



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2024년08월05일
(11) 등록번호 10-2691757
(24) 등록일자 2024년07월31일

(51) 국제특허분류(Int. Cl.)
G16B 50/00 (2019.01) G16B 30/00 (2019.01)
G16B 40/00 (2019.01) G16B 45/00 (2019.01)
(52) CPC특허분류
G16B 50/00 (2019.02)
G16B 30/00 (2019.02)
(21) 출원번호 10-2023-0048538
(22) 출원일자 2023년04월13일
심사청구일자 2023년04월13일
(56) 선행기술조사문헌
N. M. Ghiasi 외, "GenStore: A High-Performance and Energy-Efficient In-Storage Computing System for Genome Sequence Analysis", arXiv, arXiv:2202.10400v2 (2023.04.06.)*
(뒷면에 계속)

(73) 특허권자
고려대학교 산학협력단
서울특별시 성북구 안암로 145, 고려대학교 (안암동5가)
(72) 발명자
구건재
경기도 남양주시 불암로 337, 5103동 2004호
박보영
경기도 양평군 양서면 부용길 44-2
(74) 대리인
김준석, 박민욱

전체 청구항 수 : 총 8 항

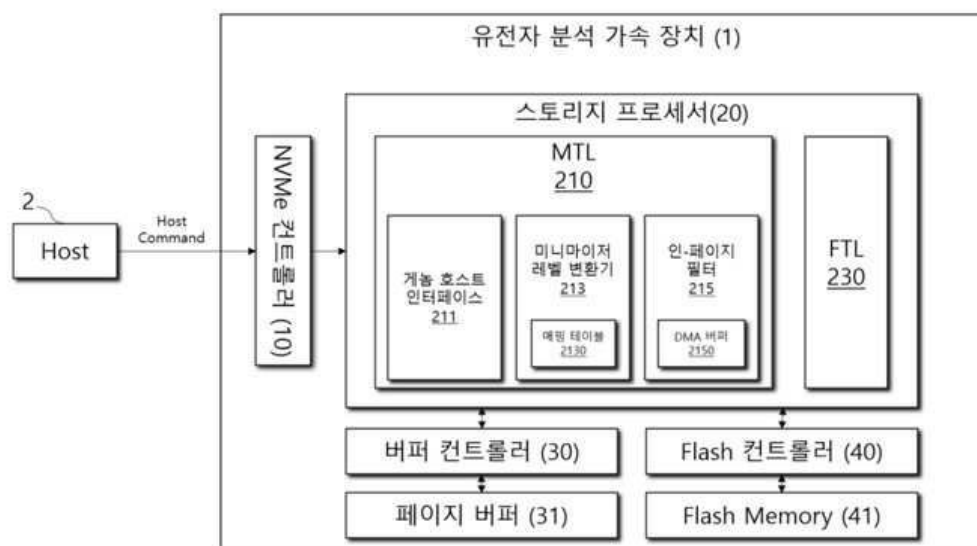
심사관 : 김채연

(54) 발명의 명칭 SSD 인-스토리지 프로세싱을 이용한 유전자 분석 가속 장치 및 방법

(57) 요약

본 발명은 SSD 인-스토리지 프로세싱 유전자 분석 가속 장치에 관한 것으로, 유전자 분석 가속 장치는 참조 유전체를 일정한 길이로 나눈 각 유전체의 대표 문자열인 복수의 미니마이어 및 복수의 미니마이어의 위치 배열을 저장한 페이지를 포함하는 플래시 메모리, 페이지의 페이지 번호와 해당 페이지에 저장된 복수의 미니마이어 중 하
(뒷면에 계속)

대표도 - 도1



나인 대표 미니마이저를 매핑시킨 매핑 테이블, 명령어를 받으면 명령어에 포함된 쿼리 미니마이저를 추출하는 계층 호스트 인터페이스, 매핑 테이블을 탐색하여 추출한 쿼리 미니마이저에 매핑되어있는 페이지 번호를 산출하여 플래시 메모리에 페이지 읽기 요청을 보내는 미니마이저 레벨 변환기, 페이지 읽기 요청에 따라 읽은 페이지를 복사하여 저장하는 페이지 버퍼 및 DMA 버퍼를 포함하여, 페이지 버퍼에 저장된 페이지 내부에서 쿼리 미니마이저에 해당하는 위치 배열을 복사하여 DMA 버퍼에 저장하는 인-페이지 필터를 포함하고, 매핑 테이블은, 모든 페이지 번호에 대해 대표 미니마이저가 매핑되어 있는 하위 테이블 및 페이지 번호를 복수개 묶어 복수개의 블록으로 나누고, 각 블록에 해시 테이블 ID를 매핑한 상위 테이블을 포함하는 2-레벨 테이블인 것을 특징으로 할 수 있다.

(52) CPC특허분류

G16B 40/00 (2019.02)

G16B 45/00 (2019.02)

(56) 선행기술조사문헌

KR1020220146679 A

KR1020150138528 A

KR1020220164409 A

KR1020140081385 A

KR1020230009878 A

*는 심사관에 의하여 인용된 문헌

이 발명을 지원한 국가연구개발사업	
과제고유번호	1711186523
과제번호	2021R1C1C1012172
부처명	과학기술정보통신부
과제관리(전문)기관명	한국연구재단
연구사업명	개인기초연구(과기정통부)
연구과제명	대규모 그래프 구조의 신경망 처리를 위한 효율적인 프로세서 및 메모리 구조에 대
한 연구	
기 여 율	70/100
과제수행기관명	고려대학교
연구기간	2023.03.01 ~ 2024.02.29
이 발명을 지원한 국가연구개발사업	
과제고유번호	1711193663
과제번호	2020-0-01819-004
부처명	과학기술정보통신부
과제관리(전문)기관명	정보통신기획평가원
연구사업명	정보통신방송혁신인재양성
연구과제명	ICT명품인재양성(고려대학교)
기 여 율	10/100
과제수행기관명	고려대학교산학협력단
연구기간	2023.01.01 ~ 2023.12.31
이 발명을 지원한 국가연구개발사업	
과제고유번호	1711193387
과제번호	2022-0-01198-002
부처명	과학기술정보통신부
과제관리(전문)기관명	정보통신기획평가원
연구사업명	정보통신방송혁신인재양성
연구과제명	융합보안대학원(고려대학교)
기 여 율	10/100
과제수행기관명	고려대학교산학협력단
연구기간	2023.01.01 ~ 2023.12.31
이 발명을 지원한 국가연구개발사업	
과제고유번호	1345375069
과제번호	2021R1A6A1A13044830
부처명	교육부
과제관리(전문)기관명	한국연구재단
연구사업명	이공학학술연구기반구축
연구과제명	융합소프트웨어연구소
기 여 율	10/100
과제수행기관명	고려대학교
연구기간	2023.03.01 ~ 2024.02.29

