 선호도에 의하여 결정되는 우선 순위를 지역 그룹에게 부여하는 단계, 저장된 오브젝트와 관련된 대상 데이터 노드들의 목록을 메타데이터 데이터베이스에서 삭제하는 단계, 프락시 서버가 클라이언트가 선택한 지역 그룹들 각각에 포함되는 데이터 노드들 중 하나의 데이터 노드를 대상 데이터 노드로서 결정하여 클라이언트로 전송하는 단계, 클라이언트가 가장 높은 우선 순위를 가지는 지역 그룹에 속하는 대상 데이터 노드에 대상 오브젝트를 저장함으로써, 우선 순위에 따라 순차적으로 더 낮은 우선 순위를 가지는 지역 그룹에 속하는 대상 데이터 노드들에 대상 오브젝트가 복제되는 복제 단계가 수행되도록 야기하는 단계, 및 프락시 서버가 대상 데이터 노드들의 목록을 이용하여 메타데이터 데이터베이스를 갱신하는 단계를 포함한다. 바람직하게는, 본 발명에 의한 오브젝트 분산 저장 방법은 프락시 서버가 동일한 지역 그룹에 포함되는 데이터 노드들의 가용 저장 용량 및 오브젝트 저장 내역을 고려하여 동일한 지역 그룹에 포함되는 데이터 노드들에게 우선 순위를 부여하는 단계, 및 프락시 서버가 가장 높은 우선 순위를 가지는 데이터 노드를 대상 데이터 노드로서 결정하는 단계를 더 포함한다. 특히, 지역 그룹 결정 단계는 위치 인식 서버가 데이터 노드들 및 클라이언트의 위치 관계에 기반하여 대상 오브젝트를 저장할 대상 데이터 노드가 속하는 지역 그룹을 선택하는 단계 및 위치 인식 서버가 선택된 지역 그룹 및 클라이언트 간의 거리에 기반하여 지역 그룹별 우선 순위를 결정하는 단계를 포함하고, 대상 데이터 노드 결정 단계는, 프락시 서버가 위치 인식 서버가 선택한 지역 그룹 당 하나의 대상 데이터 노드를 결정하는 단계, 및 프락시 서버가 대상 데이터 노드들의 목록 및 지역 그룹별 우선 순위를 클라이언트에게 전송하는 단계를 포함하며, 메타데이터 갱신 단계는, 프락시 서버가 결정된 대상 데이터 노드들의 목록을 이용하여 메타데이터 데이터베이스를 갱신하는 단계를 포함하고, 저장 단계는, 클라이언트가 가장 높은 우선 순위를 가지는 지역 그룹에 속하는 대상 데이터 노드에 대상 오브젝트를 저장하는 단계, 및 우선 순위에 따라 순차적으로 더 낮은 우선 순위를 가지는 지역 그룹에 속하는 대상 데이터 노드들에 대상 오브젝트를 복제하는 복제 단계를 포함한다. 특히, 대상 데이터 노드 결정 단계는 프락시 서버가 동일한 지역 그룹에 포함되는 데이터 노드들의 가용 저장 용량 및 오브젝트 저장 내역을 고려하여 동일한 지역 그룹에 포함되는 데이터 노드들에게 우선 순위를 부여하는 단계, 및 프락시 서버가 가장 높은 우선 순위를 가지는 데이터 노드를 대상 데이터 노드로서 결정하는 단계를 포함하는 것을 특징으로 한다. 뿐만 아니라, 복제 단계는 클라이언트가 가장 높은 우선 순위보다 낮은 우선 순위를 가지는 모든 지역 그룹들에 속하는 대상 데이터 노드에 대상 오브젝트를 저장하는 단계를 포함한다. 또는, 복제 단계는 대상 데이터 노드들이, 자신이 속한 지역 그룹의 우선 순위보다 한 단계 높은 우선 순위를 가지는 지역 그룹에 속하는 대상 데이터 노드로부터 대상 오브젝트를 수신하여 저장하는 단계를 포함한다. 더 나아가, 복제 단계는 복제 서버가 가장 높은 우선 순위를 가지는 지역 그룹에 속하는 대상 데이터 노드에 저장된 대상 오브젝트를 독출하는 단계, 및 복제 서버가 독출된 대상 오브젝트를 가장 높은 우선 순위보다 낮은 우선 순위를 가지는 모든 지역 그룹들에 속하는 대상 데이터 노드에 저장하는 단계를 포함한다.

**발명의 효과**

[0013] 본 발명에 의하여, 클라우드 스토리지 서비스에서 데이터 노드들을 지역(zone) 이라는 구역으로 그룹화하기 때문에 실제 복제본들이 안전하게 지역적으로 분산 저장되어 하나의 지역에 문제가 발생할 경우에도 더욱 신뢰성 있는 서비스가 가능하게 된다.

[0014] 또한, 오브젝트들의 복제본들을 지역 개념을 도입하여 물리적으로 분산하여 배치함으로써 오브젝트의 가용성과 신뢰성을 높이기 위한 목적으로 수행되는 복제 동작을 보다 효율적으로 지원할 수 있으며, 클라이언트의 위치와 물리적으로 가까운 데이터 노드에 오브젝트를 배치함으로써 오브젝트의 저장과 조회 지연을 최소화할 수 있다.

**도면의 간단한 설명**

- [0015] 도 1은 종래 기술에 의한 분산 저장 시스템을 개념적으로 나타내는 도면이다.
- 도 2는 본 발명의 일면에 의한 분산 저장 시스템의 일 실시예를 개념적으로 나타내는 도면이다.
- 도 3은 지역 그룹 및 대상 데이터 노드의 선택 주체에 따른 다양한 알고리즘을 정리하는 표를 도시한다.
- 도 4는 본 발명의 다른 면에 의한 분산 저장 시스템의 오브젝트 저장 방법을 개념적으로 나타내는 흐름도이다.
- 도 5는 본 발명의 다른 면에 의한 분산 저장 시스템의 오브젝트 저장 방법에 포함되는 복제 단계를 나타내는 흐름도이다.

름도이다.

도 6은 본 발명의 일면에 의한 분산 저장 시스템의 다른 실시예를 개념적으로 나타내는 도면이다.

