



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2023-0171300
(43) 공개일자 2023년12월20일

(51) 국제특허분류(Int. Cl.)
G06F 16/21 (2019.01) G06F 16/22 (2019.01)
G06N 3/12 (2023.01)
(52) CPC특허분류
G06F 16/217 (2019.01)
G06F 16/2246 (2019.01)
(21) 출원번호 10-2022-0071698
(22) 출원일자 2022년06월13일
심사청구일자 2022년06월13일

(71) 출원인
연세대학교 산학협력단
서울특별시 서대문구 연세로 50 (신촌동, 연세대학교)
(72) 발명자
박상현
서울특별시 양천구 오목로 300, 204동 3701호(목동, 현대하이페리온2)
김휘균
서울특별시 서대문구 연세로5나길 14, 래지던스퍼스트 403호(창천동)
이지은
서울특별시 서대문구 동교로 291, 101동 1004호(연희동, 연희동임광아파트)
(74) 대리인
특허법인우인

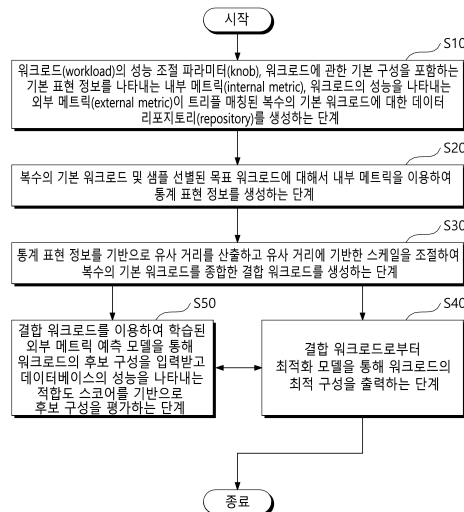
전체 청구항 수 : 총 14 항

(54) 발명의 명칭 **심층 유전 알고리즘을 사용한 데이터베이스 튜닝 방법**

(57) 요약

본 실시예들은 데이터베이스에서 가장 유사한 워크로드를 찾기 위해 워크로드에 대한 새로운 표현으로 정의하고, 결합 워크로드에 유전자 알고리즘을 적용하고 딥러닝을 통해 최적화된 파라미터를 도출하여 데이터베이스 성능을 향상시킬 수 있는 데이터베이스 튜닝 방법을 제공한다.

대표도 - 도2



(52) CPC특허분류

G06N 3/126 (2023.01)

이 발명을 지원한 국가연구개발사업

과제고유번호	1711152521
과제번호	2017-0-00477-006
부처명	과학기술정보통신부
과제관리(전문)기관명	정보통신기획평가원
연구사업명	정보통신방송연구개발사업
연구과제명	(SW스타랩) IoT 환경을 위한 고성능 플래시 메모리 스토리지 기반 인메모리 분산 DBMS 연구개발[2단계](2/2)
기 여 율	1/2
과제수행기관명	연세대학교 산학협력단
연구기간	2022.01.01 ~ 2022.12.31

이 발명을 지원한 국가연구개발사업

과제고유번호	1711152718
과제번호	2020-0-01361-003
부처명	과학기술정보통신부
과제관리(전문)기관명	정보통신기획평가원
연구사업명	정보통신방송혁신인재양성
연구과제명	인공지능대학원지원사업(3/5)
기 여 율	1/2
과제수행기관명	연세대학교 산학협력단
연구기간	2022.01.01 ~ 2023.12.31

명세서

청구범위

청구항 1

데이터베이스 튜닝 방법에 있어서,

워크로드(workload)의 성능 조절 파라미터(knob), 상기 워크로드에 관한 기본 구성을 포함하는 기본 표현 정보를 나타내는 내부 메트릭(internal metric), 상기 워크로드의 성능을 나타내는 외부 메트릭(external metric)이 트리플 매칭된 복수의 기본 워크로드에 대한 데이터 리포지토리(repository)를 생성하는 단계;

상기 복수의 기본 워크로드 및 샘플 선별된 목표 워크로드에 대해서 상기 내부 메트릭을 이용하여 통계 표현 정보를 생성하는 단계;

상기 통계 표현 정보를 기반으로 유사 거리를 산출하고 상기 유사 거리에 기반한 스케일을 조절하여 상기 복수의 기본 워크로드를 종합한 결합 워크로드를 생성하는 단계;

상기 결합 워크로드로부터 최적화 모델을 통해 상기 워크로드의 최적 구성을 출력하는 단계를 포함하는 데이터베이스 튜닝 방법.

청구항 2

제1항에 있어서,

상기 결합 워크로드를 이용하여 학습된 외부 메트릭 예측 모델을 통해 상기 워크로드의 후보 구성을 입력받고 데이터베이스의 성능을 나타내는 적합도 스코어를 기반으로 상기 후보 구성을 평가하는 단계를 포함하는 것을 특징으로 하는 데이터베이스 튜닝 방법.

청구항 3

제2항에 있어서,

상기 기본 구성은 워크로드 인덱스, 키 사이즈, 값 사이즈, 엔트리 개수, 읽기 쓰기 비율, 업데이트 정보, 또는 이들의 조합을 포함하는 것을 특징으로 하는 데이터베이스 튜닝 방법.

청구항 4

제2항에 있어서,

상기 내부 메트릭은 상기 기본 구성, 요청 데이터가 존재하지 않음을 나타내는 데이터베이스 캐시 미스, 또는 이들의 조합을 포함하는 것을 특징으로 하는 데이터베이스 튜닝 방법.

청구항 5

제2항에 있어서,

상기 외부 메트릭은 총 실행 시간, 데이터 처리 속도, 데이터 쓰기 증폭 인자, 공간 증폭 인자, 또는 이들의 조합을 포함하는 것을 특징으로 하는 데이터베이스 튜닝 방법.

청구항 6

제2항에 있어서,

상기 통계 표현 정보는 평균, 분산, 제1사분위수, 제2사분위수, 제3사분위수를 포함하는 것을 특징으로 하는 데이터베이스 튜닝 방법.

청구항 7

제2항에 있어서,

상기 유사 거리를 산출하는 것은 평균값으로부터의 거리가 표준편차의 몇 배인지를 나타내는 마할라노비스 거리(mahalanobis distance)를 적용하는 것을 특징으로 하는 데이터베이스 튜닝 방법.

청구항 8

제2항에 있어서,

상기 결합 워크로드를 생성하는 단계는,

상기 목표 워크로드에 대한 통계 표현 정보 및 상기 복수의 기본 워크로드에 대한 통계 표현 정보 간의 상기 유사 거리를 산출하여 유사 거리 매트릭스를 획득하고, 상기 유사 거리 매트릭스에 따른 최대 유사 거리를 기준으로 스케일을 조절하여 상기 복수의 기본 워크로드를 종합하며,

상기 결합 워크로드는 상기 성능 조절 파라미터, 상기 내부 메트릭, 상기 외부 메트릭이 트리플 매칭되는 것을 특징으로 하는 데이터베이스 튜닝 방법.

청구항 9

제5항에 있어서,

상기 적합도 스코어는 상기 총 실행 시간의 기본값과 예측값, 상기 데이터 처리 속도의 기본값과 예측값, 상기 데이터 쓰기 증폭 인자의 기본값과 예측값, 상기 공간 증폭 인자의 기본값과 예측값에 가중치를 적용한 후 합하여 산출되는 것을 특징으로 하는 데이터베이스 튜닝 방법.

청구항 10

제2항에 있어서,

상기 워크로드의 최적 구성을 출력하는 단계는,

상기 최적화 모델에 유전자 알고리즘을 적용하고,

상기 유전자 알고리즘의 유전자는 상기 워크로드의 성능 조절 파라미터에 해당하고, 상기 유전자 알고리즘의 염색체는 상기 워크로드의 기본 구성에 해당하도록 유전자 집단을 설정하고,

상기 외부 메트릭 예측 모델을 이용하여 산출한 상기 적합도 스코어에 따라 상기 유전자 집단의 순위를 결정하고,

상기 유전자 집단의 순위를 고려하여 상기 유전자 집단을 분류하고,

상기 성능 조절 파라미터에 대한 교차 및 변이를 적용하는 유전자 연산을 통해 상기 분류한 유전자 집단을 여러 세대를 거쳐 진화시켜서 상기 최적 구성을 찾는 것을 특징으로 하는 데이터베이스 튜닝 방법.

청구항 11

제10항에 있어서,

상기 외부 메트릭 예측 모델은,

상기 성능 조절 파라미터와 상기 기본 구성을 입력받고 복수의 선형 레이어가 연결되는 구조를 거쳐 상기 적합도 스코어를 산출하는 것을 특징으로 하는 데이터베이스 튜닝 방법.

청구항 12

프로세서를 포함하는 데이터베이스 튜닝 장치에 있어서,

상기 프로세서는,

워크로드(workload)의 성능 조절 파라미터(knob), 상기 워크로드에 관한 기본 구성을 포함하는 기본 표현 정보를 나타내는 내부 메트릭(internal metric), 상기 워크로드의 성능을 나타내는 외부 메트릭(external metric)이 트리플 매칭된 복수의 기본 워크로드에 대한 데이터 리포지토리(repository)를 생성하고,

상기 복수의 기본 워크로드 및 샘플 선별된 목표 워크로드에 대해서 상기 내부 메트릭을 이용하여 통계 표현 정

보를 생성하고,

상기 통계 표현 정보를 기반으로 유사 거리를 산출하고 상기 유사 거리에 기반한 스케일을 조절하여 상기 복수의 기본 워크로드를 종합한 결합 워크로드를 생성하고,

상기 결합 워크로드로부터 최적화 모델을 통해 상기 워크로드의 최적 구성을 출력하는 것을 특징으로 하는 데이터베이스 튜닝 장치.