명세서

청구범위

청구항 1

SSD 인-스토리지 프로세싱 유전자 분석 가속 장치에 관한 것으로,

참조 유전체를 일정한 길이로 나눈 각 유전체의 대표 문자열인 복수의 미니마이저, 상기 복수의 미니마이저의 위치 배열 및 상기 미니마이저의 위치 배열의 원소개수를 의미하는 위치 배열 크기를 저장한 페이지를 포함하는 플래시 메모리;

상기 페이지에 저장된 상기 복수의 미니마이저 중 하나인 대표 미니마이저를 해당 페이지의 페이지 번호에 매핑시킨 매핑 테이블;

명령어를 받으면 상기 명령어에 포함된 쿼리 미니마이저를 추출하는 게놈 호스트 인터페이스;

상기 매핑 테이블을 탐색하여 추출한 상기 쿼리 미니마이저에 매핑되어있는 상기 페이지 번호를 산출하여 상기 플래시 메모리에 페이지 읽기 요청을 보내는 미니마이저 레벨 변환기;

상기 페이지 읽기 요청에 따라 읽은 상기 페이지를 복사하여 저장하는 페이지 버퍼; 및

DMA 버퍼를 포함하여, 상기 페이지 버퍼에 저장된 페이지 내부에서 상기 쿼리 미니마이저에 해당하는 위치 배열을 복사하여 상기 DMA 버퍼에 저장하는 인-페이지 필터; 를 포함하고,

상기 대표 미니마이저는,

상기 페이지에 저장된 상기 미니마이저 중 가장 작은 위치 배열 크기를 가진 미니마이저로 정하고,

상기 매핑 테이블은,

모든 상기 페이지 번호에 대해 상기 대표 미니마이저가 매핑되어 있는 하위 테이블; 및

상기 페이지 번호를 복수개 묶어 복수개의 블록으로 나누고, 각 상기 블록에 해시 테이블 ID를 매핑한 상위 테이블; 을 포함하는 2-레벨 테이블인 것을 특징으로 하는

유전자 분석 가속 장치.

청구항 2

제1항에 있어서,

상기 게놈 호스트 인터페이스는,

상기 명령어에 포함된 상기 해시 테이블 ID를 더 추출하고,

상기 미니마이저 레벨 변환기는,

상기 해시 테이블 ID를 통해 상기 상위 테이블에서 상기 블록을 먼저 산출하고, 상기 하위 테이블에서 상기 미니마이저에 대해 매핑되어있는 페이지 번호를 산출하여 플래시 메모리에 페이지 읽기 요청을 보내는 것을 특징으로 하는

유전자 분석 가속 장치.

청구항 3

삭제

청구항 4

제1항에 있어서,

상기 인-페이지 필터는,

상기 페이지 버퍼에 저장된 페이지 내부에서 상기 쿼리 미니마이저에 해당하는 위치 배열 크기를 더 복사하여 상기 DMA 버퍼에 더 저장하는 것을 특징으로 하는

유전자 분석 가속 장치.

청구항 5

제1항에 있어서,

상기 계층 호스트 인터페이스는,

상기 명령어가 내부 연산 명령인지 판단하는 것을 특징으로 하는

유전자 분석 가속 장치.

청구항 6

SSD 인-스토리지 프로세싱 유전자 분석 가속 방법에 관한 것으로,

플래시 메모리의 페이지에 참조 유전체를 일정한 길이로 나눈 각 유전체의 대표 문자열인 복수의 미니마이저, 상기 복수의 미니마이저의 위치 배열 및 상기 미니마이저의 위치 배열의 원소개수를 의미하는 위치 배열 크기를 저장하는 단계;

상기 페이지의 페이지 번호와 해당 페이지에 저장된 상기 복수의 미니마이저 중 하나인 대표 미니마이저를 매핑 시킨 매핑 테이블을 작성하는 단계;

계층 호스트 인터페이스에서 명령어를 받아 상기 명령어에 포함된 쿼리 미니마이저를 추출하는 단계;

미니마이저 레벨 변환기에서 상기 매핑 테이블을 탐색하여 추출한 상기 쿼리 미니마이저에 매핑되어있는 상기 페이지 번호를 산출하여 상기 플래시 메모리에 페이지 읽기 요청을 보내는 단계;

상기 페이지 읽기 요청에 따라 읽은 상기 페이지를 복사하여 페이지 버퍼에 저장하는 단계; 및

DMA 버퍼를 포함하는 인-페이지 필터에, 상기 페이지 버퍼에 저장된 페이지 내부에서 상기 쿼리 미니마이저에 해당하는 위치 배열을 복사하여 상기 DMA 버퍼에 저장하는 단계; 를 포함하고,

상기 대표 미니마이저는,

상기 페이지에 저장된 상기 미니마이저 중 가장 작은 위치 배열 크기를 가진 미니마이저로 정하고,

상기 매핑 테이블은,

모든 상기 페이지 번호에 대해 상기 대표 미니마이저가 매핑되어 있는 하위 테이블; 및

상기 페이지 번호를 복수개 묶어 복수개의 블록으로 나누고, 각 상기 블록에 해시 테이블 ID를 매핑한 상위 테이블; 을 포함하는 2-레벨 테이블인 것을 특징으로 하는

유전자 분석 가속 방법.

청구항 7

제6항에 있어서,

상기 쿼리 미니마이저를 추출하는 단계는,

상기 명령어에 포함된 상기 해시 테이블 ID를 더 추출하고,

상기 페이지 읽기 요청을 보내는 단계는,

상기 해시 테이블 ID를 통해 상기 상위 테이블에서 상기 블록을 먼저 산출하고, 상기 하위 테이블에서 상기 미니마이저에 대해 매핑되어있는 페이지 번호를 산출하여 플래시 메모리에 페이지 읽기 요청을 보내는 것을 특징으로 하는

유전자 분석 가속 방법.

청구항 8

삭제

청구항 9

제6항에 있어서,

상기 DMA 버퍼에 저장하는 단계는,

상기 페이지 버퍼에 저장된 페이지 내부에서 상기 쿼리 미니마이저에 해당하는 위치 배열 크기를 더 복사하여 상기 DMA 버퍼에 더 저장하는 것을 특징으로 하는

유전자 분석 가속 방법.

청구항 10

제6항에 있어서,

상기 계층 호스트 인터페이스는,

상기 명령어가 내부 연산 명령인지 판단하는 것을 특징으로 하는

유전자 분석 가속 방법.

발명의 설명

기술 분야

- [0001] 본 발명은 유전자 분석 가속 장치 및 방법에 관한 것으로, 유전자 분석 과정 중 리드 매핑 과정을 SSD의 내부 연산 능력을 이용하여 가속시키는 장치 및 방법에 관한 기술이다.

