



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2017년05월29일
(11) 등록번호 10-1736406
(24) 등록일자 2017년05월10일

(51) 국제특허분류(Int. Cl.)
G06F 17/30 (2006.01)
(52) CPC특허분류
G06F 17/30339 (2013.01)
G06F 17/30094 (2013.01)
(21) 출원번호 10-2016-0066375
(22) 출원일자 2016년05월30일
심사청구일자 2016년05월30일
(56) 선행기술조사문헌
KR101331350 B1
KR1020130049111 A
US20110035359 A1

(73) 특허권자
주식회사 인피니플렉스
서울특별시 구로구 디지털로31길 19, 703호(구로동, 에이스테크노타워2차)
(72) 발명자
김성진
서울특별시 구로구 디지털로 273, 904호 (구로동)
(74) 대리인
리앤목록특허법인

전체 청구항 수 : 총 7 항

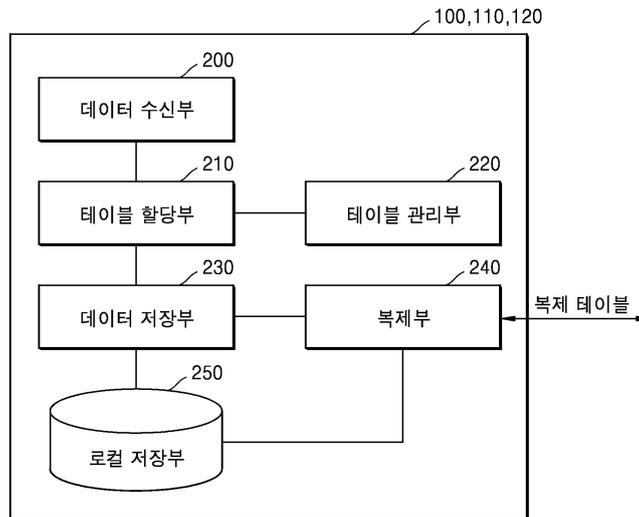
심사관 : 최정권

(54) 발명의 명칭 데이터 복제 방법 및 그 장치

(57) 요약

데이터 복제 방법 및 그 장치가 개시된다. 마스터 노드 또는 데이터 노드는, 복수 개의 데이터 영역과 복수 개의 인덱스 영역으로 구성되는 테이블 영역에 데이터와 인덱스를 저장하고, 복수 개의 인덱스 영역을 하나의 인덱스 영역으로 병합한 후 복수 개의 데이터 영역과 하나의 인덱스 영역으로 구성된 테이블 영역을 다른 노드에게 복제한다.

대표도 - 도2



(52) CPC특허분류

G06F 17/30194 (2013.01)

G06F 17/30212 (2013.01)

이 발명을 지원한 국가연구개발사업

과제고유번호 R-20150510 - 001026

부처명 미래창조과학부

연구관리전문기관 정보통신기술진흥센터

연구사업명 ICT 유망기술개발 지원사업

연구과제명 고가용성 및 고확장성을 지원하는 다중 노드 클러스터 Time-Series (시계열) DBMS 개발

기 여 율 1/1

주관기관 (주)인피니플릭스

연구기간 2015.06.01 ~ 2017.05.31

명세서

청구범위

청구항 1

마스터 노드 및 적어도 하나 이상의 데이터 노드로 구성된 시계열 데이터베이스의 데이터 복제 방법에 있어서, 데이터베이스 시스템을 구성하는 노드 장치는,

복수 개의 데이터 영역과 복수 개의 인덱스 영역으로 구성되는 테이블 영역에 데이터와 인덱스를 저장하는 단계;

상기 복수 개의 인덱스 영역을 하나의 인덱스 영역으로 병합하는 단계; 및

상기 복수 개의 데이터 영역과 하나의 인덱스 영역으로 구성된 테이블 영역을 다른 노드에게 복제하는 단계;를 수행하는 것을 특징으로 하는 데이터 복제 방법.

청구항 2

제 1항에 있어서, 상기 저장하는 단계는,

테이블 영역을 할당받는 단계; 및

상기 할당받은 테이블 영역에 데이터와 인덱스를 저장하는 단계;를 포함하는 것을 특징으로 하는 데이터 복제 방법.