청구항 13

제12항에 있어서,

상기 프로세서는,

상기 결합 워크로드를 이용하여 학습된 외부 메트릭 예측 모델을 통해 상기 워크로드의 후보 구성을 입력받고 데이터베이스의 성능을 나타내는 적합도 스코어를 기반으로 상기 후보 구성을 평가하는 것을 특징으로 하는 데이터베이스 튜닝 장치.

청구항 14

제13항에 있어서,

상기 프로세서는,

상기 최적화 모델에 유전자 알고리즘을 적용하고,

상기 유전자 알고리즘의 유전자는 상기 워크로드의 성능 조절 파라미터에 해당하고, 상기 유전자 알고리즘의 염색체는 상기 워크로드의 기본 구성에 해당하도록 유전자 집단을 설정하고,

상기 외부 메트릭 예측 모델을 이용하여 산출한 상기 적합도 스코어에 따라 상기 유전자 집단의 순위를 결정하고,

상기 유전자 집단의 순위를 고려하여 상기 유전자 집단을 분류하고,

상기 성능 조절 파라미터에 대한 교차 및 변이를 적용하는 유전자 연산을 통해 상기 분류한 유전자 집단을 여러 세대를 거쳐 진화시켜서 상기 최적 구성을 찾는 것을 특징으로 하는 데이터베이스 튜닝 장치.

발명의 설명

기술 분야

[0001] 본 발명이 속하는 기술 분야는 심층 유전 알고리즘을 이용한 로그 구조 병합 트리 기반의 데이터베이스를 튜닝하는 방법에 관한 것이다.

배경 기술

[0002] 이 부분에 기술된 내용은 단순히 본 실시예에 대한 배경 정보를 제공할 뿐 종래기술을 구성하는 것은 아니다.

[0003] 키-값 기반의 데이터베이스는 센서 데이터, 소셜 네트워크 데이터 등과 같이 비정형 데이터를 다루는데 유용하다. 키-값 기반의 데이터베이스는 로그 구조 병합 트리(Log Structured Merge Tree)를 주로 사용한다.

[0004] 로그 구조 병합 트리(Log Structured Merge Tree, LSM-Tree)는 연속적인 쓰기 연산을 수행하는 워크로드를 위해 설계되었다. LSM-Tree 구조는 하나의 인메모리 데이터 구조와 여러 개의 블록(ex. 디스크 등)에 저장을 위한 이어쓰기(Append) 방식의 데이터 구조로 이루어져 있다.

[0005] LSM-Tree는 키-값 데이터베이스에서 빈번히 발생하는 삽입 및 수정을 효율적으로 수행한다. 데이터를 우선 로그 형식으로 저장하고, 로그 상의 데이터 정렬, 수정 작업의 처리 등의 병합을 미루는 쓰기 친숙형 구조(Write Friendly Structure)이다. 하지만 나중에 발생하는 병합 동작은 쓰기 증폭을 발생시키며 시스템 성능과 저장장치 수명에 영향을 준다.

[0006] LSM-Tree는 임의적인 순서로 데이터를 쓰지 않고 순차적으로 데이터를 쓴다. 데이터를 조회할 때 주어진 데이터가 트리 내의 어느 위치에 있는지 알 수 없어서 데이터를 찾기 위해서 상위 레벨부터 순차적으로 검색해야

한다. 디스크에 데이터가 없어도 모든 레벨의 모든 파일을 읽어야 한다.

- [0007] 데이터베이스 시스템에는 일반적으로 고성능을 달성하기 위해 데이터베이스 관리자가 구성해야 하는 많은 톨이 있다. RocksDB는 로그 구조의 병합 트리를 사용하여 빠른 데이터 쓰기 성능을 달성한다. RocksDB는 중요한 성능 지표인 쓰기 및 공간 증폭과 관련된 많은 성능 조절 파라미터(노브)가 포함되어 있다. 데이터베이스 노브를 조정하여 상당한 성능 향상을 달성할 수 있다. 그러나 여러 개의 노브를 동시에 조정하는 것은 잠재적인 구성 조합과 절충안이 많은 문제가 있어 힘든 태스크에 해당한다.

선행기술문헌

특허문헌

- [0008] (특허문헌 0001) 미국공개특허공보 US 2017-0344619 (2017.11.30.)
(특허문헌 0002) 한국공개특허공보 KR 10-2016-0121819 (2016.10.21.)
(특허문헌 0003) 미국공개특허공보 US 2018-0121121 (2018.05.03.)

발명의 내용

해결하려는 과제

- [0009] 본 발명의 실시예들은 데이터베이스에서 가장 유사한 워크로드를 찾기 위해 워크로드에 대한 새로운 표현으로 정의하고, 결합 워크로드에 유전자 알고리즘을 적용하고 딥러닝을 통해 최적화된 파라미터를 도출하는 데 발명의 주된 목적이 있다.
- [0010] 본 발명의 명시되지 않은 또 다른 목적들은 하기의 상세한 설명 및 그 효과로부터 용이하게 추론할 수 있는 범위 내에서 추가적으로 고려될 수 있다.

과제의 해결 수단

- [0011] 본 실시예의 일 측면에 의하면, 데이터베이스 튜닝 방법에 있어서, 워크로드(workload)의 성능 조절 파라미터(knob), 상기 워크로드에 관한 기본 구성을 포함하는 기본 표현 정보를 나타내는 내부 메트릭(internal metric), 상기 워크로드의 성능을 나타내는 외부 메트릭(external metric)이 트리플 매칭된 복수의 기본 워크로드에 대한 데이터 리포지토리(repository)를 생성하는 단계; 상기 복수의 기본 워크로드 및 샘플 선별된 목표 워크로드에 대해서 상기 내부 메트릭을 이용하여 통계 표현 정보를 생성하는 단계; 상기 통계 표현 정보를 기반으로 유사 거리를 산출하고 상기 유사 거리에 기반한 스케일을 조절하여 상기 복수의 기본 워크로드를 종합한 결합 워크로드를 생성하는 단계; 상기 결합 워크로드로부터 최적화 모델을 통해 상기 워크로드의 최적 구성을 출력하는 단계를 포함하는 데이터베이스 튜닝 방법을 제공한다.
- [0012] 상기 데이터베이스 튜닝 방법은 상기 결합 워크로드를 이용하여 학습된 외부 메트릭 예측 모델을 통해 상기 워크로드의 후보 구성을 입력받고 데이터베이스의 성능을 나타내는 적합도 스코어를 기반으로 상기 후보 구성을 평가하는 단계를 포함할 수 있다.
- [0013] 상기 워크로드의 구성으로는 인덱스, 키 사이즈, 값 사이즈, 엔트리 개수, 읽기 쓰기 비율, 업데이트 정보, 또는 이들의 조합을 포함할 수 있다.
- [0014] 상기 내부 메트릭은 상기 기본 구성, 요청 데이터가 존재하지 않음을 나타내는 데이터베이스 캐시 미스, 또는 이들의 조합을 포함할 수 있다.
- [0015] 상기 외부 메트릭은 총 실행 시간, 데이터 처리 속도, 데이터 쓰기 증폭 인자, 공간 증폭 인자, 또는 이들의 조합을 포함할 수 있다.
- [0016] 상기 통계 표현 정보는 평균, 분산, 제1사분위수, 제2사분위수, 제3사분위수를 포함할 수 있다.
- [0017] 상기 유사 거리를 산출하는 것은 평균값으로부터의 거리가 표준편차의 몇 배인지를 나타내는 마할라노비스 거리(mahalanobis distance)를 적용할 수 있다.
- [0018] 상기 결합 워크로드를 생성하는 단계는, 상기 목표 워크로드에 대한 통계 표현 정보 및 상기 복수의 기본 워크

로드에 대한 통계 표현 정보 간의 상기 유사 거리를 산출하여 유사 거리 매트릭스를 획득하고, 상기 유사 거리 매트릭스에 따른 최대 유사 거리를 기준으로 스케일을 조절하여 상기 복수의 기본 워크로드를 종합하며, 상기 결합 워크로드는 상기 성능 조절 파라미터, 상기 내부 메트릭, 상기 외부 메트릭이 트리플 매칭될 수 있다.

[0019] 상기 적합도 스코어는 상기 총 실행 시간의 기본값과 예측값, 상기 데이터 처리 속도의 기본값과 예측값, 상기 데이터 쓰기 증폭 인자의 기본값과 예측값, 상기 공간 증폭 인자의 기본값과 예측값에 각각의 비례를 구하고 가중치를 적용한 후 합하여 산출될 수 있다.

[0020] 상기 워크로드의 최적 구성을 출력하는 단계는, 상기 최적화 모델에 유전자 알고리즘을 적용하고, 상기 유전자 알고리즘의 유전자는 상기 워크로드의 성능 조절 파라미터에 해당하고, 상기 유전자 알고리즘의 염색체는 상기 워크로드의 기본 구성에 해당하도록 유전자 집단을 설정하고, 상기 외부 메트릭 예측 모델을 이용하여 산출한 상기 적합도 스코어에 따라 상기 유전자 집단의 순위를 결정하고, 상기 유전자 집단의 순위를 고려하여 상기 유전자 집단을 분류하고, 상기 성능 조절 파라미터에 대한 교차 및 변이를 적용하는 유전자 연산을 통해 상기 분류한 유전자 집단을 여러 세대를 거쳐 진화시켜서 상기 최적 구성을 찾을 수 있다.