**발명을 실시하기 위한 구체적인 내용**

- [0016] 본 발명과 본 발명의 동작상의 이점 및 본 발명의 실시에 의하여 달성되는 목적을 충분히 이해하기 위해서는 본 발명의 바람직한 실시예를 예시하는 첨부 도면 및 첨부 도면에 기재된 내용을 참조하여야만 한다.
- [0017] 이하, 첨부한 도면을 참조하여 본 발명의 바람직한 실시예를 설명함으로써, 본 발명을 상세히 설명한다. 그러나, 본 발명은 여러 가지 상이한 형태로 구현될 수 있으며, 설명하는 실시예에 한정되는 것이 아니다. 그리고, 본 발명을 명확하게 설명하기 위하여 설명과 관계없는 부분은 생략되며, 도면의 동일한 참조부호는 동일한 부재임을 나타낸다.
- [0018] 명세서 전체에서, 어떤 부분이 어떤 구성요소를 "포함"한다고 할 때, 이는 특별히 반대되는 기재가 없는 한 다른 구성요소를 제외하는 것이 아니라, 다른 구성요소를 더 포함할 수 있는 것을 의미한다. 또한, 명세서에 기재된 "...부", "...기", "모듈", "블록" 등의 용어는 적어도 하나의 기능이나 동작을 처리하는 단위를 의미하며, 이는 하드웨어나 소프트웨어 또는 하드웨어 및 소프트웨어의 결합으로 구현될 수 있다.
- [0019] 도 2는 본 발명의 일면에 의한 분산 저장 시스템의 일 실시예를 개념적으로 나타내는 도면이다.
- [0020] 도 2에 도시된 분산 저장 시스템(200)은 네트워크(290)에 연결되는 복수 개의 클라이언트들(210, 212, 216) 및 데이터 노드들(DN11-DN1n, DN21-DN2n, DNm1-DNmn)을 포함한다. 또한, 도 2에 도시된 분산 저장 시스템(200)은 인증 서버(220), 프락시 서버(250) 및 메타데이터 데이터베이스(280)를 더 포함한다.
- [0021] 도 2를 참조하면, 데이터 노드들(DN11-DN1n, DN21-DN2n, DNm1-DNmn)은 제1 내지 제m 지역 그룹(ZG1, ZG2, ZGm) 중 어느 하나에 포함된다. 도 2에 도시되는 지역 그룹(ZG1, ZG2, ZG3)들은 효과적인 복제본의 분산 저장을 위하여, 각각 지역적으로 인접한 데이터 노드들을 그룹핑함으로써 정의된다. 또한, 동일한 지역 그룹에 속하는 데이터 노드들은 동일한 오브젝트를 저장하지 않도록 구성된다. 즉, 하나의 오브젝트의 복제본은 다른 지역 그룹에 속하는 데이터 노드들에 분산되어 저장되기 때문에 두 개의 복제본이 어느 하나의 지역 그룹에 속한 두 개의 데이터 노드에 공통적으로 저장되지 않는다. 이를 메타데이터의 관점에서 보면, 오브젝트의 물리적인 위치를 나타내는 메타데이터에서 하나의 오브젝트의 복제본들은 다른 지역 그룹에 속한 데이터 노드들에 매핑된다. 그러므로, 어느 특정한 지역 그룹이 전체 네트워크 상에 문제가 생기는 등의 물리적인 피해를 입은 경우라도 복제본이 다른 지역 그룹에 속한 데이터 노드에 분산되어 저장되어 있으므로 신뢰성을 높일 수 있다.
- [0022] 본 발명에서, 지역 그룹은 하나의 데이터 센터가 될 수도 있고 좀 더 좁은 지역으로는 하나의 서버 랙이 될 수도 있다. 지역 그룹이 설정되면, 해당 지역 그룹 내에 속한 데이터 노드들을 그 지역 그룹에 속하는 것으로 메타데이터에 등록된다. 그러면, 오브젝트의 복제본들은 다른 지역 그룹에 속한 데이터 노드에 복제된다.
- [0023] 데이터 노드를 지역 그룹으로 그룹핑함으로써 얻어지는 장점은 다음과 같다.
- [0024] 1) 본 발명에서, 모든 클라이언트들(210, 212, 216) 및 데이터 노드들(DN11-DN1n, DN21-DN2n, DNm1-DNmn)들은 네트워크(290)를 통하여 서로 통신한다. 즉, 클라이언트 각각 및 데이터 각각 간에는 가상 채널(virtual channel)이 존재한다. 그러나, 이러한 가상 채널은 클라이언트 및 데이터 노드의 모든 쌍에 대해 반드시 동일한 조건을 가지는 것은 아니다. 예를 들어, 가상 채널의 통신 환경은 클라이언트 및 데이터 노드 간의 물리적 거리에 따라 달라질 수 있다. 클라이언트 및 데이터 노드 간의 물리적 거리가 멀수록 오브젝트는 더 많은 릴레이 노드 또는 게이트웨이를 통하여 전달되기 때문에 송수신 시간이 오래 걸린다. 또한, 가상 채널의 통신 환경은 네트워크 트래픽의 양 및 가상 채널을 구성하는 네트워크 자원의 성능에 따라서도 달라질 수 있다. 가상 채널을 통해 전달되는 트래픽의 양이 많을수록 가상 채널 상에서 전송 충돌(transmission collision)이 발생할 가능성이 높아지며, 네트워크 자원의 성능이 높을수록 가상 채널의 송수신 속도는 빨라진다. 그러므로, 본 발명에서는 이와 같은 가상 채널의 통신 환경을 고려하여 클라이언트 및 데이터 노드 간에 가장 최적의 가상 채널을 선택한다. 최적의 가상 채널을 선택하기 위하여 본 발명에 의한 분산 저장 시스템에서는 클라이언트 및 지역 그룹 간의 물리적 거리를 참조할 수 있다. 따라서, 저장한 오브젝트를 포함하는 클라이언트로부터 가장 가까운 거리에 위치한 지역 그룹에 속한 데이터 노드에 오브젝트를 저장함으로써 오브젝트의 업로드 시간을 최소화할 수 있다.