청구항 3

제 1항에 있어서,

상기 인덱스는 비트맵 방식으로 일정 개수의 데이터의 저장 위치 및 데이터 값에 대한 정보를 포함하는 것을 특징으로 하는 데이터 복제 방법.

청구항 4

제 3항에 있어서, 상기 병합하는 단계는,

상기 복수 개의 인덱스 영역에 각각 저장된 비트맵들을 하나의 비트맵으로 병합하여 하나의 인덱스 영역에 저장하는 단계;를 포함하는 것을 특징으로 하는 데이터 복제 방법.

청구항 5

제 1항에 있어서,

상기 복수 개의 인덱스 영역에 저장된 각 인덱스는 LSMT(Log Structured Merge Tree) 인덱스인 것을 특징으로 하는 데이터 복제 방법.

청구항 6

데이터를 수신하는 데이터수신부;

복수 개의 데이터 영역과 복수 개의 인덱스 영역으로 구성되는 테이블 영역에 데이터와 인덱스를 저장하는 데이터저장부; 및

상기 복수 개의 인덱스 영역을 하나의 인덱스 영역으로 병합하고, 상기 복수 개의 데이터 영역과 상기 하나의 인덱스 영역으로 구성된 테이블 영역을 다른 노드에게 복제하는 복제부;를 포함하는 것을 특징으로 하는 데이터베이스의 노드 장치.

청구항 7

제 1항 내지 제 5항 중 어느 한 항에 기재된 방법을 수행하기 위한 프로그램을 기록한 컴퓨터로 읽을 수 있는

기록매체.

발명의 설명

기술 분야

[0001] 본 발명은 멀티 노드 환경에서 데이터베이스의 데이터를 복제하는 방법 및 그 장치에 관한 것이다.

배경 기술

[0002] 센서 및 장비로부터 발생하는 각종 데이터는 시계열적으로 발생하며, 이러한 시계열 데이터를 실시간으로 저장하고 분석하는 데이터베이스를 시계열 데이터베이스라고 한다.

[0003] 일반적으로 시계열 데이터베이스는 저장된 데이터의 변경, 즉 갱신(update) 연산이 거의 발생하지 않으며, 단지 데이터의 추가(insert), 삭제(delete) 또는 선택(select) 등의 연산이 주로 발생한다.

[0004] 시계열 데이터베이스는 장애 등에 따른 데이터 손실을 방지하기 위하여, 저장된 데이터를 적어도 하나 이상의 다른 곳에 복제하여 저장하는데, 하나의 시계열 데이터베이스가 시계열 데이터의 수신 및 복제 등을 모두 수신하는 경우에 부하가 집중되어 데이터 처리 속도가 늦어지는 등 성능 저하의 문제가 발생한다.

선행기술문헌

특허문헌

[0005] (특허문헌 0001) 특허출원번호 10-2012-0051558

발명의 내용

해결하려는 과제

[0006] 본 발명이 이루고자 하는 기술적 과제는, 데이터베이스의 특정 노드에 부하가 집중되는 것을 방지하고, 복제에 따른 소요시간을 단축시킬 수 있는 데이터 복제 방법 및 그 장치를 제공하는 데 있다.

과제의 해결 수단

[0007] 상기의 기술적 과제를 달성하기 위한, 본 발명에 따른 데이터 복제 방법의 일 예는, 마스터 노드 및 적어도 하나 이상의 데이터 노드로 구성된 시계열 데이터베이스의 데이터 복제 방법에 있어서, 상기 마스터 노드 또는 데이터 노드는, 복수 개의 데이터 영역과 복수 개의 인덱스 영역으로 구성되는 테이블 영역에 데이터와 인덱스를 저장하는 단계; 상기 복수 개의 인덱스 영역을 하나의 인덱스 영역으로 병합하는 단계; 및 상기 복수 개의 데이터 영역과 하나의 인덱스 영역으로 구성된 테이블 영역을 다른 노드에게 복제하는 단계;를 포함한다.