[0021] 상기 외부 메트릭 예측 모델은, 상기 성능 조절 파라미터와 상기 기본 구성을 입력받고 복수의 선형 레이어가 연결되는 구조를 거쳐 상기 적합도 스코어를 산출할 수 있다.

[0022] 본 실시예의 다른 측면에 의하면, 프로세서를 포함하는 데이터베이스 튜닝 장치에 있어서, 상기 프로세서는, 워크로드(workload)의 성능 조절 파라미터(knob), 상기 워크로드에 관한 기본 구성을 포함하는 기본 표현 정보를 나타내는 내부 메트릭(internal metric), 상기 워크로드의 성능을 나타내는 외부 메트릭(external metric)이 트리플 매칭된 복수의 기본 워크로드에 대한 데이터 리포지토리(repository)를 생성하고, 상기 복수의 기본 워크로드 및 샘플 선별된 목표 워크로드에 대해서 상기 내부 메트릭을 이용하여 통계 표현 정보를 생성하고, 상기 통계 표현 정보를 기반으로 유사 거리를 산출하고 상기 유사 거리에 기반한 스케일을 조절하여 상기 복수의 기본 워크로드를 종합한 결합 워크로드를 생성하고, 상기 결합 워크로드로부터 최적화 모델을 통해 상기 워크로드의 최적 구성을 출력하는 것을 특징으로 하는 데이터베이스 튜닝 장치를 제공한다.

[0023] 상기 프로세서는, 상기 결합 워크로드를 이용하여 학습된 외부 메트릭 예측 모델을 통해 상기 워크로드의 후보 구성을 입력받고 데이터베이스의 성능을 나타내는 적합도 스코어를 기반으로 상기 후보 구성을 평가할 수 있다.

[0024] 상기 프로세서는, 상기 최적화 모델에 유전자 알고리즘을 적용하고, 상기 유전자 알고리즘의 유전자는 상기 워크로드의 성능 조절 파라미터에 해당하고, 상기 유전자 알고리즘의 염색체는 상기 워크로드의 기본 구성에 해당하도록 유전자 집단을 설정하고, 상기 외부 메트릭 예측 모델을 이용하여 산출한 상기 적합도 스코어에 따라 상기 유전자 집단의 순위를 결정하고, 상기 유전자 집단의 순위를 고려하여 상기 유전자 집단을 분류하고, 상기 성능 조절 파라미터에 대한 교차 및 변이를 적용하는 유전자 연산을 통해 상기 분류한 유전자 집단을 여러 세대를 거쳐 진화시켜서 상기 최적 구성을 찾을 수 있다.

발명의 효과

[0025] 이상에서 설명한 바와 같이 본 발명의 실시예들에 의하면, 데이터베이스에서 가장 유사한 워크로드를 찾기 위해 워크로드에 대한 새로운 표현으로 정의하고, 결합 워크로드에 유전자 알고리즘을 적용하고 딥러닝을 통해 최적화된 파라미터를 도출하여 데이터베이스 성능을 향상시킬 수 있는 효과가 있다.

[0026] 여기에서 명시적으로 언급되지 않은 효과라 하더라도, 본 발명의 기술적 특징에 의해 기대되는 이하의 명세서에서 기재된 효과 및 그 잠정적인 효과는 본 발명의 명세서에 기재된 것과 같이 취급된다.

도면의 간단한 설명

[0027] 도 1은 로그 구조 병합 트리 기반의 데이터베이스가 데이터 컴팩션 동작을 수행하는 것을 예시한 도면이다.

도 2 및 도 3은 본 발명의 일 실시예에 따른 데이터베이스 튜닝 방법을 예시한 블록도이다.

도 4 및 도 5는 본 발명의 일 실시예에 따른 데이터베이스 튜닝 방법이 처리하는 워크로드의 구성을 예시한 도면이다.

도 6은 본 발명의 일 실시예에 따른 데이터베이스 튜닝 방법이 처리하는 워크로드의 기본 표현 정보를 예시한 도면이다.

도 7은 본 발명의 일 실시예에 따른 데이터베이스 튜닝 방법이 처리하는 워크로드의 통계 표현 정보를 예시한

도면이다.

도 8은 본 발명의 일 실시예에 따른 데이터베이스 튜닝 방법이 처리하는 결합 워크로드를 예시한 도면이다.

도 9 및 도 10은 본 발명의 일 실시예에 따른 데이터베이스 튜닝 방법이 처리하는 외부 메트릭 예측 모델을 예시한 도면이다.

도 11 및 도 12는 본 발명의 일 실시예에 따른 데이터베이스 튜닝 방법이 처리하는 최적화 모델을 예시한 도면이다.

도 13은 본 발명의 다른 실시예에 따른 데이터베이스 튜닝 장치를 예시한 도면이다.

도 14 내지 도 19는 실시예들에 따라 수행된 모의실험 결과를 도시한 것이다.

발명을 실시하기 위한 구체적인 내용

- [0028] 이하, 본 발명을 설명함에 있어서 관련된 공지기능에 대하여 이 분야의 기술자에게 자명한 사항으로서 본 발명의 요지를 불필요하게 흐릴 수 있다고 판단되는 경우에는 그 상세한 설명을 생략하고, 본 발명의 일부 실시예들을 예시적인 도면을 통해 상세하게 설명한다.
- [0029] 도 1은 로그 구조 병합 트리 기반의 데이터베이스가 데이터 컴팩션 동작을 수행하는 것을 예시한 도면이다.
- [0030] LSM-Tree를 이용한 대표적인 데이터베이스로는 RocksDB 등이 있다.
- [0031] LSM-Tree는 삽입 연산이 수행되면 먼저 메모리 영역에 데이터를 저장한다. 메모리의 일정 용량까지 데이터가 쌓이면 메모리의 내용을 디스크로 플러시(Flush)를 수행한다. 플러시되는 데이터는 디스크에 저장되어 있던 기존 데이터와 병합 정렬을 하여 기록된다. 디스크 영역의 각 레벨이 임계치를 넘으면 병합 정렬을 실행하여 하위 레벨을 생성한다.
- [0032] LSM-Tree 기반의 데이터베이스는 키-값 형태로 데이터를 저장한다. LSM-Tree 기반의 데이터베이스에 데이터의 삽입 연산 요청이 들어오면 데이터를 메모리에 기록하기 전에 우선적으로 로그 파일에 로그를 기록한다. 로그를 기록한 다음 메모리 영역에 있는 맴테이블(Memtable)에 데이터를 저장한다. 쓰기 요청이 계속되어 맴테이블(Memtable)에 데이터가 일정 용량까지 기록되면, 맴테이블(Memtable)은 변경이 불가능한 불변 맴테이블(Immutable Memtable, Read-Only Memtable)로 변경된다. 불변 맴테이블이 가득 차게 되면 블록(디스크) 영역으로 플러시가 발생한다.
- [0033] 플러시 동작을 수행하면, 맴테이블의 파일은 키 순서에 따라 정렬되어 SST(Stored String Table) 파일로 변경된다. SST 파일은 복수의 블록을 갖는다. 블록의 예시로는 데이터를 저장하는 데이터 블록(Data Block), 데이터 블록의 위치를 인덱싱하는 인덱스 블록(Index Block), 인덱스 블록의 위치를 처리하는 푸터 블록(Footer Block) 등이 있다.
- [0034] SST 파일은 디스크 영역에서 컴팩션(Compaction)을 통해 업데이트된다. 한 번 생성된 SST 파일은 사라지지 않을 수 있다. 하위 레벨에 상주하는 SST 파일일수록 상위 레벨의 SST 파일보다 오래된 데이터가 위치할 수 있다.
- [0035] 트랜잭션 수행 도중에 시스템 오류 또는 전원 차단 등과 같은 문제가 발생하면, 아직 디스크에 반영되지 않고 버퍼에 남아있는 데이터는 유실된다. 시스템이 재부팅된 후 데이터베이스가 복구를 수행할 때, 트랜잭션이 어떤 갱신 연산을 수행했는지 기록하는 로그를 사용한다. 로그 기록 방식으로는 WAL(Write-Ahead-Logging) 규칙이 있다. WAL은 트랜잭션으로 인해 변경된 데이터가 디스크에 기록되기 전에 관련된 로그를 로그 파일에 기록하는 규칙이다.
- [0036] LSM-Tree 기반의 데이터베이스는 두 개의 명령어를 수행한다. 하나는 메모리에서 디스크로 넘어가는 플러시 명령어이고, 다른 하나는 디스크의 레벨들을 조정하는 컴팩션 명령어이다.
- [0037] 플러시 명령어를 수행할 때, 불변 맴테이블은 단일 SST 파일로 변경된다. 대량의 데이터가 한꺼번에 입력되면, 플러시 속도의 균형을 맞추고 SST 파일의 레벨의 용량 한계치를 유지하기 위해서 의도적으로 플러시 속도를 조절한다. 이러한 의도된 지연을 'Write Stall'이라고 한다.
- [0038] LSM-Tree의 각각의 레벨은 특성이 구분된다. 컴팩션 비용과 디스크 쓰기는 특정 레벨에 집중된다. 계층적 저장 구조는 상위 레벨에서 하위 레벨로 점진적인 데이터 누적을 야기한다. SST 파일의 수명은 해당하는 레벨에서 존재하는 동안 컴팩션을 수행하는 횟수를 의미한다. 컴팩션을 수행하는 동안 SST 파일이 특정 레벨에서 삭제되지

않으면, 해당 SST 파일의 수명은 높게 나타난다.