- [0025] 2) 또한, 본 발명에 의한 분산 저장 시스템은 오브젝트를 복제할 때 동일한 지역 그룹에 속하는 데이터 노드로 복제되지 않도록 한다. 따라서, 저장할 대상 오브젝트는 여러 개의 지역 그룹에 분산 저장된다. 일반적으로, 네트워크 장애가 발생하면 인접 지역의 데이터 노드들의 동작도 불가능한 경우가 많다. 예를 들어, 어느 데이터 센터에 여러 개의 데이터 노드들이 존재하고, 이 데이터 센터가 하나의 지역 그룹으로 설정된다고 가정한다. 이러한 가정은 본 발명을 용이하게 설명하기 위한 것으로 본 발명을 한정하는 것이 아님은 명백하다. 갑작스런 정전 등의 사고로 해당 데이터 센터가 동작 불능의 상태에 빠지는 경우가 발생할 수 있다. 이 경우, 본 발명에 의한 분산 저장 시스템은 대상 오브젝트를 해당 데이터 센터의 데이터 노드들 중 오직 하나의 대상 데이터 노드에만 저장하고, 그 복제본은 다른 지역 그룹에 속하는 대상 데이터 노드에 저장한다. 그러므로, 데이터 센터의 모든 데이터 노드들이 장애를 일으키더라도, 다른 지역 그룹에 속하는 대상 데이터 노드로부터 원하는 대상 오브젝트를 용이하게 조회할 수 있다.
- [0026] 다시 도 2를 참조하면, 분산 저장 시스템(200)에서 데이터 노드를 할당하는 방식으로는 수동 선택 방식과 자동 선택 방식의 두 가지 방식이 제안될 수 있다. 수동 선택 방법은 클라이언트가 최초로 계정을 만들 때 프락시 서버에서 선호하는 지역 그룹을 클라이언트에게 질의하여 클라이언트에게 선택권을 주는 방법이다. 복제본의 개수에 따라 지역 그룹의 개수가 정해지기 때문에, 클라이언트는 허용되는 가장 많은 복제본 개수만큼 선호하는 지역 그룹을 우선 순위에 따라 선택한다. 즉, 클라이언트는 오브젝트에 따라 복제본 개수를 다르게 선택할 수 있는 경우를 고려해 허용되는 최대 복제본 개수만큼 선호 지역 그룹을 선택한다. 이 후 선호하는 지역 그룹을 바꾸고자 할 경우에는 프락시 서버에게 지역 그룹 변경을 요청하고, 이 요청이 받아들여지면 그 이 후에는 대상 오브젝트를 변경된 선호 지역 그룹에 속하는 대상 데이터 노드에 저장한다.
- [0027] 자동 선택 방법은 위치 인식 서버(location-aware server)를 이용하여 지역 그룹을 선택하는 방식이다. 자동 선택 방법에 대해서는 도 6을 참조하여 상세히 후술된다.
- [0028] 도 3은 지역 그룹 및 대상 데이터 노드의 선택 주체에 따른 다양한 알고리즘을 정리하는 표를 도시한다.
- [0029] 도 3을 참조하면, 분산 저장 시스템(200)에서 대상 오브젝트를 저장하기 위해서는 우선 클라이언트가 선호하는 지역 그룹을 선택하고, 선택된 지역 그룹에 속하는 데이터 노드 중에서 하나의 데이터 노드를 대상 데이터 노드로서 선택해야 한다. 이 때, 지역 그룹 및 대상 데이터 노드를 선택하는 주체가 무엇인지에 따라서 도 3에 제공되는 바와 같은 알고리즘들이 존재할 수 있다.
- [0030] 알고리즘 1
- [0031] 알고리즘 1에서는 지역 그룹 및 대상 데이터 노드를 모두 클라이언트가 선택한다. 본 명세서에서 '대상 오브젝트(target object)'란 클라이언트가 저장하고자 하는 오브젝트나 데이터 노드로부터 조회하고자 하는 관심 대상인 오브젝트를 의미한다. 또한, '대상 데이터 노드(target data node)'란 여러 개의 데이터 노드 중에서 대상 오브젝트가 저장된 데이터 노드를 의미한다.
- [0032] 알고리즘 1에서는 클라이언트는 지역 그룹 및 대상 데이터 노드를 선택하고, 선택된 지역 그룹들에 대한 클라이언트의 선호도(preference)에 의하여 결정되는 우선 순위(priority)를 지역 그룹에게 부여한다. 본 명세서에서 '우선 순위'란 특정 대상 오브젝트를 저장하기에 어떤 지역 그룹 또는 데이터 노드가 다른 지역 그룹 또는 데이터 노드에 비하여 더 적합한지 판단하여 매긴 순위를 나타낸다. 우선 순위에는 특정 지역 그룹이 다른 지역 그룹에 비하여 가지는 우선 순위 및 동일한 지역 그룹 내에 속한 데이터 노드들 간의 우선 순위가 포함될 수 있다. 또한, 우선 순위는 클라이언트가 직접 어느 대상 오브젝트에 관련한 특정 지역 및 데이터 노드에 대한 선호도에 기반하여 매길 수도 있고, 또는 프락시 서버 또는 위치 인식 서버에 의하여 자동으로 결정될 수도 있다. 이에 대해서는 명세서의 해당 부분에서 상세히 후술된다.
- [0033] 이와 같이, 클라이언트에 의하여 지역 그룹 및 대상 데이터 노드가 결정되면, 클라이언트는 자신이 선택한 대상 데이터 노드들에 대상 오브젝트를 저장하고 복제한다. 복제 동작에 대해서는 도 5를 이용하여 상세히 후술된다.
- [0034] 클라이언트가 대상 데이터 노드를 결정하면, 프락시 서버는 대상 오브젝트가 저장된 대상 데이터 노드들의 목록을 클라이언트에게 질의(query)하고, 회신된 대상 데이터 노드들의 목록을 이용하여 메타데이터 데이터베이스를 갱신한다.
- [0035] 알고리즘 1에 따르면, 클라이언트는 자신이 선택한 지역 그룹 및 대상 데이터 노드를 변경할 수도 있다. 이러한 변경 동작을 수행하기 위하여, 클라이언트는 우선 프락시 서버에게 지역 그룹 변경 요청(zone group change