[0008] 상기의 기술적 과제를 달성하기 위한, 본 발명에 따른 데이터베이스의 노드 장치의 일 예는, 데이터를 수신하는 데이터수신부; 복수 개의 데이터 영역과 복수 개의 인덱스 영역으로 구성되는 테이블 영역에 데이터와 인덱스를 저장하는 데이터저장부; 및 상기 복수 개의 인덱스 영역을 하나의 인덱스 영역으로 병합하고, 상기 복수 개의 데이터 영역과 상기 하나의 인덱스 영역으로 구성된 테이블 영역을 다른 노드에게 복제하는 복제부;를 포함한다.

발명의 효과

[0009] 본 발명에 따르면, 복수의 노드가 분산하여 데이터를 저장하므로, 데이터베이스 시스템의 부하를 분산시킬 수 있다. 또한, 각 노드는 데이터가 일정 이상 쌓이면 이를 다른 노드에 복제할 때 인덱스 파일을 하나의 파일로 만들어 복제를 수행하므로 복제에 따른 소요시간을 절감할 수 있다.

도면의 간단한 설명

[0010] 도 1은 본 발명에 따른 복수의 노드로 구성된 데이터베이스 시스템의 일 예를 도시한 도면,
 도 2는 본 발명에 따른 데이터베이스 시스템을 구성하는 노드의 일 예의 구성을 도시한 도면,

도 3은 본 발명에 따른 일반 테이블의 일 실시 예를 도시한 도면,
 도 4는 본 발명에 따른 복제 테이블의 일 실시 예를 도시한 도면,
 도 5는 본 발명에 따른 비트맵 인덱스의 병합 예를 도시한 도면, 그리고,
 도 6은 본 발명에 따른 데이터베이스 복제 방법의 일 예를 도시한 도면이다.