배경 기술

- [0002] 유전자 분석 과정의 리드 매핑(read mapping)은 유전자 시퀀싱(sequencing)에 의해 생성된 리드(read)라는 유전자 정보 문자열을 종별 참조하여 유전체(genome)에 대해 정렬하여 어느 염색체에 속하는지 결정하는 단계이다.
- [0003] 일반적으로 사용되는 대부분의 전장 염기 서열 규정 어플리케이션은 seed-chain-align 방법에 따라 진행된다. 이 방법은 인덱싱(indexing), 시딩(seeding), 체이닝(chaining) 및 얼라이닝(aligning) 과정을 포함한다.
- [0004] 인덱싱 과정에서는 참조 유전체를 일정한 길이로 나누어 각각의 대표 문자열인 미니마이저(minimizer)와 그 위치 배열을 해시 테이블(hash table) 형태로 저장한다. 시딩 과정에서는 매핑을 수행하고자 하는 리드인 쿼리 리드(query read)의 미니마이저를 생성하여 참조 인덱스에서 그 위치를 찾고 시드(seed)를 정한다. 체이닝 과정에서는 가까이 위치한 시드들을 다이내믹 프로그래밍에 의해 엮어서 체인을 생성해 글로벌 매핑을 수행할 수 있다. 얼라이닝 과정은 엮기 단위의 정렬을 진행하여 로컬 매핑을 수행할 수 있다.
- [0005] 참조 인덱스는 수 기가바이트에 이르는 매우 많은 정보를 담고 있기 때문에, 시딩 과정 중 미니마이저를 참조 인덱스 내에서 찾는 과정은 매우 많은 데이터 검색 및 이동을 필요로 한다. 이러한 데이터 이동은 유전자 분석 과정의 오버헤드를 증가시켜 성능 및 효율이 저하되는 문제가 있었다.

발명의 내용

해결하려는 과제

- [0006] 본 발명의 일 실시예에 따른 유전자 분석 가속 장치 및 방법의 목적은, SSD의 내부 연산 능력을 이용하여 유전자 분석에 필요한 데이터의 이동을 줄여 리드 매핑 과정을 가속하고자 하는데 있다.

과제의 해결 수단

- [0007] 본 발명의 일 실시예는 SSD 인-스토리지 프로세싱 유전자 분석 가속 장치에 관한 것으로, 참조 유전체를 일정한

길이로 나눈 각 유전체의 대표 문자열인 복수의 미니마이저 및 복수의 미니마이저의 위치 배열을 저장한 페이지를 포함하는 플래시 메모리, 페이지의 페이지 번호와 해당 페이지에 저장된 복수의 미니마이저 중 하나인 대표 미니마이저를 매핑시킨 매핑 테이블, 명령어를 받으면 명령어에 포함된 쿼리 미니마이저를 추출하는 게놈 호스트 인터페이스, 매핑 테이블을 탐색하여 추출한 쿼리 미니마이저에 매핑되어있는 페이지 번호를 산출하여 플래시 메모리에 페이지 읽기 요청을 보내는 미니마이저 레벨 변환기, 페이지 읽기 요청에 따라 읽은 페이지를 복사하여 저장하는 페이지 버퍼 및 DMA 버퍼를 포함하여, 페이지 버퍼에 저장된 페이지 내부에서 쿼리 미니마이저에 해당하는 위치 배열을 복사하여 DMA 버퍼에 저장하는 인-페이지 필터를 포함하고, 매핑 테이블은, 모든 페이지 번호에 대해 대표 미니마이저가 매핑되어 있는 하위 테이블 및 페이지 번호를 복수개 묶어 복수개의 블록으로 나누고, 각 블록에 해시 테이블 ID를 매핑한 상위 테이블을 포함하는 2-레벨 테이블인 것을 특징으로 할 수 있다.

[0008] 또한, 일 실시예에 따라서, 게놈 호스트 인터페이스는, 명령어에 포함된 해시 테이블 ID를 더 추출하고, 미니마이저 레벨 변환기는, 해시 테이블 ID를 통해 상위 테이블에서 블록을 먼저 산출하고, 하위 테이블에서 미니마이저에 대해 매핑되어있는 페이지 번호를 산출하여 플래시 메모리에 페이지 읽기 요청을 보내는 것을 특징으로 할 수 있다.

[0009] 또한, 일 실시예에 따라서, 대표 미니마이저는, 페이지에 저장된 미니마이저 중 가장 작은 위치 배열 크기를 가진 미니마이저로 정하는 것을 특징으로 할 수 있다.

[0010] 또한, 일 실시예에 따라서, 플래시 메모리는, 복수의 미니마이저의 위치 배열 크기를 더 저장하고, 인-페이지 필터는, 페이지 버퍼에 저장된 페이지 내부에서 쿼리 미니마이저에 해당하는 위치 배열 크기를 더 복사하여 DMA 버퍼에 더 저장하는 것을 특징으로 할 수 있다.

[0011] 또한, 일 실시예에 따라서, 게놈 호스트 인터페이스는, 명령어가 내부 연산 명령인지 판단하는 것을 특징으로 할 수 있다.

[0012] 본 발명의 일 실시예는 SSD 인-스토리지 프로세싱 유전자 분석 가속 방법에 관한 것으로, 플래시 메모리의 페이지에 참조 유전체를 일정한 길이로 나눈 각 유전체의 대표 문자열인 복수의 미니마이저 및 복수의 미니마이저의 위치 배열을 저장하는 단계, 페이지의 페이지 번호와 해당 페이지에 저장된 복수의 미니마이저 중 하나인 대표 미니마이저를 매핑시킨 매핑 테이블을 작성하는 단계, 게놈 호스트 인터페이스에서 명령어를 받아 명령어에 포함된 쿼리 미니마이저를 추출하는 단계, 미니마이저 레벨 변환기에서 매핑 테이블을 탐색하여 추출한 쿼리 미니마이저에 매핑되어있는 페이지 번호를 산출하여 플래시 메모리에 페이지 읽기 요청을 보내는 단계, 페이지 읽기 요청에 따라 읽은 페이지를 복사하여 페이지 버퍼에 저장하는 단계 및 DMA 버퍼를 포함하는 인-페이지 필터에, 페이지 버퍼에 저장된 페이지 내부에서 쿼리 미니마이저에 해당하는 위치 배열을 복사하여 DMA 버퍼에 저장하는 단계를 포함하고, 매핑 테이블은, 모든 페이지 번호에 대해 대표 미니마이저가 매핑되어 있는 하위 테이블 및 페이지 번호를 복수개 묶어 복수개의 블록으로 나누고, 각 블록에 해시 테이블 ID를 매핑한 상위 테이블; 을 포함하는 2-레벨 테이블인 것을 특징으로 할 수 있다.

[0013] 또한, 일 실시예에 따라서, 쿼리 미니마이저를 추출하는 단계는, 명령어에 포함된 해시 테이블 ID를 더 추출하고, 페이지 읽기 요청을 보내는 단계는, 해시 테이블 ID를 통해 상위 테이블에서 블록을 먼저 산출하고, 하위 테이블에서 미니마이저에 대해 매핑되어있는 페이지 번호를 산출하여 플래시 메모리에 페이지 읽기 요청을 보내는 것을 특징으로 할 수 있다.

[0014] 또한, 일 실시예에 따라서, 대표 미니마이저는, 페이지에 저장된 미니마이저 중 가장 작은 위치 배열 크기를 가진 미니마이저로 정하는 것을 특징으로 할 수 있다.