request)을 송신한다. 프락시 서버 서버는 이러한 변경 요청을 수락할 경우 저장된 오브젝트와 관련된 대상 데이터 노드들의 목록을 메타데이터 데이터베이스에서 삭제한다. 지역 그룹 변경 요청이 수락되면, 클라이언트는 대상 오브젝트를 저장할 새로운 대상 데이터 노드들이 속하는 지역 그룹들 및 각 지역 그룹 당 하나의 대상 데이터 노드를 선택한다. 또한, 클라이언트는 선택된 지역 그룹들에 대한 선호도에 의하여 결정되는 우선 순위를 지역 그룹에게 부여할 수 있다. 그러면, 프락시 서버는 클라이언트가 선택한 새로운 지역 그룹에 속하는 대상 데이터 노드들의 목록을 클라이언트에게 질의하여 회신된 목록을 이용하여 메타데이터 데이터베이스를 갱신한다. 본 명세서에서 A라는 오브젝트가 B라는 데이터 노드와 '관련' 된다는 의미는 다음과 같이 이해된다. 우선, 관리 동작이 오브젝트 저장 동작일 경우에는, A 오브젝트가 저장될 대상 데이터 노드로서 데이터 노드 B가 결정된다는 것을 의미할 수 있다. 또는, 관리 동작이 오브젝트 독출 동작일 경우에는, A 오브젝트가 이미 데이터 노드 B에 저장되어 있다는 것을 의미할 수 있다.

[0036] 본 명세서에서 대상 오브젝트의 관리 동작이란, 대상 오브젝트를 대상 데이터 노드에 저장하는 저장 동작, 저장된 대상 오브젝트를 대상 데이터 노드로부터 독출하는 독출 동작, 대상 오브젝트가 저장된 지역 그룹 또는 대상 데이터 노드를 변경하려는 변경 동작, 및 저장된 대상 오브젝트를 수정하려는 동작 등을 모두 포함하는 것으로 이해되어야 한다. 즉, 클라이언트가 오브젝트를 분산 저장하고, 저장된 오브젝트를 독출하여 수정하는 모든 동작이 관리 동작에 포함되는 것으로 이해된다. 또한, 본 명세서에서 분산 저장 시스템이란 클라우드 컴퓨팅을 지원하는 클라우드 스토리지 시스템과 같이, 동일한 오브젝트를 두 개 이상의 저장소에 기록하여 데이터의 손실을 방지하고 저장 부담을 경감시키려는 모든 시스템을 총칭하여 가리키는 용어로 이해되어야 한다.

[0037] 알고리즘 1에 의하면, 클라이언트가 직접 지역 그룹 및 데이터 노드를 선택하기 때문에 사용자가 선호하는 데이터 노드에 우선적으로 대상 오브젝트를 저장할 수 있다는 장점이 있다.

[0038] 알고리즘 2

[0039] 알고리즘 2에서는 지역 그룹은 클라이언트가 선택하지만 대상 데이터 노드는 프락시 서버에 의하여 선택된다. 클라이언트가 주로 관심을 가지는 것은 자신으로부터 가까운 지역 그룹이 어디 있는지가 될 수 있다. 또한, 지역 그룹을 선택한 후 클라이언트는 해당 지역 그룹 내에 어느 데이터 노드가 대상 데이터 노드로서 선택되는지에 대해서는 별로 관심이 없을 수도 있다. 특히, 지역 그룹에 속하는 데이터 노드의 개수가 많아질수록 클라이언트가 이들 중 하나를 대상 데이터 노드로서 선택하는 작업은 매우 지루하고 불필요한 것일 수 있다. 따라서, 클라이언트가 지역 그룹만 선택하면, 프락시 서버가 자동적으로 대상 데이터 노드를 선택할 필요가 있다.

[0040] 알고리즘 2를 상세히 설명하면 다음과 같다. 우선, 클라이언트는 대상 오브젝트를 저장할 대상 데이터 노드들이 속하는 지역 그룹들 및 선택된 지역 그룹들에 대한 클라이언트의 선호도에 의하여 결정되는 우선 순위를 지역 그룹에게 부여한다. 그러면, 프락시 서버는 클라이언트가 선택한 지역 그룹들 각각에 포함되는 데이터 노드들 중 하나의 데이터 노드를 대상 데이터 노드로서 결정하여 클라이언트로 전송하고, 대상 데이터 노드들의 목록을 이용하여 메타데이터 데이터베이스를 갱신한다.

[0041] 클라이언트는, 지역 그룹 당 하나의 대상 데이터 노드의 목록을 프락시 서버로부터 수신한 뒤 대상 오브젝트의 저장 및 복제 동작을 수행할 수 있다. 이 때, 프락시 서버는 대상 데이터 노드를 선택하기 위하여 동일한 지역 그룹에 포함되는 데이터 노드들의 가용 저장 용량 및 오브젝트 저장 내역을 고려할 수 있다. 즉, 프락시 서버는 가용 저장 용량 및 오브젝트 저장 내역을 고려하여 동일한 지역 그룹에 포함되는 데이터 노드들에게 우선 순위를 부여한 후, 가장 높은 우선 순위를 가지는 데이터 노드를 대상 데이터 노드로서 결정할 수 있다.

[0042] 알고리즘 2에서, 클라이언트가 지역 그룹을 변경하고자 할 경우에도 지역 그룹만을 선택하면 되고 대상 데이터 노드까지는 선택할 필요가 없다. 따라서, 클라이언트는 지역 그룹 변경 요청이 수락되면, 대상 오브젝트를 저장할 새로운 대상 데이터 노드들이 속하는 지역 그룹을 선택하고 선택된 지역 그룹에 우선 순위를 부여한다. 그러면, 프락시 서버는 저장된 오브젝트와 관련된 대상 데이터 노드들의 목록을 메타데이터 데이터베이스에서 삭제하고, 클라이언트가 선택한 지역 그룹들 각각에 포함되는 데이터 노드들 중 하나의 데이터 노드를 대상 데이터 노드로서 결정하여 클라이언트로 전송한다. 또한, 프락시 서버는 대상 데이터 노드들의 목록을 이용하여 메타데이터 데이터베이스를 갱신한다. 지역 그룹 당 하나의 대상 데이터 노드의 목록이 프락시 서버로부터 수신되면, 클라이언트는 가장 높은 우선 순위를 가지는 지역 그룹에 속하는 대상 데이터 노드에 대상 오브젝트를 저장함으로써, 대상 오브젝트의 복제 동작이 수행되도록 할 수 있다.