발명을 실시하기 위한 구체적인 내용

- [0011] 이하에서, 첨부된 도면들을 참조하여 본 발명에 따른 데이터 복제 방법 및 그 장치에 대해 상세히 설명한다.
- [0012] 도 1은 본 발명에 따른 복수의 노드로 구성된 데이터베이스 시스템의 일 예를 도시한 도면이다.
- [0013] 도 1을 참조하면, 데이터베이스 시스템은 마스터 노드(100)와 적어도 하나 이상의 데이터 노드(110,120)로 구성된다. 마스터 노드(100)와 데이터 노드(110,120)는 각각 테이블을 할당받아 데이터를 저장하며, 또한 입력된 데이터가 일정 이상 쌓이면 이를 다른 노드에 복제한다.
- [0014] 본 실시 예는 시계열 데이터를 저장하는 시계열 데이터베이스에 적용될 수 있으며, 이 경우 데이터베이스의 갱신 과정이 거의 일어나지 아니하므로 복수 개의 인덱스 파일을 병합하여 파일 입출력(I/O)의 오버헤드를 줄여 빠른 복제를 수행할 수 있다.
- [0015] 또한 본 실시 예의 데이터에는 시간에 따라 순차적으로 증가 또는 감소하는 형태의 데이터 식별자(RID)가 부여될 수 있다. 예를 들어, 데이터 식별자는 년월일시간분초(나노초) 등과 같은 시간 기반으로 생성될 수 있다.
- [0016] 도 2는 본 발명에 따른 데이터베이스 시스템을 구성하는 노드의 일 예의 구성을 도시한 도면이다.
- [0017] 도 2를 참조하면, 마스터 노드(100) 또는 데이터 노드(110,120)는 데이터 수신부(200), 테이블 할당부(210), 테이블 관리부(220), 데이터 저장부(230), 복제부(240) 및 로컬 저장부(250)를 포함한다.
- [0018] 데이터 수신부(200)는 외부로부터 데이터를 수신한다.
- [0019] 테이블 할당부(210)는 테이블 관리부(220)로부터 테이블 영역을 할당받는다.
- [0020] 테이블 관리부(220)는 각 테이블을 어떤 노드에 할당할지에 대한 정보를 가지고 있다. 예를 들어, 마스터 노드에게 제1,3 테이블을 할당하고, 제1 데이터 노드에게 제2,4 테이블을 할당하고, 제2 데이터 노드에게 제5,6 테이블을 할당하는 내용이 테이블 관리부(220)에 미리 설정되어 있다.
- [0021] 실시 예에 따라, 테이블 관리부(220)는 마스터 노드(100)에만 존재할 수 있다. 이 경우, 데이터 노드(110,120)의 테이블 할당부(210)는 마스터 노드(100)의 테이블 관리부(220)에게 테이블 할당을 요청하여 수신할 수 있다. 다른 예로, 마스터 노드(100)를 비롯한 각 데이터 노드(110,120)가 모두 동일한 테이블 할당 정보를 포함하는 테이블 관리부(220)를 포함할 수 있다.
- [0022] 데이터 저장부(230)는 테이블 할당부(210)에 의해 할당받은 테이블 영역에 데이터를 저장한다. 데이터가 저장되는 일반 테이블은 복수의 데이터 영역 및 복수의 인덱스 영역으로 구성된다. 각 데이터 영역은 각 데이터 파일로 구성되고, 각 인덱스 영역은 각 인덱스 파일로 구성된다.
- [0023] 데이터 저장부(230)는 데이터 영역에 데이터를 저장하고, 인덱스 영역에 해당 데이터의 저장 위치를 포함하는 각종 인덱스 정보를 저장한다. 데이터가 저장되는 일반 테이블의 일 예는 도 3에 도시되어 있다.
- [0024] 복제부(240)는 테이블에 일정 크기 이상의 데이터가 저장되면, 로컬 저장부(250)에 저장된 데이터를 다른 노드에게 복제한다. 테이블을 복제할 대상 노드는 다양한 방법에 의해 동적 또는 정적으로 설정될 수 있으며, 본 실시 예에서는 각 노드별 테이블의 복제 대상 노드가 미리 설정되어 있다고 가정한다.
- [0025] 복제부(240)는 데이터를 다른 노드에 복제할 때 로컬 저장부에 저장된 일반 테이블 그 자체를 이용하여 복제하는 것이 아니라, 테이블 내 존재하는 복수 개의 인덱스 영역을 하나의 영역을 병합하고, 이를 이용하여 복제를 수행한다. 다시 말해, 복제부(240)는 복수 개의 인덱스 파일을 하나의 파일로 병합한 후 병합된 하나의 인덱스 파일을 이용하여 각 데이터 영역의 데이터를 복제하므로, I/O 오버헤드를 줄일 수 있다. 인덱스 영역이 병합된 복제 테이블의 일 예는 도 4에 도시되어 있다. 인덱스 영역을 통합하는 방법에 대해서는 도 5에서 다시 설명한다. 복제부(240)는 다른 노드로부터 복제 테이블을 전달받아 이를 로컬 저장부(250)에 저장할 수 있다.
- [0026] 로컬 저장부(250)는 수신한 데이터를 저장하는 도 3과 같은 일반 테이블(300)과 다른 노드로부터 전달받은 도 4

와 같은 복제 테이블(400)을 포함한다.