[0015] 또한, 일 실시예에 따라서, 플래시 메모리는, 복수의 미니마이저의 위치 배열 크기를 더 저장하고, DMA 버퍼에 저장하는 단계는, 페이지 버퍼에 저장된 페이지 내부에서 쿼리 미니마이저에 해당하는 위치 배열 크기를 더 복사하여 DMA 버퍼에 더 저장하는 것을 특징으로 할 수 있다. 또한, 일 실시예에 따라서, 게놈 호스트 인터페이스는, 명령어가 내부 연산 명령인지 판단하는 것을 특징으로 할 수 있다.

발명의 효과

[0016] 본 발명의 일 실시예에 의한 유전자 분석 가속 장치 및 방법에 따르면, SSD 내부의 연산 능력을 이용하여 리드 매핑 과정 상의 데이터 이동량을 감소시킴으로써 유전자 분석의 효율 및 성능을 향상시킬 수 있는 효과가 있다.

도면의 간단한 설명

- [0017] 도 1은 일 실시예에 따른 유전자 분석 가속 장치의 블록도이다.
- 도 2는 일 실시예에 따라 플래시 메모리의 각 페이지에 참조 인덱스 데이터가 저장된 모습을 도시한 도면이다.
- 도 3은 일 실시예에 따른 매핑 테이블을 도시한 도면이다.
- 도 4는 일 실시예에 따른 매핑 테이블을 도시한 도면이다.
- 도 5는 일 실시예에 따른 스토리지 프로세서 및 페이지 버퍼를 도시한 블록도이다.
- 도 6a는 일 실시예에 따라 메모리 사용량을 측정한 결과 그래프이다.
- 도 6b는 일 실시예에 따라 데이터 이동 시간을 측정한 결과 그래프이다.
- 도 7은 일 실시예에 따른 유전자 분석 가속 방법의 순서도이다.

발명을 실시하기 위한 구체적인 내용

- [0018] 본 발명의 이점 및 특징, 그리고 그것들을 달성하는 방법은 첨부되는 도면과 함께 후술되어 있는 실시예들을 참조하면 명확해질 것이다. 그러나 본 발명은 이하에서 개시되는 실시예들에 한정되는 것이 아니라 서로 다른 다양한 형태로 구현될 수 있으며, 단지 본 실시예들은 본 발명의 개시가 완전하도록 하고, 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자에게 발명의 범주를 완전하게 알려주기 위해 제공되는 것이며, 본 발명은 청구항의 범주에 의해 정의될 뿐이다.
- [0019] 본 명세서에서 사용되는 용어에 대해 간략히 설명하고, 본 발명에 대해 구체적으로 설명하기로 한다.
- [0020] 본 발명에서 사용되는 용어는 본 발명에서의 기능을 고려하면서 가능한 현재 널리 사용되는 일반적인 용어들을 선택하였으나, 이는 당 분야에 종사하는 기술자의 의도 또는 관례, 새로운 기술의 출현 등에 따라 달라질 수 있다. 또한, 특정한 경우는 출원인이 임의로 선정한 용어도 있으며, 이 경우 해당되는 발명의 설명 부분에서 상세히 그 의미를 기재할 것이다. 따라서 본 발명에서 사용되는 용어는 단순한 용어의 명칭이 아닌, 그 용어가 가지는 의미와 본 발명의 전반에 걸친 내용을 토대로 정의되어야 한다.
- [0021] 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 있음을 의미한다. 또한, 명세서에서 사용되는 "부", "모듈", "유닛" 등의 용어는 적어도 하나의 기능 또는 동작을 처리하는 단위를 의미하며, 소프트웨어, FPGA 또는 ASIC과 같은 하드웨어 구성요소, 또는 소프트웨어와 하드웨어의 결합으로 구현될 수 있다. 그렇지만 "부", "모듈", "유닛" 등의 용어가 소프트웨어 또는 하드웨어에 한정되는 의미는 아니다. "부", "모듈", "유닛" 등은 어드레싱할 수 있는 저장 매체에 있도록 구성될 수도 있고 하나 또는 그 이상의 프로세서들을 재생시키도록 구성될 수도 있다. 따라서, 일 예로서 "부", "모듈", "유닛" 등의 용어는 소프트웨어 구성요소들, 객체지향 소프트웨어 구성요소들, 클래스 구성요소들 및 태스크 구성요소들과 같은 구성요소들과, 프로세스들, 함수들, 속성들, 프로시저들, 서브루틴들, 프로그램 코드의 세그먼트들, 드라이버들, 펌웨어, 마이크로 코드, 회로, 데이터, 데이터베이스, 데이터 구조들, 테이블들, 어레이들 및 변수들을 포함한다.
- [0022] 아래에서는 첨부한 도면을 참고하여 본 발명의 실시예에 대하여 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자가 용이하게 실시할 수 있도록 상세히 설명한다. 그리고 도면에서 본 발명을 명확하게 설명하기 위해서 설명과 관계없는 부분은 생략한다. "제1", "제2" 등과 같이 서수를 포함하는 용어는 다양한 구성 요소들을 설명하는데 사용될 수 있지만, 구성 요소들은 용어들에 의해 한정되지는 않는다. 용어들은 하나의 구성요소를 다른 구성요소로부터 구별하는 목적으로만 사용된다. 예를 들어, 본 발명의 권리 범위를 벗어나지 않으면서 제1 구성요소는 제2 구성요소로 명명될 수 있고, 유사하게 제2 구성요소도 제1 구성요소로 명명될 수 있다. "및/또는"이라는 용어는 복수의 관련된 항목들의 조합 또는 복수의 관련된 항목들 중의 어느 하나의 항목을 포함한다.
- [0023] 본 명세서에 설명된 다양한 시스템, 장치, 저장 매체, 모듈 및 기구는 전용 연산 디바이스, 또는 하나 이상의 전용 연산 디바이스의 하나 이상의 연산 칩에서 구현될 수도 있다.
- [0024] 유전자 분석 과정 중 리드 매핑(read mapping)은 유전자 시퀀싱(sequencing)에 의해 생성된 리드(read)라는 유전자 정보 문자열을 중별 참조 유전체에 대해 정렬하여 어느 염색체에 속하는 지 결정하는 단계이다. 유전자 분

석 과정 중의 리드 매핑은 인덱싱(Indexing), 시딩(Seeding), 체이닝(Chaining) 및 얼라이닝(Aligning) 과정으로 구성되어 있다.