[0043] 알고리즘 1에 비하여, 알고리즘 2는 클라이언트가 오직 지역 그룹만을 선택하면 되기 때문에 데이터 저장 및 복제 동작이 신속하게 수행될 수 있다는 장점을 가진다. 또한, 동일한 지역 그룹 내의 데이터 노드들 중 가장 높

은 우선 순위를 가지는 데이터 노드를 프락시 서버가 자동으로 선택하기 때문에, 데이터 노드들 간 로드 밸런싱(load balancing) 효과가 달성된다.

[0044] 알고리즘 3

[0045] 알고리즘 3에서는 지역 그룹 및 대상 데이터 노드 선택 동작이 모두 자동적으로 이루어진다.

[0046] 클라이언트로부터 가장 근접한 지역 그룹을 선택하기 위하여 위치 인식 서버와 같은 장치가 이용될 수 있음은 전술된 바와 같다. 위치 인식 서버는 프락시 서버와 별개의 장치로서 존재할 수도 있지만, 프락시 서버 내에 통합될 수도 있음은 물론이다. 위치 인식 서버는 데이터 노드들 및 클라이언트의 위치 관계에 기반하여 대상 오브젝트를 저장할 대상 데이터 노드가 속하는 지역 그룹을 선택하고 선택된 지역 그룹 및 클라이언트 간의 거리에 기반하여 지역 그룹별 우선 순위를 결정한다. 그러면, 프락시 서버는 위치 인식 서버가 선택한 지역 그룹 당 하나의 대상 데이터 노드를 결정하고, 결정된 대상 데이터 노드들의 목록을 이용하여 메타데이터 데이터베이스를 갱신한다. 또한, 프락시 서버가 대상 데이터 노드들의 목록 및 지역 그룹별 우선 순위를 클라이언트에게 전송하면, 클라이언트는 가장 높은 지역 그룹별 우선 순위를 가지는 지역 그룹에 속하는 대상 데이터 노드에 대상 오브젝트를 저장함으로써 복제 동작이 수행되도록 한다.

[0047] 알고리즘 3에서는 지역 그룹 및 대상 데이터 노드의 선택 동작이 자동적으로 수행되기 때문에, 알고리즘 1 및 2에 비하여 가장 신속하게 오브젝트를 저장할 수 있다는 장점을 가진다. 또한, 위치 인식 서버는 실시간으로 변동되는 네트워크 환경을 고려하여 최적 지역 그룹을 동적으로 선택할 수 있다. 또한, 선택 동작에 클라이언트가 관여하지 않기 때문에, 알고리즘 1 및 2와는 달리 지역 그룹 변경 동작이 불필요할 수 있다. 하지만, 클라이언트가 원할 경우 자동 선택된 지역 그룹 및 대상 데이터 노드를 변경할 수도 있음은 명백하다. 지역 그룹을 변경하고자 원하면, 클라이언트는 대상 오브젝트를 저장할 새로운 대상 데이터 노드들이 속하는 지역 그룹 및 대상 데이터 노드를 모두 선택할 수 있음은 전술된 바와 같다. 또는, 클라이언트는 지역 그룹만을 선택하고, 대상 데이터 노드는 프락시 서버에 의하여 결정될 수도 있음도 역시 전술된 바와 같다.

[0048] 본 발명에서 대상 오브젝트가 대상 데이터 노드에 복제되는 복제 동작에 대해서는 도 5를 이용하여 상세히 후술한다.

[0049] 도 4는 본 발명의 다른 면에 의한 분산 저장 시스템의 오브젝트 저장 방법을 개념적으로 나타내는 흐름도이다.

[0050] 우선, 인증 서버가 클라이언트를 인증한다(S410). 또한, 본 발명에 의한 오브젝트 저장 방법은 데이터 노드들을 각 데이터 노드들의 위치 정보를 기반으로 지역별로 그룹핑하여 지역 그룹을 생성한다(S320). 동일한 지역 그룹에 속하는 데이터들 간에는 대상 오브젝트의 복제가 이루어지지 않는다는 것은 전술한 바와 같다.

[0051] 클라이언트가 대상 오브젝트를 저장할 필요가 있다고 판단하면, 대상 오브젝트가 저장될 대상 데이터 노드가 속하는 지역 그룹 및 동일한 지역 그룹에 속하는 데이터 노드들 중 하나의 대상 데이터 노드가 결정된다(S430, S440). 이 경우, 수동 선택 방식일 경우에는, 지역 그룹 및 대상 데이터 노드를 모두 클라이언트가 선택하거나, 클라이언트는 지역 그룹만을 선택하고, 대상 데이터 노드는 프락시 서버가 선택할 수 있음은 전술된 바와 같다. 또한, 지역 그룹 및 대상 데이터 노드는 프락시 서버 또는 위치 인식 서버에 의하여 자동으로 선택될 수도 있다는 것도 전술된 바와 같다.

[0052] 이와 같이, 대상 오브젝트를 저장한 대상 데이터 노드가 결정되면, 대상 오브젝트를 해당 대상 데이터 노드에 저장 및 복제한다(S450). 이 때, 클라이언트는 가장 높은 우선 순위를 가지는 지역 그룹에 속하는 데이터 노드에 대상 오브젝트를 저장한다. 이후의 복제 동작은 각각 클라이언트, 데이터 노드, 및 복제 서버를 통하여 수행될 수 있는데, 이에 대해서는 도 5를 이용하여 후술된다.