- [0039] 컴팩션 명령어를 수행한 각 레벨의 결과를 살펴보면, 데이터 컴팩션을 수행하는 짧은 시간 동안에 상위 레벨의 SST 파일들은 생성되고 삭제됨을 나타낸다.
- [0040] 본 실시예는 로그 구조 병합 트리 기반의 데이터베이스에서 가장 유사한 워크로드를 신속하고 정확하게 찾기 위해 워크로드에 대한 새로운 표현으로 정의하고, 결합 워크로드에 유전자 알고리즘을 적용하고 딥러닝을 통해 최적화된 함수를 도출하여 데이터베이스 성능을 향상시킨다.
- [0041] 도 2 및 도 3은 본 발명의 일 실시예에 따른 데이터베이스 튜닝 방법을 예시한 블록도이다.
- [0042] 데이터베이스 튜닝 방법은 데이터베이스 튜닝 장치에 의해 수행될 수 있다.
- [0043] 단계 S10에서는 워크로드(workload)의 성능 조절 파라미터(knob), 워크로드에 관한 기본 구성을 포함하는 기본 표현 정보를 나타내는 내부 메트릭(internal metric), 워크로드의 성능을 나타내는 외부 메트릭(external metric)이 트리플 매칭된 복수의 기본 워크로드에 대한 데이터 리포지토리(repository)를 생성하는 단계를 수행할 수 있다.
- [0044] 단계 S20에서는 복수의 기본 워크로드 및 샘플 선별된 목표 워크로드에 대해서 내부 메트릭을 이용하여 통계 표현 정보를 생성하는 단계를 수행할 수 있다.
- [0045] 단계 S30에서는 통계 표현 정보를 기반으로 유사 거리를 산출하고 유사 거리에 기반한 스케일을 조절하여 복수의 기본 워크로드를 종합한 결합 워크로드를 생성하는 단계를 수행할 수 있다. 결합 워크로드를 목표 워크로드로 간주할 수 있다.
- [0046] 단계 S40에서는 결합 워크로드로부터 최적화 모델을 통해 워크로드의 최적 구성을 출력하는 단계를 수행할 수 있다.
- [0047] 단계 S50에서는 결합 워크로드를 이용하여 학습된 외부 메트릭 예측 모델을 통해 워크로드의 후보 구성을 입력받고 데이터베이스의 성능을 나타내는 적합도 스코어를 기반으로 후보 구성을 평가하는 단계를 수행할 수 있다.
- [0048] 도 4는 기본 워크로드의 구성을 예시하고, 도 5는 목표 워크로드의 구성을 예시한다.
- [0049] 기본 구성은 워크로드 인덱스, 키 사이즈, 값 사이즈, 엔트리 개수, 읽기 쓰기 비율, 업데이트 정보, 또는 이들의 조합을 포함할 수 있다.
- [0050] 기본 워크로드는 구성을 임의 생성한 워크로드이고, 목표 워크로드는 속도를 기준으로 선별한 워크로드이다.
- [0051] 도 6 및 도 7은 워크로드의 내부 메트릭을 나타내며, 도 6은 기본 표현 정보를 예시하고, 도 7은 통계 표현 정보를 예시한다.
- [0052] 내부 메트릭은 기본 구성, 요청 데이터가 존재하지 않음을 나타내는 데이터베이스 캐시 미스(예컨대, rocksdb.block.cache.miss), 또는 이들의 조합을 포함할 수 있다.
- [0053] 통계 표현 정보는 평균, 분산, 제1사분위수, 제2사분위수, 제3사분위수를 포함할 수 있다. 통계 표현 정보를 적용하면 워크로드 표현을 간략하게 하면서 검색 공간의 차원을 줄일 수 있다.
- [0054] 도 8은 본 발명의 일 실시예에 따른 데이터베이스 튜닝 방법이 처리하는 결합 워크로드를 예시한 도면이다.
- [0055] 결합 워크로드는 성능 조절 파라미터, 내부 메트릭, 외부 메트릭이 트리플 매칭될 수 있다.
- [0056] 결합 워크로드를 생성하는 것은 목표 워크로드에 대한 통계 표현 정보 및 복수의 기본 워크로드에 대한 통계 표현 정보 간의 유사 거리를 산출하여 유사 거리 매트릭스를 획득하고, 유사 거리 매트릭스에 따른 최대 유사 거리를 기준으로 스케일을 조절하여 복수의 기본 워크로드를 종합한다. 유사 거리를 산출하는 것은 평균값으로부터의 거리가 표준편차의 몇 배인지를 나타내는 마할라노비스 거리(mahalanobis distance)를 적용할 수 있다.
- [0057] 도 9 및 도 10은 본 발명의 일 실시예에 따른 데이터베이스 튜닝 방법이 처리하는 외부 메트릭 예측 모델을 예시한 도면이다.
- [0058] 외부 메트릭은 총 실행 시간(TIME), 데이터 처리 속도(RATE), 데이터 쓰기 증폭 인자(WAF), 공간 증폭 (SA), 또는 이들의 조합을 포함할 수 있다.
- [0059] TIME은 데이터베이스 튜닝(ex. RocksDB 튜닝)의 시작부터 끝까지의 시간 간격을 나타낸다.

[0060] RATE는 컴팩션, 압축, 읽기 및 쓰기 작업과 같이 RocksDB가 초당 처리하는 작업 수를 나타낸다.

[0061] WAF는 스토리지에 기록된 데이터의 양(물리적 데이터 크기)과 데이터베이스에 기록된 데이터의 양(논리적 데이터 크기)의 비율을 나타내는 성능 지표이다. WAF는 RocksDB를 사용하는 동안 작성되는 추가 데이터의 양을 의미한다.

수학식 1

$$WAF = \frac{\text{Physical data size}}{\text{Logical data size}}$$

[0062]

[0063] SA는 실제 LSM-Tree에 기록된 데이터의 크기로 측정된다. LSM-Tree는 유효한 SST 파일과 유효하지 않은 SST 파일을 동시에 포함하기 때문에 LSM-Tree에 기록된 데이터의 크기는 RocksDB가 이용하는 실제 물리적 공간의 크기에 대한 효과적인 지표로 사용될 수 있다.

[0064] 외부 메트릭 예측 모델은, 성능 조절 파라미터와 기본 구성을 입력받고 복수의 선형 레이어가 연결되는 구조를 거쳐 상기 적합도 스코어를 산출할 수 있다.

[0065] 외부 메트릭 예측 모델이 산출하는 적합도 스코어는 총 실행 시간의 기본값과 예측값, 데이터 처리 속도의 기본값과 예측값, 데이터 쓰기 증폭 인자의 기본값과 예측값, 공간 증폭 인자의 기본값과 예측값에 가중치를 적용한 후 합하여 산출될 수 있다.

수학식 2

$$Score = \alpha_1 \frac{TIME_D}{TIME_P} + \alpha_2 \frac{RATE_P}{RATE_D} + \alpha_3 \frac{WAF_D}{WAF_P} + \alpha_4 \frac{SA_D}{SA_P}$$

[0066]

[0067] 도 11 및 도 12는 본 발명의 일 실시예에 따른 데이터베이스 튜닝 방법이 처리하는 최적화 모델을 예시한 도면이다.

[0068] 워크로드의 최적 구성을 출력하는 것은 최적화 모델에 유전자 알고리즘을 적용할 수 있다.

[0069] 유전자 알고리즘의 유전자는 워크로드의 성능 조절 파라미터에 해당하고, 유전자 알고리즘의 염색체는 워크로드의 기본 구성에 해당하도록 유전자 집단을 설정할 수 있다.

[0070] 외부 메트릭 예측 모델을 이용하여 산출한 적합도 스코어에 따라 유전자 집단의 순위를 결정하고, 유전자 집단의 순위를 고려하여 유전자 집단을 분류할 수 있다.

[0071] 성능 조절 파라미터에 대한 교차 및 변이를 적용하는 유전자 연산을 통해 분류한 유전자 집단을 여러 세대를 거쳐 진화시켜서 최적 구성을 찾을 수 있다.

[0072] 도 13은 본 발명의 다른 실시예에 따른 데이터베이스 튜닝 장치를 예시한 도면이다.

[0073] 데이터베이스 튜닝 장치(110)는 적어도 하나의 프로세서(120), 컴퓨터 판독 가능한 저장매체(130) 및 통신 버스(170)를 포함한다.

[0074] 프로세서(120)는 데이터베이스 튜닝 장치(110)를 동작하도록 제어할 수 있다. 예컨대, 프로세서(120)는 컴퓨터 판독 가능한 저장 매체(130)에 저장된 하나 이상의 프로그램들을 실행할 수 있다. 하나 이상의 프로그램들은 하나 이상의 컴퓨터 실행 가능 명령어를 포함할 수 있으며, 컴퓨터 실행 가능 명령어는 프로세서(120)에 의해 실행되는 경우 데이터베이스 튜닝 장치(110)로 하여금 예시적인 실시예에 따른 동작들을 수행하도록 구성될 수 있다.

[0075] 컴퓨터 판독 가능한 저장 매체(130)는 컴퓨터 실행 가능 명령어 내지 프로그램 코드, 프로그램 데이터 및/또는 다른 적합한 형태의 정보를 저장하도록 구성된다. 컴퓨터 판독 가능한 저장 매체(130)에 저장된 프로그램(140)은 프로세서(120)에 의해 실행 가능한 명령어의 집합을 포함한다. 일 실시예에서, 컴퓨터 판독 가능한 저장 매

체(130)는 메모리(랜덤 액세스 메모리와 같은 휘발성 메모리, 비휘발성 메모리, 또는 이들의 적절한 조합), 하나 이상의 자기 디스크 저장 디바이스들, 광학 디스크 저장 디바이스들, 플래시 메모리 디바이스들, 그 밖에 데이터베이스 튜닝 장치(110)에 의해 액세스되고 원하는 정보를 저장할 수 있는 다른 형태의 저장 매체, 또는 이들의 적합한 조합일 수 있다.