- [0025] 인덱싱 과정에서는 참조 유전체를 일정한 길이로 나누고, 일정한 길이로 나누어진 유전체 각각의 대표 문자열인 미니마이저(minimizer)를 추출할 수 있다. 또한, 추출한 미니마이저와 해당 미니마이저 위치 배열을 해시 테이블(hash table) 형태로 저장할 수 있다. 이러한 미니마이저 및 미니마이저 위치 배열 정보를 참조 인덱스로 정의할 수 있다.
- [0026] 시딩 과정에서는 매핑을 수행하고자 하는 리드인 쿼리 리드(Query Read)의 미니마이저를 생성하고, 참조 인덱스에서 해당 미니마이저 및 미니마이저의 위치를 찾아 시드(Seed)로 정할 수 있다.
- [0027] 체이닝 과정에서는 가까이 위치한 시드들을 다이نام릭 프로그래밍(dynamic programming)에 의해 엮어서 체인을 생성하는 글로벌 매핑을 수행할 수 있다.
- [0028] 얼라이닝 과정에서는 염기 단위의 정렬을 진행하여 로컬 매핑을 수행할 수 있다.
- [0029] 기존의 유전자 분석 과정은 시딩 과정에서 저장장치 내부 연산 능력을 활용하지 않고 호스트의 메인 메모리를 주로 사용하여 호스트와 디바이스간 방대한 데이터 이동이 수행되기 때문에 이로 인한 오버헤드 증가 및 성능 저하가 발생하였다.
- [0030] 본 발명의 유전자 분석 가속 장치(1) 및 방법은 유전자 분석 과정 중 시딩 과정을 SSD(Solid-State Drive) 내부에서 소프트웨어적으로 처리하여 호스트와 디바이스간 데이터 이동량을 감소시킬 수 있어 유전자 분석 성능을 향상시킬 수 있다.
- [0031] 또한, 유전자 분석을 SSD 내부의 펌웨어만으로 수행하도록 함으로써, 새로운 하드웨어를 추가하지 않고 위 성능 향상을 달성할 수 있다.
- [0032] 이하, 도면을 참조하여 본 발명의 유전자 분석 가속 장치(1) 및 방법에 관하여 설명하도록 한다.
- [0033] 도 1은 본 발명의 일 실시예에 따른 유전자 분석 가속 장치(1)의 블록도를 도시한 도면이다.
- [0034] 본 발명의 유전자 분석 가속 장치(1)는 인-스토리지 프로세싱 유전자 분석 가속 장치(1)로, 저장 장치 상에서 구현될 수 있다.
- [0035] 본 발명의 유전자 분석 가속 장치(1)는 SSD 상에서 구현될 수 있으며, 바람직하게는 PCIe(Peripheral Component Interconnect Express) 및 NVMe(NON-Volatile Memory Express) 인터페이스를 사용하는 SSD 상에서 구현될 수 있다.
- [0036] 본 발명의 유전자 분석 가속 장치(1)는 전술한 리드 매핑 과정 중 시딩 과정을 SSD 내부 연산으로 처리하여 호스트 시스템(2)의 메인 메모리 사용량을 줄이고, 호스트와 디바이스간 데이터 이동량을 감소시켜 오버헤드를 줄임으로써 유전자 분석 성능 및 에너지 효율을 향상시킬 수 있다.
- [0037] 본 발명의 일 실시예에 따른 유전자 분석 가속 장치(1)는 플래시 메모리(41), 매핑 테이블(2130), 게놈 호스트 인터페이스(211), 미니마이저 레벨 변환기(213), 페이지 버퍼(31) 및 인-페이지 필터(215)를 포함할 수 있다. 또한, 이 외에도 유전자 분석 가속 장치(1)는 NVMe 컨트롤러(10), 버퍼 컨트롤러(30), 플래시 메모리(41) 컨트롤러, FTL(Flash Translation Layer, 230) 등 일반적인 SSD의 펌웨어로 포함될 수 있는 구성들을 포함할 수 있다.
- [0038] 전술한 구성 중 게놈 호스트 인터페이스(211), 미니마이저 레벨 변환기(213) 및 인-페이지 필터(215)는 MTL(210)(Minimizer Translation Layer)에 포함될 수 있다. 또한, MTL(210) 및 FTL(230)은 스토리지 프로세서(20) 내에 포함될 수 있다.
- [0039] 전술한 구성 중 매핑 테이블(2130)은 미니마이저 레벨 변환기(213)에 포함된 형태로 존재할 수 있다. 또한, 인-페이지 필터(215)는 DMA(Direct Memory Access) 버퍼(2150)를 포함할 수 있다.
- [0040] 본 발명의 유전자 검사 장치에 포함되는 NVMe 컨트롤러(10)는 일반적인 NVMe SSD에서 사용되는 NVMe 컨트롤러(10)와 동일 내지 유사한 기능을 수행할 수 있다. NVMe 컨트롤러(10)는 NVMe SSD를 제어하는 데 사용되는 구성으로, 호스트 시스템(2)과 SSD 디바이스 간 통신을 처리하고, 장치 제어 관리, 데이터 쓰기 및 읽기, 에러 검사 및 복구, 메모리 관리 등의 기능을 수행할 수 있다.
- [0041] 본 발명의 유전자 검사 장치에 포함되는 버퍼 컨트롤러(30)는 일반적인 SSD에서 사용되는 버퍼 컨트롤러(30)와

동일 내지 유사한 기능을 수행할 수 있다. 버퍼 컨트롤러(30)는 페이지 버퍼(31)를 제어하는 데 사용되는 구성으로, 페이지 버퍼(31)와 다른 구성 간 데이터 흐름 제어, 데이터 손실 및 오류를 방지하기 위한 검사 및 복구 등의 다양한 기능을 수행할 수 있다.

[0042] 본 발명의 유전자 검사 장치에 포함되는 플래시 컨트롤러(40)는 일반적인 SSD에서 사용되는 플래시 메모리 컨트롤러와 동일 내지 유사한 기능을 수행할 수 있다. 플래시 컨트롤러(40)는 플래시 메모리(41)를 제어하는데 사용되는 구성으로, 플래시 메모리(41) 및 다른 구성 간 데이터 흐름 제어, 플래시 메모리(41) 관리 등의 기능을 수행할 수 있다.

[0043] 본 발명의 유전자 분석 가속 장치(1)에 포함된 각 구성들은 서로 유기적으로 동작하여 상호작용할 수 있다.

[0044] 본 발명의 유전자 분석 가속 장치(1)는 호스트 시스템(2)으로부터 호스트 커맨드를 전달받아 이를 수행할 수 있다. 일반적으로, 호스트 커맨드는 컴퓨터 시스템에서 호스트 시스템(2) 및 저장장치 간 상호 작용하기 위한 명령어 혹은 데이터 등을 의미하는 것일 수 있다.

[0045] 본 발명에서 유전자 분석 가속 장치(1)가 전달 받는 호스트 커맨드는 유전자 분석 가속 장치(1)의 각 구성을 제어하도록 새로 정의된 명령어일 수 있다. 새롭게 정의된 명령어는 플래그(flag)에 따라 종류가 구분될 수 있으며, 기존의 SLBA(Starting Logical block Address) 대신 미니마이저를 페이지 매핑 주소로 사용할 수 있다.

[0046] 도 2를 참조하여 본 발명의 일 실시예에 따른 플래시 메모리(41)를 설명하도록 한다.