[0053] 이와 같이, 대상 오브젝트가 대상 데이터 노드에 저장되면, 프락시 서버는 대상 오브젝트의 고유 정보 및 대상 오브젝트가 저장된 데이터 노드의 고유 정보를 이용하여 메타데이터를 갱신한다(S460). 갱신된 메타데이터는 추후에 클라이언트가 저장된 대상 오브젝트를 조회할 경우에 해당 대상 오브젝트가 저장된 대상 데이터 노드를 통지하는데 이용된다.

[0054] 또한, 본 발명에 의한 분산 저장 방법은 이미 선택된 지역 그룹들을 변경하기 위한 지역 그룹 변경 단계를 더 포함할 수 있다. 지역 그룹 변경 동작을 수행하기 위해서는, 우선 클라이언트가 프락시 서버에게 지역 그룹 변경 요청을 송신한다. 그러면, 프락시 서버는 지역 그룹 변경 요청을 수락할지 여부를 결정하고, 수락할 경우 저장된 오브젝트와 관련된 대상 데이터 노드들의 목록을 메타데이터 데이터베이스에서 삭제한다.

- [0055] 대상 오브젝트를 저장할 새로운 지역 그룹 및 대상 데이터 노드를 선택하는 주체는 클라이언트일 수도 있고 프락시 서버일 수도 있다. 즉, 클라이언트가 지역 그룹 및 대상 데이터 노드를 모두 선택할 수도 있고, 클라이언트는 지역 그룹만을 선택하고, 대상 데이터 노드는 프락시 서버에 의하여 선택될 수도 있다. 다른 실시예에서는, 지역 그룹 및 대상 데이터 노드 모두가 프락시 서버 또는 위치 인식 서버에 의하여 자동으로 선택될 수도 있다.
- [0056] 이와 같이, 지역 그룹 및 대상 데이터 노드가 결정되면, 대상 오브젝트의 저장 및 복제 동작이 수행된다. 그러면, 프락시 서버가 새로운 대상 데이터 노드의 목록을 이용하여 메타데이터 데이터베이스를 갱신한다.
- [0057] 본 발명에 의한 분산 저장 방법에서, 오브젝트의 고유 정보는 오브젝트의 ID, 크기, 데이터 타입, 및 작성자 등을 포함할 수 있고, 데이터 노드의 고유 정보는 데이터 노드의 ID, IP(Internet Protocol) 주소, 및 물리적 위치 등을 포함할 수 있다.
- [0058] 이하, 본 발명에 의한 분산 저장 방법에서 수행되는 복제 단계를 상세히 설명한다.
- [0059] 도 5는 본 발명의 다른 면에 의한 분산 저장 시스템의 오브젝트 저장 방법에 포함되는 복제 단계를 나타내는 흐름도이다.
- [0060] 지역 그룹 및 대상 데이터 노드가 결정되면, 클라이언트는 결정된 지역 그룹 중 가장 높은 우선 순위를 가지는 지역 그룹에 속하는 대상 데이터 노드에 대상 오브젝트를 저장한다(S510). 그러면, 가장 높은 우선 순위를 가지는 지역 그룹에 저장된 대상 오브젝트를 어느 장치로 하여금 복제할지를 결정한다(S520).
- [0061] 본 발명에 의한 분산 저장 방법에서 이용될 수 있는 복제 방법으로는 세 가지가 존재하는데, 이는 다음과 같다.
- [0062] 첫째로, 클라이언트가 업로드할 오브젝트를 n개의 복제본을 위해 자신에게 할당된 n개의 데이터 노드들에게 모두 전송하는 것이다(S530). 이 경우, n 개의 대상 데이터 노드가 병렬적으로 n개의 복제본을 저장할 수 있기 때문에 트랜잭션 시간이 단축될 수 있다. 다만, 첫째 방법의 경우 클라이언트가 n 개의 대상 데이터 노드에 한꺼번에 오브젝트를 전송하기 때문에 클라이언트의 전송 대역폭이 과다하게 소요되는 단점이 있다.
- [0063] 둘째로, 파이프라인 방식이다. 이 방식에 따르면, 특정 우선 순위를 가지는 지역 그룹에 속하는 대상 데이터 노드는 자신이 수신하여 저장한 오브젝트를 자신이 속한 지역 그룹의 우선 순위보다 한 단계 낮은 우선 순위를 가지는 지역 그룹에 속하는 대상 데이터 노드에 송신한다. 그러면, 차순위 지역 그룹에 속한 대상 데이터 노드가 이 대상 오브젝트를 수신하여 저장하고, 다시 자신보다 한 단계 낮은 우선 순위를 가지는 지역 그룹에 속하는 대상 데이터 노드에 전송한다(S550). 이때, 대상 데이터 노드는 자신이 속한 보다 낮은 우선 순위를 가지는 지역 그룹에 속한 모든 대상 데이터 노드의 목록도 함께 전송한다. 어느 대상 데이터 노드가 대상 오브젝트를 수신하면, 해당 대상 데이터 노드는 자신이 우선 순위의 최하위에 위치하는지 여부를 판단한다(S560). 만일 자신이 속한 지역 그룹이 가장 낮은 우선 순위를 가지지 않는다면, 아직 대상 오브젝트를 저장할 다른 대상 데이터 노드들이 존재한다는 것을 의미한다. 그러므로, 우선 순위를 낮춘 후(S570) 위의 동작을 반복한다. 이 경우, n개의 대상 데이터 노드가 존재할 경우, n-1번의 복제 동작이 수행되어야 하기 때문에 클라이언트가 전체 트랜잭션에 시간이 지연되는 단점이 있다. 하지만, 파이프라인 방식에 따르면 클라이언트의 부하를 많이 감소해 주고 클라이언트에서의 전송 트래픽을 줄여주는 장점이 있다.
- [0064] 셋째로, 클라이언트는 자신이 가장 선호하는 첫 번째 데이터 노드에만 대상 오브젝트를 전송하고, 나머지 대상 데이터 노드들의 복제 동작은 복제를 전담하는 복제 서버가 수행할 수 있다(S540).
- [0065] 도 6은 본 발명의 일면에 의한 분산 저장 시스템의 다른 실시예를 개념적으로 나타내는 도면이다.
- [0066] 도 6에 도시된 분산 저장 시스템(600)은 네트워크(690)에 연결되는 복수 개의 클라이언트들(610, 612, 616) 및 데이터 노드들(DN11-DN1n, DN21-DN2n, DNm1-DNm)을 포함한다. 또한, 도 6에 도시된 분산 저장 시스템(600)은 인증 서버(620), 프락시 서버(650), 위치 인식 서버(660), 복제 서버(670) 및 메타데이터 데이터베이스(680)를 더 포함한다.
- [0067] 도 6에 도시된 클라이언트들(610, 612, 616), 인증 서버(620) 및 메타데이터 데이터베이스(680)의 구성 및 동작은 도 2에 도시되는 대응 구성 요소들의 그것과 유사하다. 그러므로, 명세서의 간략화를 위하여 반복적인 설명이 생략된다.
- [0068] 도 6에 도시된 분산 저장 시스템(600)에 포함되는 위치 인식 서버(660)는 지역 그룹 또는 대상 데이터 노드를