[0076] 통신 버스(170)는 프로세서(120), 컴퓨터 판독 가능한 저장 매체(140)를 포함하여 데이터베이스 튜닝 장치(110)의 다른 다양한 컴포넌트들을 상호 연결한다.

[0077] 데이터베이스 튜닝 장치(110)는 또한 하나 이상의 입출력 장치를 위한 인터페이스를 제공하는 하나 이상의 입출력 인터페이스(150) 및 하나 이상의 통신 인터페이스(160)를 포함할 수 있다. 입출력 인터페이스(150) 및 통신 인터페이스(160)는 통신 버스(170)에 연결된다. 입출력 장치(미도시)는 입출력 인터페이스(150)를 통해 데이터베이스 튜닝 장치(110)의 다른 컴포넌트들에 연결될 수 있다.

[0078] 데이터베이스 튜닝 장치의 프로세서는 워크로드(workload)의 성능 조절 파라미터(knob), 워크로드에 관한 기본 구성을 포함하는 기본 표현 정보를 나타내는 내부 메트릭(internal metric), 워크로드의 성능을 나타내는 외부 메트릭(external metric)이 트리플 매칭된 복수의 기본 워크로드에 대한 데이터 리포지토리(repository)를 생성할 수 있다.

[0079] 데이터베이스 튜닝 장치의 프로세서는 복수의 기본 워크로드 및 샘플 선별된 목표 워크로드에 대해서 내부 메트릭을 이용하여 통계 표현 정보를 생성할 수 있다.

[0080] 데이터베이스 튜닝 장치의 프로세서는 통계 표현 정보를 기반으로 유사 거리를 산출하고 유사 거리에 기반한 스케일을 조절하여 복수의 기본 워크로드를 종합한 결합 워크로드를 생성할 수 있다.

[0081] 데이터베이스 튜닝 장치의 프로세서는 결합 워크로드로부터 최적화 모델을 통해 상기 워크로드의 최적 구성을 출력하는 것을 특징으로 하는 데이터베이스 튜닝 장치를 제공한다.

[0082] 데이터베이스 튜닝 장치의 프로세서는 결합 워크로드를 이용하여 학습된 외부 메트릭 예측 모델을 통해 워크로드의 후보 구성을 입력받고 데이터베이스의 성능을 나타내는 적합도 스코어를 기반으로 후보 구성을 평가할 수 있다.

[0083] 데이터베이스 튜닝 장치의 프로세서는 최적화 모델에 유전자 알고리즘을 적용하고, 유전자 알고리즘의 유전자는 워크로드의 성능 조절 파라미터에 해당하고, 유전자 알고리즘의 염색체는 워크로드의 기본 구성에 해당하도록 유전자 집단을 설정하고, 외부 메트릭 예측 모델을 이용하여 산출한 적합도 스코어에 따라 유전자 집단의 순위를 결정하고, 유전자 집단의 순위를 고려하여 유전자 집단을 분류하고, 성능 조절 파라미터에 대한 교차 및 변이를 적용하는 유전자 연산을 통해 분류한 유전자 집단을 여러 세대를 거쳐 진화시켜서 최적 구성을 찾을 수 있다.

[0084] 도 14 내지 도 19는 실시예들에 따라 수행된 모의실험 결과를 도시한 것이다. 도 14는 전체 성능을 나타내고, 도 15는 워크로드 결합 여부 및 내부 메트릭 축소 여부에 따른 성능을 나타내고, 도 16은 노브의 개수에 따른 성능을 나타내고, 도 17은 적합도 스코어의 가중치에 따른 성능을 나타내고, 도 18은 코사인 유사도와 마할라노비스 거리에 따른 성능을 나타내고, 도 19는 베이지안 최적화에 따른 성능을 나타낸다.

[0085] 기본값(D)과 실제값(A)을 이용하여 총 실행 시간(TIME), 데이터 처리 속도(RATE), 데이터 쓰기 증폭 인자(WAF), 공간 증폭 인자(SA)에 관한 연산을 통해 평가 스코어를 매길 수 있다.

수학식 3

$$Score = \sqrt[4]{\frac{TIME_D}{TIME_A} \times \frac{RATE_A}{RATE_D} \times \frac{WAF_D}{WAF_A} \times \frac{SA_D}{SA_A}}$$

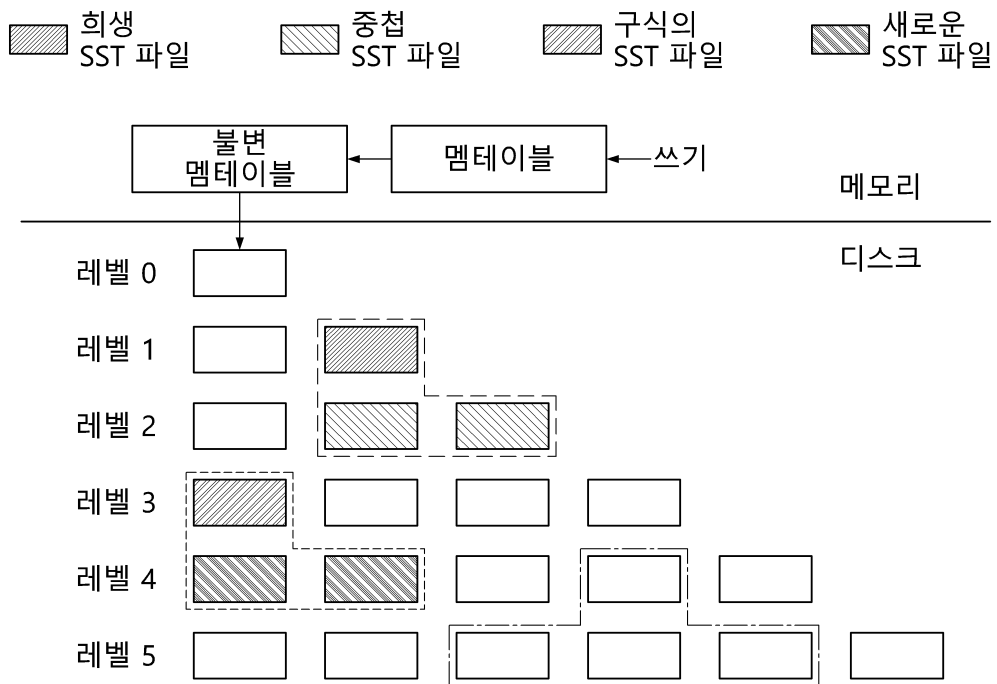
[0086]

[0087] 본 실시예에 의하면 데이터베이스에서 가장 유사한 워크로드를 찾기 위해 워크로드에 대한 새로운 통계 표현으로 정의하고, 유사 거리를 판단하는데 마할라노비스 거리를 적용하고, 결합 워크로드에 유전자 알고리즘을 적용하고 딥러닝을 통해 최적화된 함수를 도출하여 튜닝 시간을 줄이면서 데이터베이스 성능을 향상시킬 수 있다.