[0047] 본 발명의 유전자 분석 가속 장치(1)의 플래시 메모리(41)는 페이지를 가진 비휘발적 저장장치(예를 들어, NAND 플래시 메모리, NOR 플래시 메모리, OneNAND 플래시 메모리 및 AG-AND 플래시 메모리 등)로 구현될 수 있다. 여기서 페이지는 일정한 크기의 메모리 블록을 구성하는 복수의 저장공간으로, 플래시 메모리(41)에 데이터를 저장하고 읽을 때 페이지 단위로 작업이 이루어질 수 있다. 각 페이지는 고유한 주소인 페이지 번호(page number)를 가질 수 있으며, 데이터를 읽고 저장하는데 페이지 번호를 이용하여 작업이 가능할 수 있다.

[0048] 도 2는 플래시 메모리(41)의 각 페이지에 참조 인덱스 데이터가 저장되어 있는 모습을 도시한 도면이다.

[0049] 본 발명의 플래시 메모리(41)는 참조 인덱스 정보를 저장하고 있을 수 있다. 구체적으로 플래시 메모리(41)의 각 페이지 내에는 참조 인덱스 데이터가 저장되어 있을 수 있다. 각각의 페이지는 복수의 엔트리로 구성되어 일정한 길이로 나뉜 참조 유전체의 각 미니마이저 및 그의 위치 배열을 나타내는 참조 인덱스 정보를 저장하고 있을 수 있다. 페이지는 저장하고 있는 미니마이저의 위치 배열에 관한 페이지 구성 정보를 저장하여 메타 데이터로 활용할 수 있다. 여기서 위치 배열은 해당 미니마이저가 참조 유전체의 어느 부분에 위치하는지를 나타낼 수 있는 정보를 의미할 수 있다.

[0050] 참조 인덱스는 전술한 것과 같이, 참조 유전체를 일정한 길이로 나누고, 일정한 길이로 나누어진 참조 유전체 각각의 대표 문자열인 미니마이저를 추출하여 해당 미니마이저와 그 위치 배열을 해시 테이블 형태로 나타낸 것이다. 여기서 참조 유전체는 유전체 서열을 의미하며, 일반적으로 염기서열분석이 완료되어 공공 데이터베이스로 구축되어 있는 표준 염기서열을 사용할 수 있다.

[0051] 플래시 메모리(41)의 각 페이지에는 참조 인덱스에서 추출한 미니마이저, 해당 미니마이저의 위치 배열 및 메타 데이터로 활용할 수 있는 위치 배열 정보(위치 배열 및 위치 배열 크기 등)가 저장되어 있을 수 있다. 도 2의 예시를 참조하여 설명하면, P1에는 총 3개의 미니마이저 및 미니마이저들의 위치 배열이 저장되어 있다. 0 부터 1번까지의 엔트리에는 미니마이저 'ACG'의 위치 배열이 저장되어 있으며, 2번 엔트리에는 미니마이저 'ACT'의 위치 배열, 3 내지 5번 엔트리에는 미니마이저 'AGA'의 위치 배열이 저장되어 있을 수 있다. 또한, 6 내지 8번 엔트리에는 각 {미니마이저, 해당 미니마이저의 위치 배열 크기}가 저장되어 있을 수 있으며, 각 위치 배열의 크기는 해당 페이지내에 저장되어 있는 미니마이저 위치 배열의 원소 개수(참조 유전체에 해당 미니마이저가 몇 개 있는지)를 의미할 수 있다. 다시 도 2의 예시를 참조하면, 엔트리 6번에는 {AGA, 3}, 7번에는 {ACT, 1}, 8번에는 {ACG, 2}가 저장되어 있을 수 있다. 저장된 미니마이저 및 위치 배열 정보를 포함하는 페이지 구성 정보는 메타 데이터로 활용되어 페이지 내부에서 미니마이저 검색에 활용될 수 있다.

[0052] 또한 플래시 메모리(41)는 일 미니마이저의 위치 배열이 한 페이지의 크기보다 큰 경우, 여러 페이지를 동시에 점유하여 참조 인덱스를 저장할 수 있다. 도 2의 P2 및 P3을 참조하여 설명하면, 미니마이저 'AGT'는 11에 해당하는 위치 배열 크기를 가질 수 있어 한 페이지 안에 모두 저장하지 못하고, P2의 0 내지 7번 엔트리에 위치 배열을 저장하고, 이어 다음 페이지인 P3에 나머지 위치 배열을 저장하였다. 이 경우에도 마찬가지로 {AGT, 해당

페이지에 저장된 위치 배열의 크기}의 형태로 페이지 구성 정보가 페이지의 마지막에 함께 저장 될 수 있다.