- [0027] 도 3은 본 발명에 따른 일반 테이블의 일 실시 예를 도시한 도면이다.
- [0028] 도 3을 참조하면, 각 노드가 수신한 데이터를 저장하는 일반 테이블은 복수 개의 데이터 영역(310,312,314)과 복수 개의 인덱스 영역(320,322,324)을 포함한다. 본 실시 예는 설명의 편의를 위하여 영역이라고 표현하고 있으나, 각 데이터 영역(310,312,314) 및 각 인덱스 영역(320,322,324)은 각각의 파일로 구성되며, 이하의 실시 예에서도 동일하다. 각 인덱스 영역은 LSM(Log Structured Merge)-Tree 인덱스 파일로 생성될 수 있다.
- [0029] 각 데이터 영역(310,312,314)은 일정 크기를 가지므로, 저장되는 데이터의 크기에 따라 각 데이터 영역에 포함되는 데이터의 개수는 서로 상이할 수 있다. 반면에 인덱스 영역(320,322,324)은 일정 개수의 인덱스를 포함한다.
- [0030] 따라서 각 데이터 영역(310,312,314)에 포함되는 데이터의 크기에 따라 데이터 영역(310,312,314)과 인덱스 영역(320,322,324)이 일대일 관계가 아닐 수 있다. 예를 들어, 제1 데이터 영역(310)과 제2 데이터 영역(312)의 크기는 동일한 데 각 데이터 영역에 저장되는 데이터 크기가 서로 달라서 제1 데이터 영역(310)에는 20개의 데이터가 저장되고, 제2 데이터 영역(312)에는 10개의 데이터가 저장될 수 있다. 각 인덱스 영역이 10개의 인덱스를 저장한다면, 제1 인덱스 영역(320)은 제1 데이터 영역(310)에 저장된 10개의 데이터에 대한 인덱스를 저장하고, 제2 인덱스 영역(322)은 제1 데이터 영역(312)의 나머지 10개의 데이터에 대한 인덱스를 저장한다.
- [0031] 이와 같은 경우에, 제1 데이터 영역의 데이터를 복제하기 위해서는 제1 인덱스 영역(320)에 해당하는 파일과 제2 인덱스 영역(322)에 해당하는 파일을 각각 참조하여야 하므로, 인덱스 파일의 I/O 오버헤드가 발생한다. 인덱스 영역이 1000개라면, 복제를 위하여 최소 1000개의 인덱스 파일을 참조하여야 한다.
- [0032] 따라서 본 실시 예는 도 3의 일반 테이블을 그대로 복제하는 것이 아니라 도 4의 복제 테이블을 이용하여 복제를 수행한다.
- [0033] 도 4는 본 발명에 따른 복제 테이블의 일 실시 예를 도시한 도면이다.
- [0034] 도 4를 참조하면, 복제 테이블(400)은 복수의 데이터 영역(310,312,314)과 하나의 인덱스 영역(410)으로 구성된다. 복수의 데이터 영역(310,312,314)은 도 3에서 살핀 복수의 데이터 영역과 동일하다.
- [0035] 인덱스 영역(410)은 도 3의 복수의 인덱스 영역(320,322,324)을 하나로 통합한 영역이다. 즉, 인덱스 영역(410)은 도 3의 복수의 인덱스 파일을 하나의 파일로 통합한 것이다. 그러나 도 3의 인덱스 영역(320,322,324)의 통합은 일반적인 문서 통합과 같이 단지 하나의 파일로 만드는 것이 아니라, 복제에 따른 소요시간을 줄이기 위하여 도 5와 같은 비트맵 형태로 구성된 인덱스 영역의 통합이다.
- [0036] 도 5는 본 발명에 따른 비트맵 인덱스의 병합 예를 도시한 도면이다.
- [0037] 도 5를 참조하면, 도 3의 각 인덱스 영역(320,322,324)에 포함된 인덱스는 비트맵 인덱스(500,510)이다. 비트맵 인덱스(500,510)는 데이터값과 각 데이터 식별자의 행렬로 구성되고 각 요소는 데이터의 존재 여부를 나타내는 0 또는 1의 값을 가진다.
- [0038] 데이터는 시간에 따라 순차적으로 증가 또는 감소하는 형태의 데이터 식별자를 가지므로, 각 인덱스 영역에 저장된 비트맵 인덱스(500,510)를 통합할 때, 데이터 식별자 구간을 늘려 하나의 비트맵 인덱스(500)로 통합할 수 있다. 예를 들어, 제1 비트맵 인덱스(500)에 저장된 데이터 식별자 구간이 RID_1 ~ RID_k 이고, 제2 비트맵 인덱스(510)에 저장된 데이터 식별자 구간이 RID_(k+1) ~ RID_m 인 경우에, 비트맵 인덱스의 열을 RID_1 ~ RID_m 으로 만들어 두 비트맵 인덱스를 통합할 수 있다.
- [0039] 도 6은 본 발명에 따른 데이터베이스 복제 방법의 일 예를 도시한 도면이다.
- [0040] 도 6을 참조하면, 데이터베이스 시스템을 구성하는 마스터 노드 또는 데이터 노드의 장치(이하, '노드 장치'라 함)는 각자 데이터를 수신한다(S600). 노드 장치는 할당받은 테이블 영역을 할당받아 데이터를 저장한다(S610, S620). 노드 장치는 도 3에서 살핀 복수의 데이터 영역과 복수의 인덱스 영역으로 구성된 일반 테이블을 이용하여 데이터를 저장한다.
- [0041] 노드 장치는 복제 조건을 만족하는지 파악한다(S630). 복제 조건은 실시 예에 따라 다양하게 미리 설정될 수 있으며, 일 예로 테이블에 일정 크기 이상의 데이터가 쌓이면 복제 조건을 만족하는 것으로 파악할 수 있다.
- [0042] 노드 장치는 복제를 위하여 도 4와 같이 복수 개의 인덱스 영역을 하나의 영역으로 통합한다(S640). 그리고 하