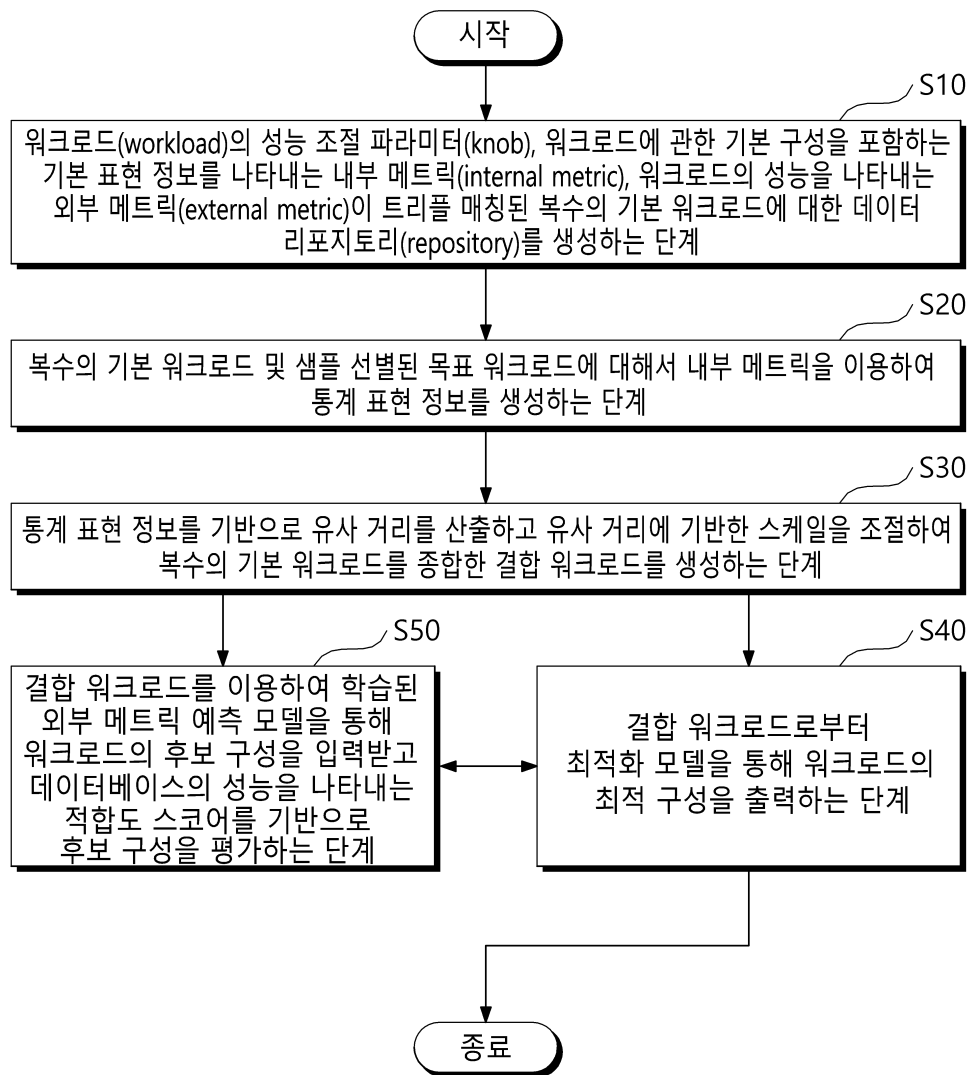
- [0088] 데이터베이스에 포함된 구성요소들이 분리되어 연결될 수 있고, 복수의 구성요소들은 상호 결합되어 적어도 하나의 모듈로 구현될 수 있다. 구성요소들은 장치 내부의 소프트웨어적인 모듈 또는 하드웨어적인 모듈을 연결하는 통신 경로에 연결되어 상호 간에 유기적으로 동작한다. 이러한 구성요소들은 하나 이상의 통신 버스 또는 신호선을 이용하여 통신한다.
- [0089] 데이터베이스는 하드웨어, 펌웨어, 소프트웨어 또는 이들의 조합에 의해 로직회로 내에서 구현될 수 있고, 범용 또는 특정 목적 컴퓨터를 이용하여 구현될 수도 있다. 장치는 고정배선형(Hardwired) 기기, 필드 프로그램 가능한 게이트 어레이(Field Programmable Gate Array, FPGA), 주문형 반도체(Application Specific Integrated Circuit, ASIC) 등을 이용하여 구현될 수 있다. 또한, 장치는 하나 이상의 프로세서 및 컨트롤러를 포함한 시스템온칩(System on Chip, SoC)으로 구현될 수 있다.
- [0090] 데이터베이스는 하드웨어적 요소가 마련된 컴퓨팅 디바이스에 소프트웨어, 하드웨어, 또는 이들의 조합하는 형태로 탑재될 수 있다. 컴퓨팅 디바이스는 각종 기기 또는 유무선 통신망과 통신을 수행하기 위한 통신 모듈 등의 통신장치, 프로그램을 실행하기 위한 데이터를 저장하는 메모리, 프로그램을 실행하여 연산 및 명령하기 위한 마이크로프로세서 등을 전부 또는 일부 포함한 다양한 장치를 의미할 수 있다.
- [0091] 도 2에서는 각각의 과정을 순차적으로 실행하는 것으로 기재하고 있으나 이는 예시적으로 설명한 것에 불과하고, 이 분야의 기술자라면 본 발명의 실시예의 본질적인 특성에서 벗어나지 않는 범위에서 도 5에 기재된 순서를 변경하여 실행하거나 또는 하나 이상의 과정을 병렬적으로 실행하거나 다른 과정을 추가하는 것으로 다양하게 수정 및 변형하여 적용 가능할 것이다.
- [0092] 본 실시예들에 따른 동작은 다양한 컴퓨터 수단을 통하여 수행될 수 있는 프로그램 명령 형태로 구현되어 컴퓨터 판독 가능한 매체에 기록될 수 있다. 컴퓨터 판독 가능한 매체는 실행을 위해 프로세서에 명령어를 제공하는 데 참여한 임의의 매체를 나타낸다. 컴퓨터 판독 가능한 매체는 프로그램 명령, 데이터 파일, 데이터 구조 또는 이들의 조합을 포함할 수 있다. 예를 들면, 자기 매체, 광기록 매체, 메모리 등이 있을 수 있다. 컴퓨터 프로그램은 네트워크로 연결된 컴퓨터 시스템 상에 분산되어 분산 방식으로 컴퓨터가 읽을 수 있는 코드가 저장되고 실행될 수도 있다. 본 실시예를 구현하기 위한 기능적인(Functional) 프로그램, 코드, 및 코드 세그먼트들은 본 실시예가 속하는 기술분야의 프로그래머들에 의해 용이하게 추론될 수 있을 것이다.
- [0093] 본 실시예들은 본 실시예의 기술 사상을 설명하기 위한 것이고, 이러한 실시예에 의하여 본 실시예의 기술 사상의 범위가 한정되는 것은 아니다. 본 실시예의 보호 범위는 아래의 청구범위에 의하여 해석되어야 하며, 그와 동등한 범위 내에 있는 모든 기술 사상은 본 실시예의 권리범위에 포함되는 것으로 해석되어야 할 것이다.

도면

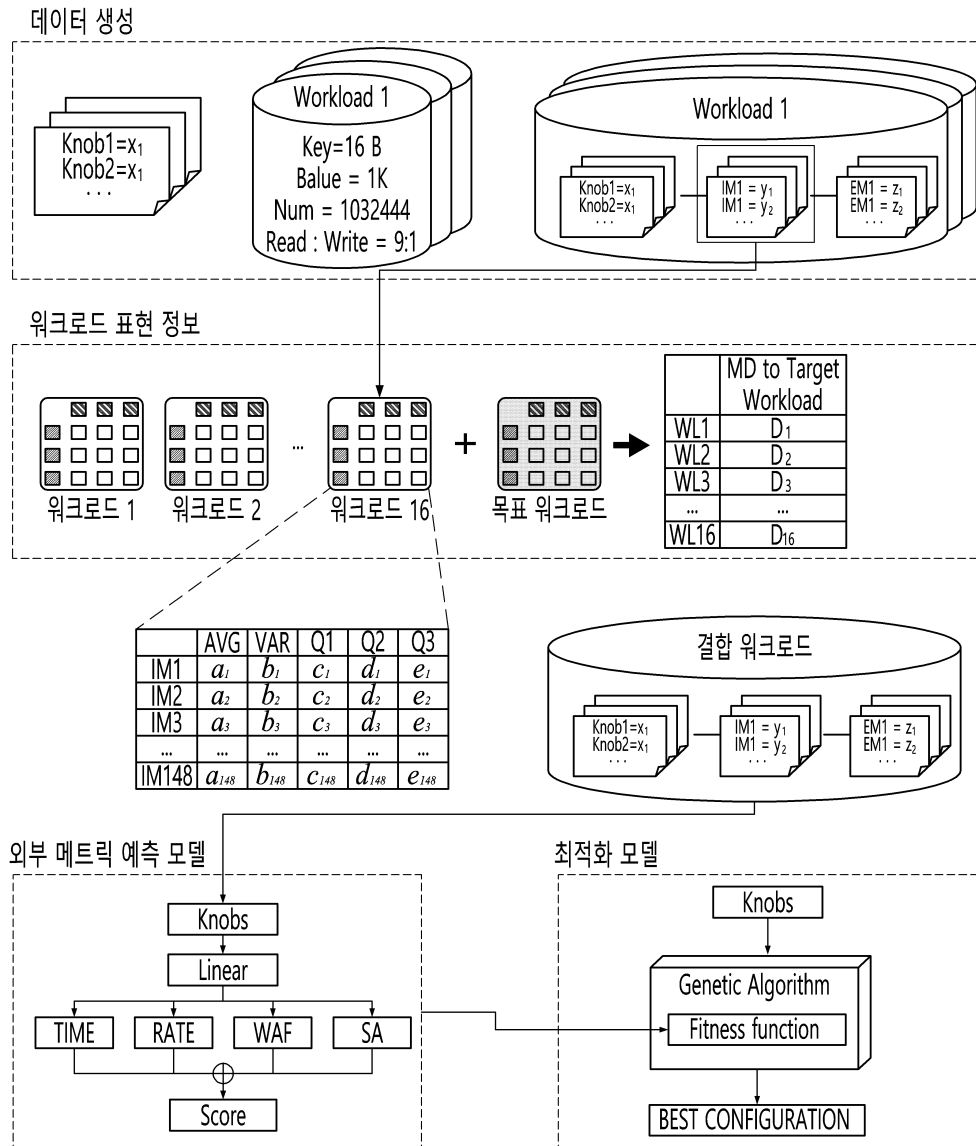
도면1



도면2



도면3



도면4

워크로드 인덱스	값 사이즈 (B), 엔트리 개수	읽기 쓰기 비율	업데이트 정보
0	1024,1032444	9 : 1	-
1	1024,1032444	1 : 1	-
2	1024,1032444	1 : 9	-
3	1024,1032444	-	TRUE
4	4096,261124	9 : 1	-
5	4096,261124	1 : 1	-
6	4096,261124	1 : 9	-
7	4096,261124	-	TRUE
8	16384,65472	9 : 1	-
9	16384,65472	1 : 1	-
10	16384,65472	1 : 9	-
11	16384,65472	-	TRUE
12	65536,16380	9 : 1	-
13	65536,16380	1 : 1	-
14	65536,16380	1 : 9	-
15	65536,16380	-	TRUE

도면5

워크로드 인덱스	값 사이즈 (B), 엔트리 개수	읽기 쓰기 비율	업데이트 정보
16	8192,130816	7 : 3	-
17	8192,130816	3 : 7	-
18	8192,130816	-	TRUE
19	32768,32752	7 : 3	-
20	32768,32752	3 : 7	-
21	32768,32752	-	TRUE