- [0053] 계놈 호스트 인터페이스(211)는 호스트로부터 호스트 커맨드 명령어를 받으면, 명령어에 포함된 미니마이저를 추출할 수 있다. 명령어는 메시지를 포함하고, 메시지는 찾고자하는 쿼리 미니마이저(Query minimizer)를 포함하고 있을 수 있다. 따라서, 계놈 호스트 인터페이스(211)는 명령어 메시지에서 쿼리 미니마이저를 추출할 수 있다.
- [0054] 일 실시예에 의하면 명령어는 NVMe 명령어일 수 있으며, 명령어의 메시지는 64B의 NVMe 메시지일 수 있다.
- [0055] 일 실시예에 의하면, 계놈 호스트 인터페이스(211)는 명령어를 받으면, 명령어에 포함된 해시 테이블 ID를 더 추출할 수 있다. 해시 테이블 ID는 후술할 매핑 테이블(2130)에서 상위 테이블의 주소로 이용될 수 있는 정보로, 명령어의 메시지는 쿼리 미니마이저의 해시 테이블 ID 정보를 더 포함하고 있을 수 있다. 따라서 계놈 호스트 인터페이스(211)는 쿼리 미니마이저의 해시 테이블 ID를 명령어 메시지에서 더 추출할 수 있다.
- [0056] 일 실시예에 의하면, 계놈 호스트 인터페이스(211)는 명령어가 내부 연산 명령인지 판단할 수 있다. 계놈 호스트 인터페이스(211)는 명령어의 내부 연산 명령 여부를 판단하여 명령어가 내부 연산 명령인 경우에만 쿼리 미니마이저 및 해시 테이블 ID를 추출하도록 할 수 있다.
- [0057] 미니마이저 레벨 변환기(213)는 매핑 테이블(2130)을 포함할 수 있다.
- [0058] 도 3 및 도 4를 참조하여 매핑 테이블(2130)에 관하여 설명하도록 한다.
- [0059] 도 3 및 도 4는 실시예에 따른 매핑 테이블(2130)을 도시한 도면이다.
- [0060] 매핑 테이블(2130)은 각 페이지 번호에 대표 미니마이저를 매핑시킨 테이블일 수 있다. 여기서 대표 미니마이저는 매핑시키고자 하는 해당 페이지에 저장된 복수의 미니마이저 중 어느 하나의 미니마이저를 의미하는 것일 수 있다. 일 실시예에 의하면, 대표 미니마이저는 페이지에 저장된 복수의 미니마이저 중 위치 배열의 크기가 가장 작은 미니마이저일 수 있다. 대표 미니마이저를 정하는 기준은 전술한 예시에 한정되지 않고 다양한 기준이 사용될 수 있다.
- [0061] 도 3의 예시를 참조하면, 페이지 번호 200에는 대표 미니마이저 'AAA', 페이지 번호 201에는 대표 미니마이저 'CAA', 페이지 번호 5167에는 대표 미니마이저 'GGT', 페이지 번호 5168에는 대표 미니마이저 'ACG', 물리페이지 번호 5169 내지 5171에는 미니마이저 'AGT'가 매핑되어 있을 수 있다.
- [0062] 일 실시예에 의하면, 매핑 테이블(2130)은 하위 테이블 및 상위 테이블로 구성된 2-레벨 테이블일 수 있다. 하위 테이블은 전술한 바와 같이, 페이지 번호에 대표 미니마이저를 매핑시킨 테이블일 수 있다. 상위 테이블은 페이지 번호를 복수개 묶어 복수개의 블록으로 나누고, 각 블록에 해시 테이블 ID를 매핑시킨 테이블일 수 있다. 즉, 매핑 테이블(2130)은 모든 페이지 번호에 대해 각 미니마이저를 매핑시킨 하위 테이블과 복수의 물리 페이지 넘버를 묶어 복수의 블록을 생성하고, 각 블록에 해시 테이블 ID를 매핑시킨 상위테이블로 구성될 수 있다.
- [0063] 도 4의 예시를 참조하면, 하위 테이블은 모든 페이지 번호에 대해 미니마이저가 매핑되어 있으며, 상위 테이블은 페이지 번호 200 내지 5167까지를 하나의 블록으로 묶은 후, 해시 테이블 ID를 0으로 하여 매핑시킬 수 있다. 또한, 상위 테이블은 페이지 번호 5168 내지 5172까지를 하나의 블록으로 묶은 후, 해시 테이블 ID를 1로 하여 매핑시킬 수 있다.
- [0064] 매핑 테이블(2130)이 전술한 바와 같이 2-레벨 테이블로 형성되는 경우, 해시 테이블 ID를 더 한정하여 탐색하면, 찾고자하는 쿼리 미니마이저의 페이지 주소를 더 효율적으로 찾을 수 있고, 내부 메모리의 사용량을 줄일 수 있는 장점이 있을 수 있다.
- [0065] 미니마이저 레벨 변환기(213)는 매핑 테이블(2130)을 이용하여 앞서 추출된 쿼리 미니마이저에 대해 매핑되어있는 페이지 번호를 구할 수 있다. 또한, 미니마이저 레벨 변환기(213)는 앞서 구한 페이지 번호에 따라 플래시 메모리(41)에 해당 페이지를 읽도록 읽기 요청을 보낼 수 있다.
- [0066] 미니마이저 레벨 변환기(213)는 명령어에서 추출된 쿼리 미니마이저에 대해, 매핑 테이블(2130)을 탐색하여 해당 미니마이저에 매핑된 페이지 번호를 구할 수 있다. 또한, 미니마이저 레벨 변환기(213)는 앞서 구한 페이지 번호를 이용하여 플래시 메모리(41)에 페이지 읽기 요청을 보내 해당 페이지를 읽어오도록 할 수 있다.
- [0067] 도 5를 참조하여 페이지 버퍼(31) 및 인-페이지 필터(215)를 설명하고자 한다.

- [0068] 도 5는 스토리지 프로세서(20) 및 페이지 버퍼(31)를 도시한 도면이다.
- [0069] 페이지 버퍼(31)는 페이지 읽기 요청에 따라 플래시 메모리(41)로부터 읽은 페이지를 임시적으로 저장할 수 있다.
- [0070] 도 5에 예시된 바와 같이, 페이지 버퍼(31)는 페이지 읽기 요청에 따라 읽은 201번 페이지를 저장할 수 있다.
- [0071] 페이지 버퍼(31)는 바람직하게는 디램(DRAM, Dynamic Random Access Memory)으로 구현될 수 있으나, 이에 한정되는 것은 아니며 SDRAM(Synchronous Dynamic Random Access Memory), DDR SDRAM(Double data rate synchronous dynamic random access memory), GDDR(Graphics Double Data Rate), HBM(High Bandwidth Memory), RDRAM(Rambus DRAM), MRAM((Magnetic Random Access Memory), PRAM(Phase-change Random Access Memory), ReRAM(Resistive Random Access memory) 등 다양한 저장장치로 구현될 수 있다.
- [0072] 다시 도 5를 참조하여 인-페이지 필터(215)에 대하여 설명하도록 한다.
- [0073] 인-페이지 필터(215)는 페이지 버퍼(31)에 저장된 페이지 내부에서 찾고자하는 쿼리 미니마이저의 위치 배열 및/혹은 위치 배열의 크기를 복사하여 저장할 수 있다. 인-페이지 필터(215)는 내부에 DMA 버퍼(2150)를 포함하여 미니마이저 위치 배열을 복사하여 DMA 버퍼(2150)에 저장할 수 있다. 도 5의 예시를 참조하면, 인-페이지 필터(215)는 페이지에 저장된 {ATA, 2} {AGG, 2} 및 {AGA, 1}의 미니마이저 위치 배열 정보를 참조하여 미니마이저 'ATA', 'AGG' 및 'AGA' 중 쿼리 미니마이저 'AGG'를 찾아 위치 배열 및/혹은 위치 배열의 크기를 복사하고, 이를 DMA 버퍼(2150)에 저장할 수 있다.
- [0074] 인-페이지 필터(215) 내 DMA 버퍼(2150)에 쿼리 미니마이저의 위치 배열 정보가 저장되면, 계층 호스트 인터페이스(211)는 NVMe 메시지를 통해 내부 연산 명령을 완료할 수 있다.
- [0075] FTL(230)은 플래시 메모리(41)의 특성을 숨겨 운영체제가 논리 블록 장치로 입출력을 관리할 수 있도록 하는 SSD 펌웨어 구성요소 중 하나로, 매핑 테이블(2130)을 포함하여 논리적 블록 매핑(logical block mapping) 운영체제가 입출력을 관리하는 단위인 논리적 블록 주소(logical block address)를 플래시 메모리(41)가 읽기/쓰기 작업을 페이지 주소에 대응되도록 할 수 있다. 본 발명의 유전자 분석 가속 장치(1)에 포함되는 FTL(230)은 일반적인 SSD에 포함되는 FTL(230)과 동일 내지 유사하여 본 발명에서 FTL(230)에 관한 자세한 설명은 생략하도록 한다.
- [0077] 이하 실험예에서 실험을 통해 본 발명의 유전자 분석 가속 장치(1)의 메모리 사용량 및 데이터 이동 시간을 비교하도록 한다.
- [0078] <실험예 1>
- [0079] 본 실험예에서는 Intel i5-4690, 4 cores CPU, 16GB 메모리 및 Linux kernel 3.19의 호스트 시스템(2)을 사용하였다. 또한, 1GHz Dual-Core Cortex-A9 Embedded Core, DDR3 1GB DRAM, 2TB MLC 및 8-Channel 8-way NAND Flash, PCIe Gen 2.0 8-lane NVMe Host 인터페이스의 OpenSSD를 사용하였다.
- [0080] 본 실험예에서는 기존의 DRAM-based system, SSD 내부 연산을 하지 않은 유전자 분석 시스템(SSD-baseline system) 과 본 발명의 유전자 분석 가속 장치(1) GenomeSSD를 사용하여 유전자 분석을 실시하고, 메모리 사용량 및 데이터 이동량을 측정하였다.
- [0081] 도 6a는 기존의 DRAM-based system, SSD-baseline system 및 본 발명의 유전자 분석 가속 장치(1) GenomeSSD를 사용하여 유전자 분석을 하였을 때, 메모리 사용량을 측정한 그래프이다.
- [0082] 도 6a에 따르면, SSD-baseline system 및 GenomeSSD는 기존의 DRAM-based system에 비하여 메모리 사용량이 90.38% 감소한 것으로 나타났다. 따라서, 본 발명의 유전자 분석 가속 장치(1)는 기존의 DRAM-based system에 비하여 현저한 메모리 사용량 감소 효과를 가질 수 있다.
- [0083] 도 6b는 SSD-baseline system 및 본 발명의 유전자 분석 가속 장치(1) GenomeSSD를 사용하여 유전자 분석을 하였을 때, 데이터 이동 시간을 측정한 그래프이다.
- [0084] 도 6b에 따르면, SSD-baseline system는 GenomeSSD에 비하여 데이터 이동 시간이 2.84배에 이르는 것으로 나타났다. 이는 내부 연산 과정이 없는 SSD-baseline system에 비해 GenomeSSD는 내부 연산을 통해 데이터 이동량을 99.98%가량 줄임으로써 나타난 것으로 보인다. 따라서, 본 발명의 유전자 분석 가속 장치(1)는 SSD-baseline system에 비하여 현저한 데이터 이동량 및 이동 시간 감소를 가질 수 있다.