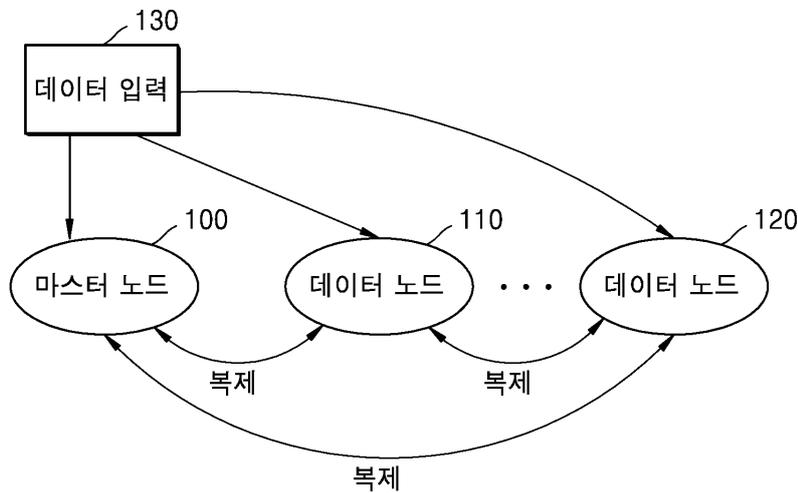
나의 인덱스 영역으로 통합된 테이블을 이용하여 다른 노드에 데이터를 복제한다(S650). 이 경우 노드 장치는 각 데이터 영역에 대한 인덱스 파일을 일일이 확인하는 것이 아니라 통합된 하나의 인덱스 파일을 이용하여 복제시 인덱스 파일의 I/O 오버헤드를 줄여 복제를 보다 빠르게 수행할 수 있다.

[0043] 본 발명은 또한 컴퓨터로 읽을 수 있는 기록매체에 컴퓨터가 읽을 수 있는 코드로서 구현하는 것이 가능하다. 컴퓨터가 읽을 수 있는 기록매체는 컴퓨터 시스템에 의하여 읽혀질 수 있는 데이터가 저장되는 모든 종류의 기록장치를 포함한다. 컴퓨터가 읽을 수 있는 기록매체의 예로는 ROM, RAM, CD-ROM, 자기 테이프, 플로피디스크, 광데이터 저장장치 등이 있다. 또한 컴퓨터가 읽을 수 있는 기록매체는 네트워크로 연결된 컴퓨터 시스템에 분산되어 분산방식으로 컴퓨터가 읽을 수 있는 코드가 저장되고 실행될 수 있다.