자동으로 선택하기 위하여 이용된다. 인증된 클라이언트가 대상 오브젝트를 저장할 대상 데이터 노드에 대해서 프락시 서버(650)에 질의하면, 프락시 서버(650)는 가장 유리한 지역 그룹에 대해서 위치 인식 서버(660)에 문의한다.

[0069] 위치 인식 서버(660)는 여러 가지 방법으로 클라이언트의 위치를 파악할 수 있는데, 일반적으로는 클라이언트의 IP 주소로 클라이언트의 물리적 위치를 파악할 수 있다. 위치 인식 서버(660)는 프락시 서버(650)의 요청에 따라 클라이언트의 기본 복제본 개수만큼의 지역 그룹을 선정한 후 선정된 지역 그룹 리스트를 프락시 서버(650)에게 전송한다. 위치 인식 서버(660)는 물리적으로 프락시 서버(650) 내에 통합되어 구현될 수 있다.

[0070] 위치 인식 서버(660)에 의하여 결정된 지역 그룹 각각에 속하는 대상 데이터 노드를 결정하는 동작은 프락시 서버(650) 또는 위치 인식 서버(660)에 의하여 수행될 수 있다. 위치 인식 서버(660)가 대상 데이터 노드까지 결정하는 경우, 위치 인식 서버(660)는 메타데이터 데이터베이스(680)를 참조하여 선택된 지역 그룹 내에서 대상 오브젝트를 가지고 있는 클라이언트와 가장 근접한 대상 데이터 노드를 선택할 수 있다. 반면에, 프락시 서버(650)가 대상 데이터 노드를 선택한다면, 프락시 서버(650)는 로드 밸런서(load balancer, 655)를 이용하여 각 지역 그룹에 속하는 데이터 노드들의 상태를 점검하고, 이 중에서 최적 조건을 가지는 데이터 노드를 대상 데이터 노드로서 선택할 수 있다. 로드 밸런서(655)는 프락시 서버(650)에 포함되는 것으로 도시되었으나, 이는 본 발명을 한정하는 것이 아님이 이해되어야 한다.

[0071] 또한, 프락시 서버(650)는 각 지역 그룹 내 데이터 노드들의 정보를 메타데이터에 관리하고, 데이터 노드들의 로드 밸런싱을 위해 각 데이터 노드의 스토리지 용량을 고려하여 각 데이터 노드의 가중치 값을 사전에 결정한다. 현재까지 각 데이터 노드의 오브젝트 저장 내역과 데이터 노드의 가중치 값을 고려해 요청 클라이언트의 데이터 노드를 선정함으로써 지역 그룹 내 데이터 노드간 로드 밸런싱을 유지한다.

[0072] 본 발명은 도면에 도시된 실시예를 참고로 설명되었으나 이는 예시적인 것에 불과하며, 본 기술 분야의 통상의 지식을 가진 자라면 이로부터 다양한 변형 및 균등한 타 실시예가 가능하다는 점을 이해할 것이다.

[0073] 또한, 본 발명에 따르는 방법은 컴퓨터로 읽을 수 있는 기록 매체에 컴퓨터가 읽을 수 있는 코드로서 구현하는 것이 가능하다. 컴퓨터가 읽을 수 있는 기록 매체는 컴퓨터 시스템에 의하여 읽혀질 수 있는 데이터가 저장되는 모든 종류의 기록 장치를 포함할 수 있다. 컴퓨터가 읽을 수 있는 기록 매체의 예로는 ROM, RAM, CD-ROM, 자기 테이프, 플로피 디스크, 광 데이터 저장 장치 등이 있으며, 또한 캐리어 웨이브(예를 들어 인터넷을 통한 전송)의 형태로 구현되는 것도 포함한다. 또한 컴퓨터가 읽을 수 있는 기록 매체는 네트워크로 연결된 분산 컴퓨터 시스템에 의하여 분산 방식으로 실행될 수 있는 컴퓨터가 읽을 수 있는 코드를 저장할 수 있다.

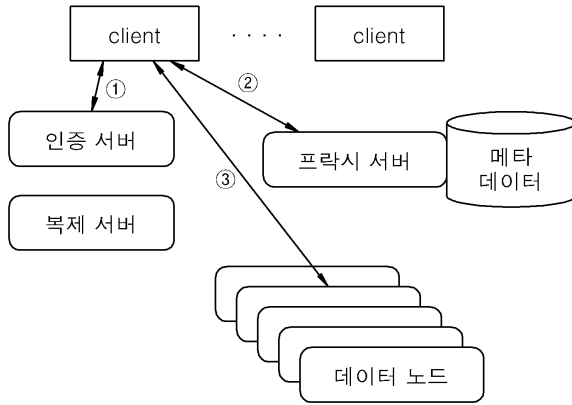
[0074] 따라서, 본 발명의 진정한 기술적 보호 범위는 첨부된 등록청구범위의 기술적 사상에 의해 정해져야 할 것이다.