도면6

구성 #	1	2	...	20000
IM 1	12965	13040	...	14586
IM 2	1239	837	...	297
...
IM n	44	63	78	208

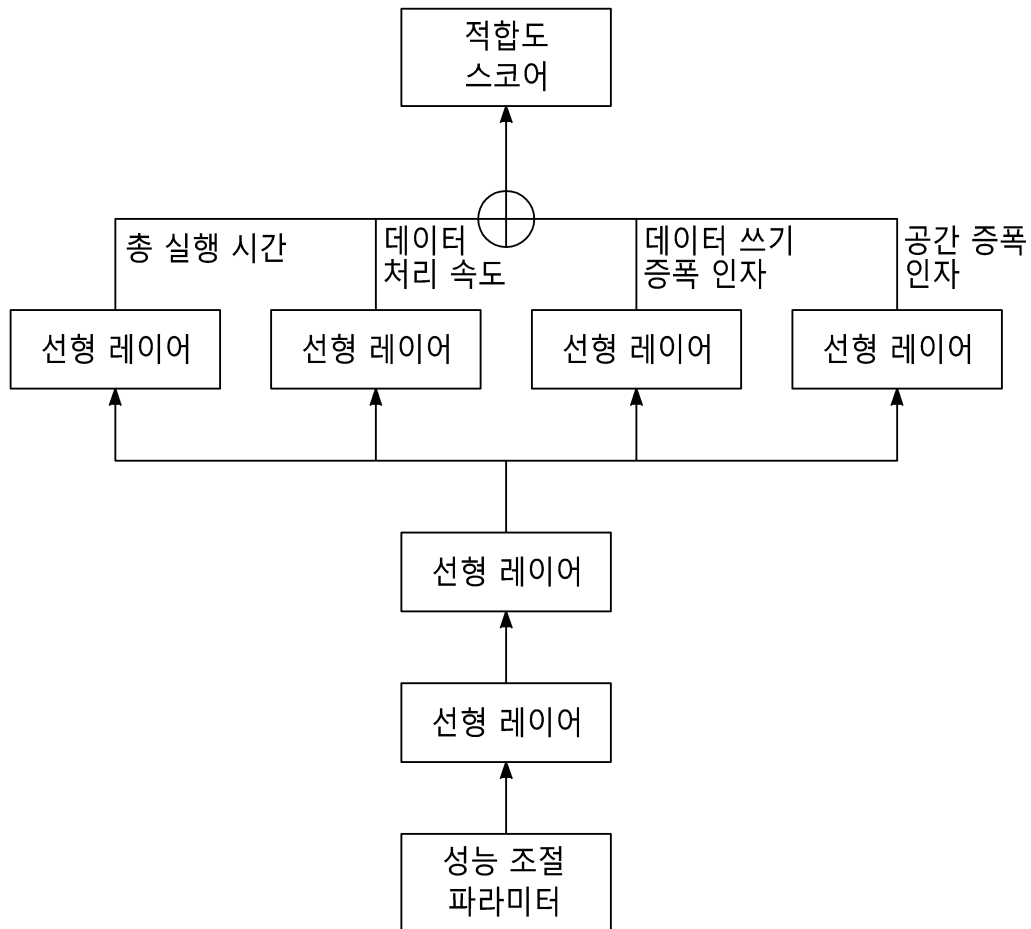
도면7

구성 #	평균	분산	제1사분위수	제2사분위수	제3사분위수
IM 1	13544	55615	10513	13448	15798
IM 2	834	2315	564	912	1132
...
IM n	80	1213	68	81	121

도면8

$$\begin{array}{c} \boxed{} \\ \text{결합 워크로드} \end{array} = \left\{ \begin{array}{l} \begin{array}{c} \boxed{} \\ \text{워크로드 0} \end{array} \times \frac{MD_{max} - MD_0}{\sum_{i=0}^{15} (MD_{max} - MD_i)} \\ \begin{array}{c} \boxed{} \\ \text{워크로드 1} \end{array} \times \frac{MD_{max} - MD_1}{\sum_{i=0}^{15} (MD_{max} - MD_i)} \\ \vdots \\ \begin{array}{c} \boxed{} \\ \text{워크로드 15} \end{array} \times \frac{MD_{max} - MD_{15}}{\sum_{i=0}^{15} (MD_{max} - MD_i)} \end{array} \right.$$

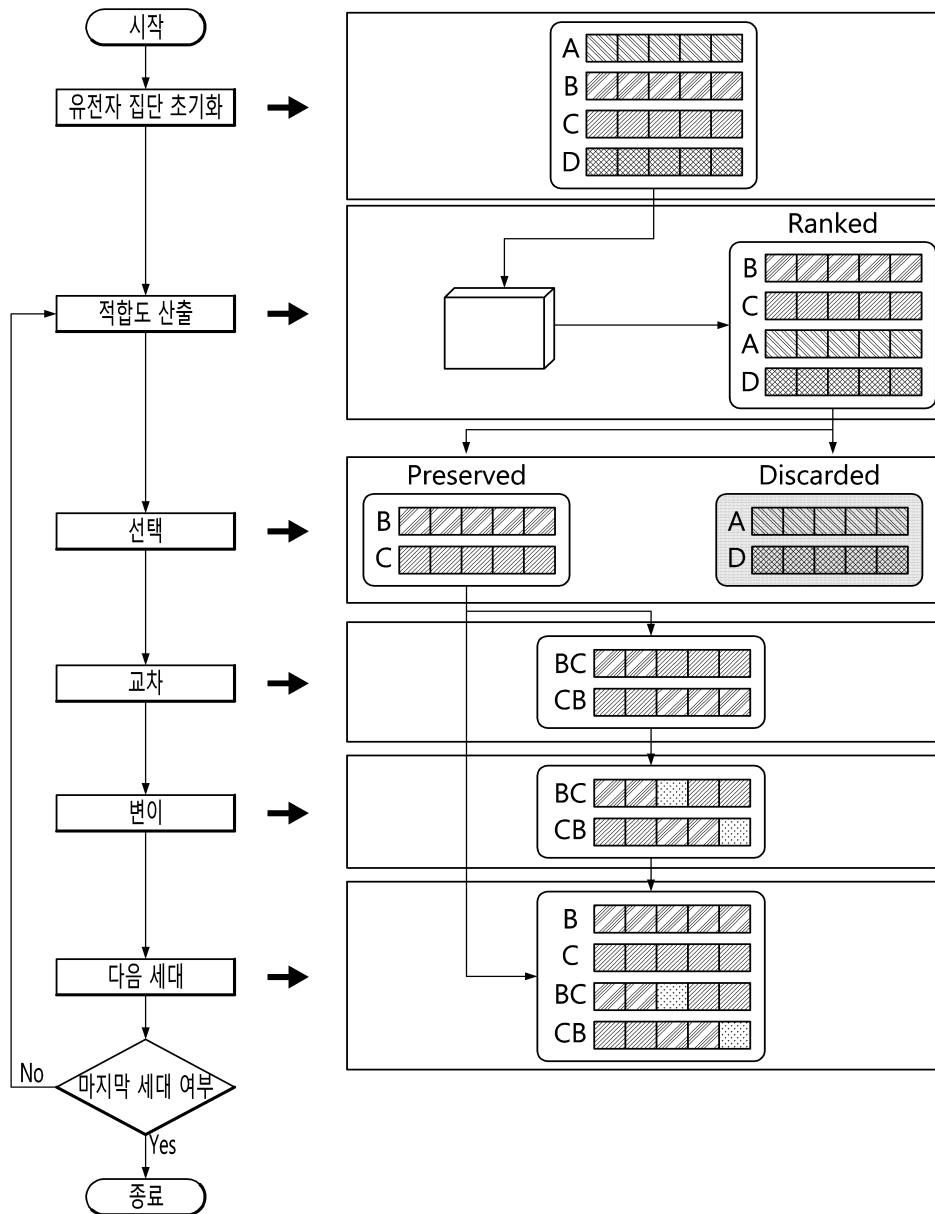
도면9



도면10

최적화	Adamw
학습률	0.0002
훈련 사이클	300
손실 함수	MSE
X 스케일러	MinMaxScaler
Y 스케일러	StandarScaler
레이어	(22, 64, 16, 1)
활성화 함수	ReLU