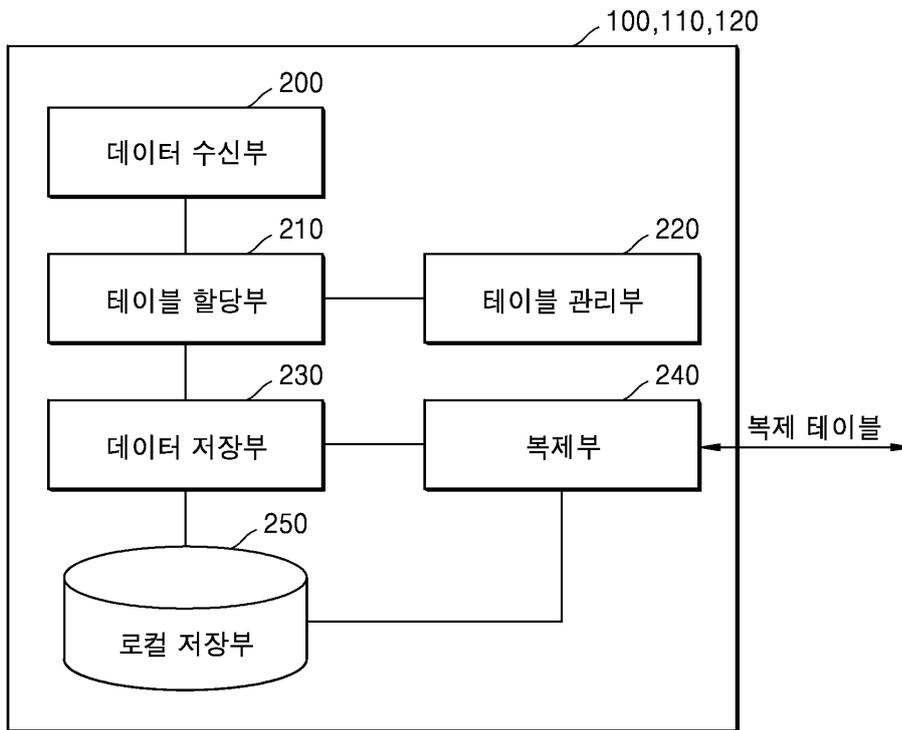
[0044] 이제까지 본 발명에 대하여 그 바람직한 실시예들을 중심으로 살펴보았다. 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자는 본 발명이 본 발명의 본질적인 특성에서 벗어나지 않는 범위에서 변형된 형태로 구현될 수 있음을 이해할 수 있을 것이다. 그러므로 개시된 실시예들은 한정적인 관점이 아니라 설명적인 관점에서 고려되어야 한다. 본 발명의 범위는 전술한 설명이 아니라 특허청구범위에 나타나 있으며, 그와 동등한 범위 내에 있는 모든 차이점은 본 발명에 포함된 것으로 해석되어야 할 것이다.