**산업상 이용가능성**

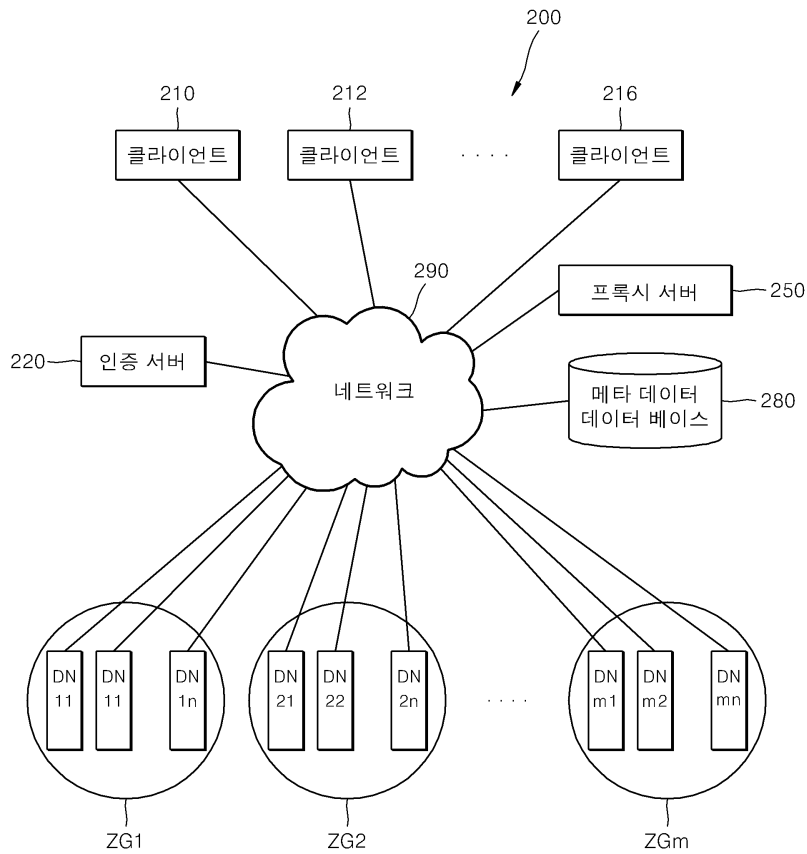
[0075] 본 발명은 클라우드 컴퓨팅의 스토리지 서비스를 제공할 수 있는 오브젝트 스토리지에서 클라이언트의 위치를 파악하여 가장 가까운 데이터 노드에 오브젝트를 배치하는 기술에 적용될 수 있다.

도면

도면1



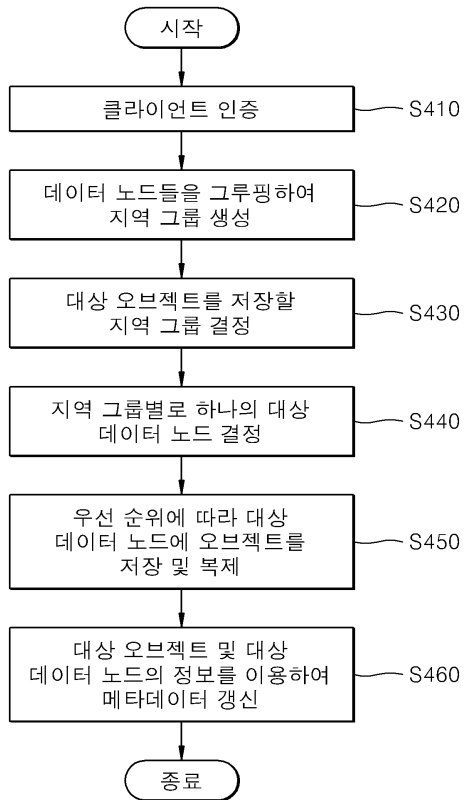
도면2



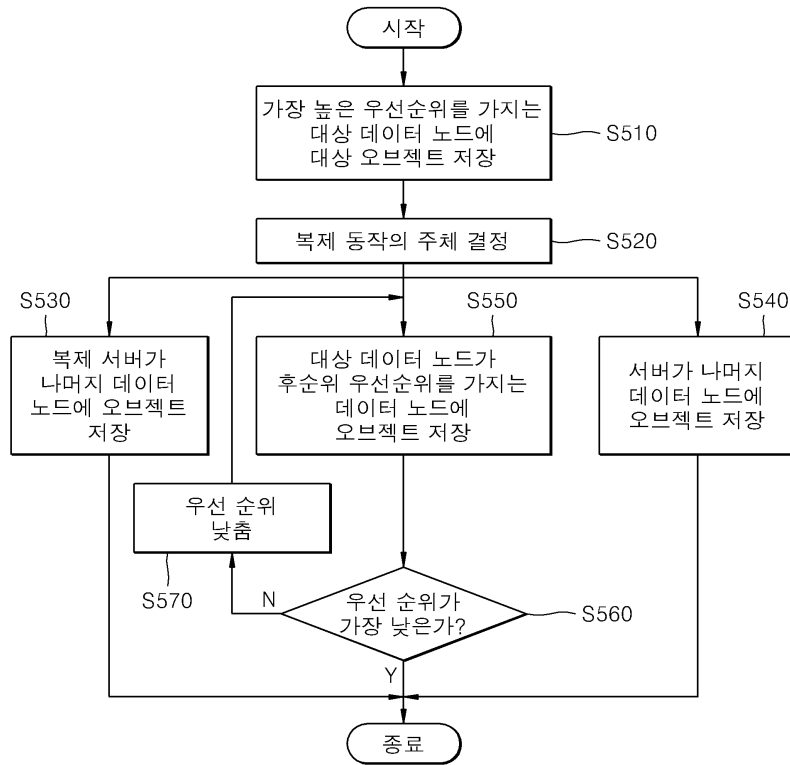
도면3

	저장 동작 (storage process)		지역 그룹 변경 동작 (zone group change process)	
	지역 그룹 선택	대상 데이터 노드 선택	지역 그룹 선택	대상 데이터 노드 선택
알고리즘 1	클라이언트	클라이언트	클라이언트	클라이언트
알고리즘 2	클라이언트	프락시 서버	클라이언트	프락시 서버
알고리즘 3	위치 인식 서버 (프락시 서버)	위치 인식 서버 (프락시 서버)	(자동 선택 방식에서는 일반적으로 지역 그룹 변경이 불필요하나, 자동/수동 겸용 선택 방식에서는 수행 가능)	

도면4



도면5



도면6