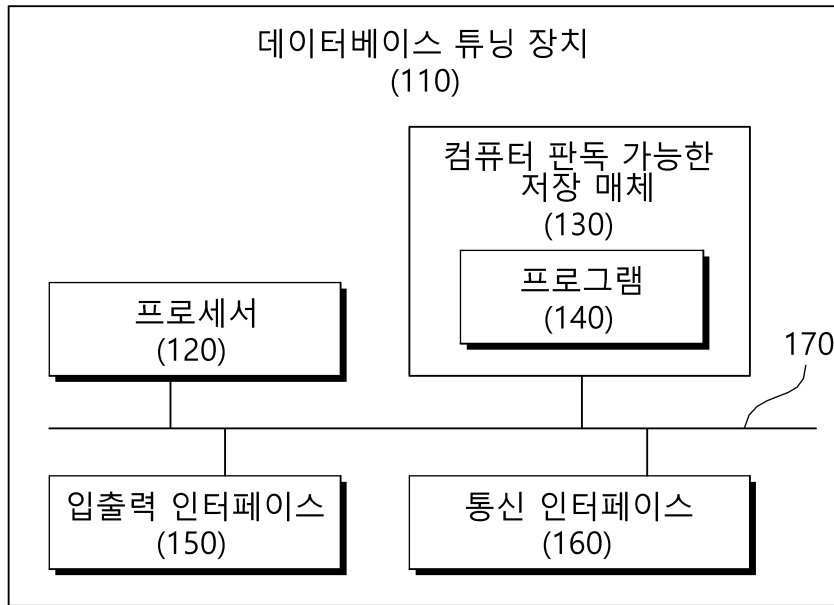
도면11



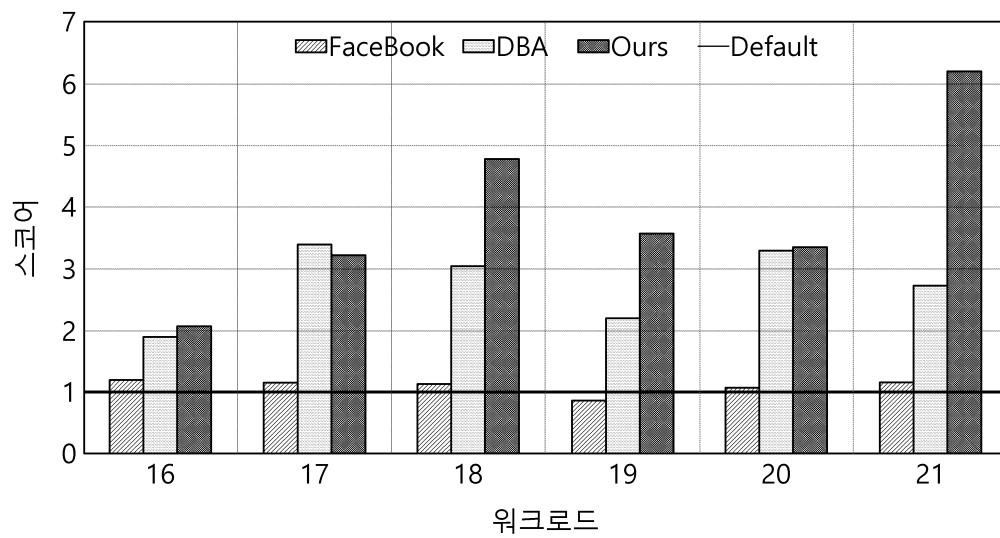
도면12

변이 비율	0.4
교차 비율	0.5
유전자 집단 사이즈	128
세대	1000
선택 알고리즘	Rank Selection
선택 사이즈	64

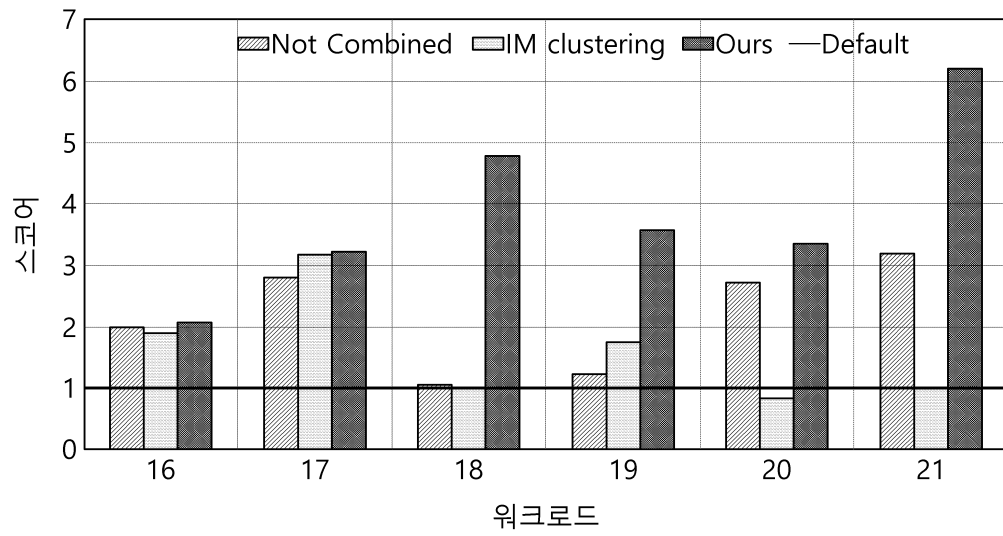
도면13



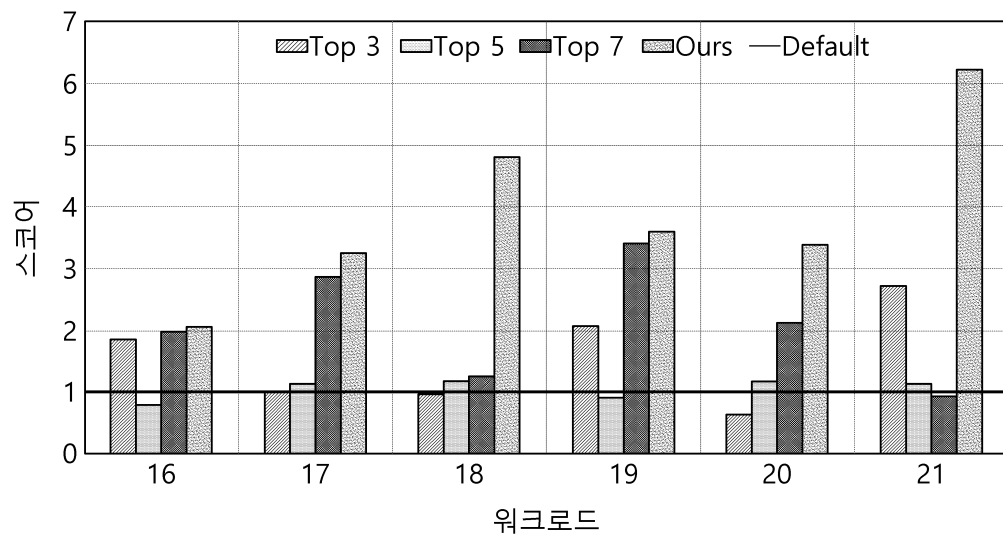
도면14



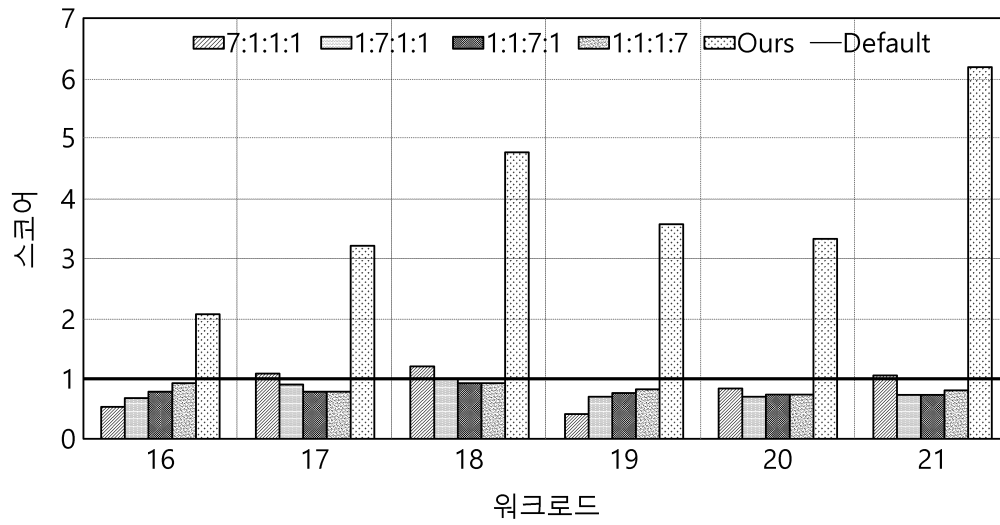
도면15



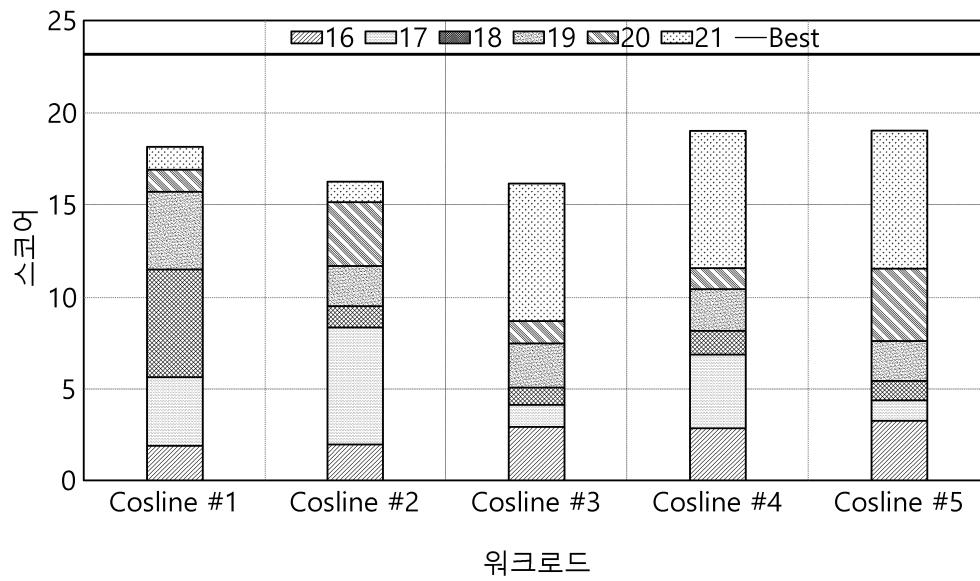
도면16



도면17



도면18



도면19