도면

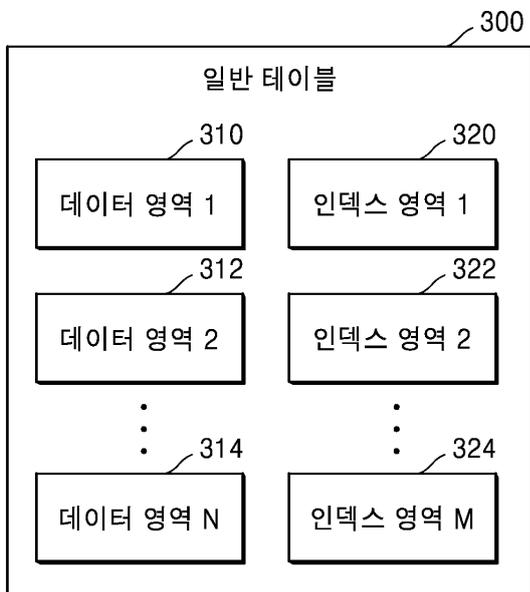
도면1



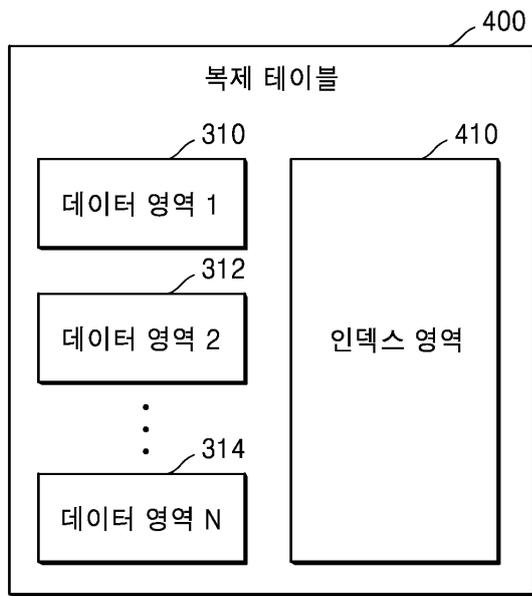
도면2



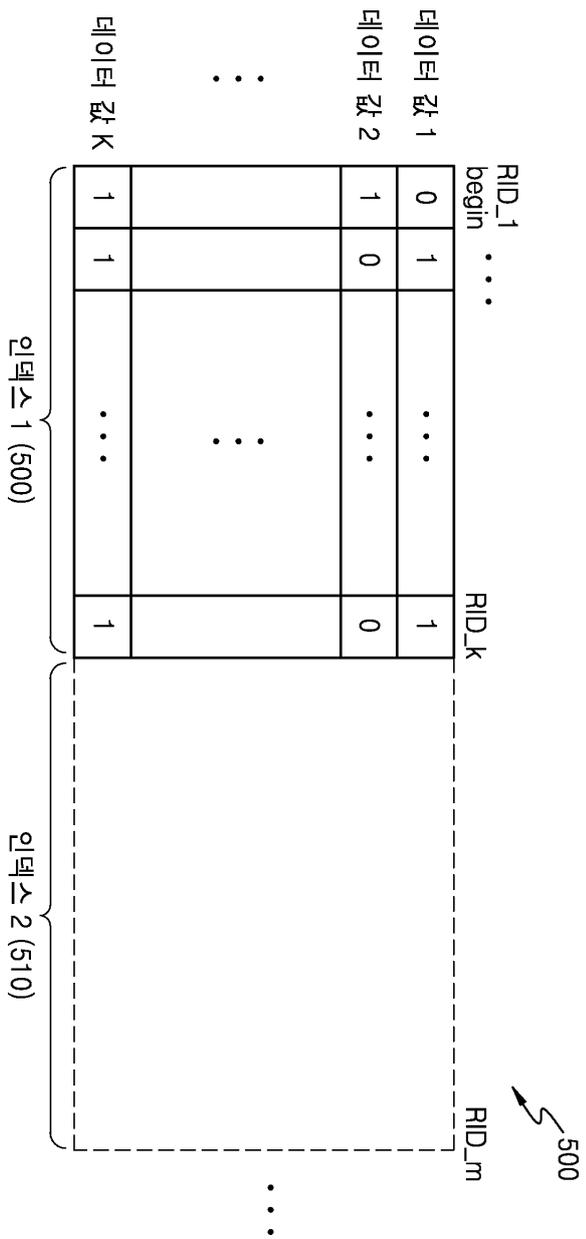
도면3



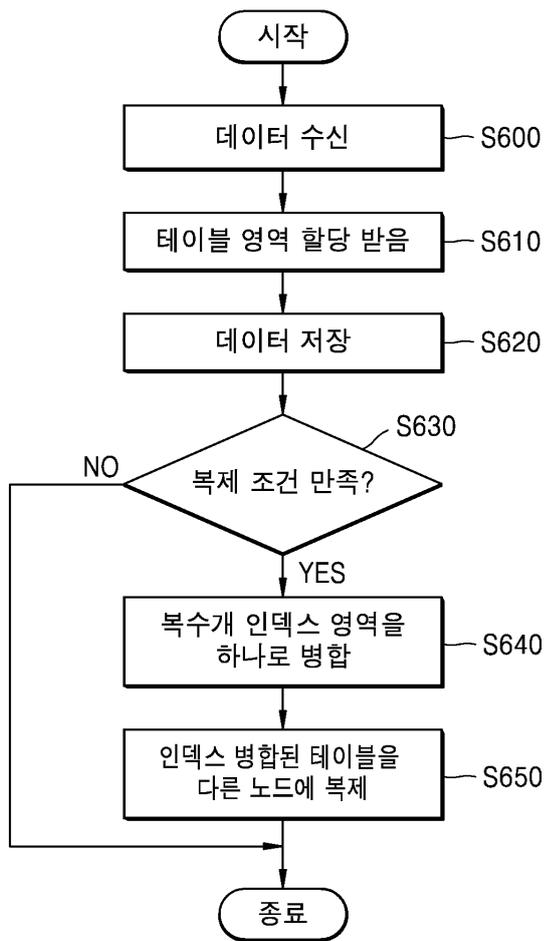
도면4



도면5



도면6



【심사관 직권보정사항】

【직권보정 1】

【보정항목】 청구범위

【보정세부항목】 청구항 1

【변경전】

포함하는 것을 특징으로 하는 데이터 복제 방법

【변경후】

수행하는 것을 특징으로 하는 데이터 복제 방법