- [0085] 결론적으로, 실험예에 따르면, 본 발명의 유전자 분석 가속 장치(1)는 SSD 인-스토리지 프로세싱을 통해 기존의 유전자 분석 장치들에 비하여 현저한 메모리 사용량, 데이터 이동량 및 데이터 이동 시간 감소의 효과를 가질 수 있다.
- [0087] 도 7을 참조하여 본 발명의 유전자 분석 가속 방법에 관하여 설명하도록 한다.
- [0088] 도 7은 일 실시예에 따른 유전자 분석 가속 방법의 순서도이다.
- [0089] 일 실시예에 의하면, 본 발명의 SSD 인-스토리지 프로세싱 유전자 분석 가속 방법은 플래시 메모리(41)의 페이지에 참조 유전체를 일정한 길이로 나눈 각 유전체의 대표 문자열인 복수의 미니마이어 및 복수의 미니마이어의 위치 배열을 저장하는 단계(S101)를 포함할 수 있다.
- [0090] 또한, 본 발명의 SSD 인-스토리지 프로세싱 유전자 분석 가속 방법은 페이지의 페이지 번호와 해당 페이지에 저장된 복수의 미니마이어 중 하나인 대표 미니마이어를 매핑시킨 매핑 테이블(2130)을 작성하는 단계(S102)를 포함할 수 있다.
- [0091] 또한, 본 발명의 SSD 인-스토리지 프로세싱 유전자 분석 가속 방법은 계층 호스트 인터페이스(211)에서 명령어를 받아 상기 명령어에 포함된 쿼리 미니마이어를 추출하는 단계(S103)를 포함할 수 있다.
- [0092] 또한, 본 발명의 SSD 인-스토리지 프로세싱 유전자 분석 가속 방법은 미니마이어 레벨 변환기(213)에서 매핑 테이블(2130)을 탐색하여 추출한 쿼리 미니마이어에 매핑되어있는 페이지 번호를 산출하여 플래시 메모리(41)에 페이지 읽기 요청을 보내는 단계(S104)를 포함할 수 있다.
- [0093] 또한, 본 발명의 SSD 인-스토리지 프로세싱 유전자 분석 가속 방법은 페이지 읽기 요청에 따라 읽은 페이지를 복사하여 페이지 버퍼(31)에 저장하는 단계(S105)를 포함할 수 있다.
- [0094] 또한, 본 발명의 SSD 인-스토리지 프로세싱 유전자 분석 가속 방법은 DMA 버퍼(2150)를 포함하는 인-페이지 필터(215)에, 페이지 버퍼(31)에 저장된 페이지 내부에서 쿼리 미니마이어에 해당하는 위치 배열을 복사하여 DMA 버퍼(2150)에 저장하는 단계(S106)를 포함할 수 있다.
- [0095] 또한, 이외에도 본 발명의 SSD 인-스토리지 프로세싱 유전자 분석 가속 방법은 DMA 버퍼(2150)에 쿼리 미니마이어의 위치 배열이 저장되면 계층 호스트 인터페이스(211)에서 메시지를 보내 내부 연산 명령을 완료하는 단계를 더 포함할 수 있다.
- [0096] 또 다른 실시예에 의하면, 쿼리 미니마이어를 추출하는 단계는, 명령어에 포함된 해시 테이블 ID를 더 추출하고, 페이지 읽기 요청을 보내는 단계는, 해시 테이블 ID를 통해 상위 테이블에서 블록을 먼저 산출하고, 하위 테이블에서 미니마이어에 대해 매핑되어있는 페이지 번호를 산출하여 플래시 메모리에 페이지 읽기 요청을 보내는 것을 특징으로 할 수 있다.
- [0098] 이상 SSD 인-스토리지 프로세싱 유전자 분석 가속 장치(1) 및 방법의 여러 실시예에 대해 설명하였으나, 장치 또는 방법은 오직 상술한 실시예에 한정되는 것은 아니다. 해당 기술 분야에서 통상의 지식을 가진 자가 상술한 실시예를 기초로 수정 및 변형하여 구현할 수 있는 다른 다양한 장치나 방법 역시 상술한 유전자 분석 가속 장치(1) 또는 유전자 분석 가속 방법의 일 실시예가 될 수 있다. 예를 들어, 설명된 방법(들)이 설명된 바와 다른 순서로 수행되거나, 및/또는 설명된 시스템, 구조, 장치, 회로 등의 구성 요소(들)가 설명된 바와 다른 형태로 결합, 연결 또는 조합되거나 다른 구성 요소 또는 균등물 등에 의하여 대체 또는 치환되더라도, 상술한 유전자 분석 가속 장치(1) 및/또는 유전자 분석 가속 방법의 일 실시예가 될 수 있다.
- [0099] 본원 발명의 실시예 등과 관련된 기술 분야에서 통상의 지식을 가진 자는 상기 기재의 본질적인 특성에서 벗어나지 않는 범위에서 변형된 형태로 구현될 수 있음을 이해할 수 있을 것이다. 그러므로, 개시된 방법들은 한정적인 관점이 아닌 설명적 관점에서 고려되어야 한다. 본 발명의 범위는 발명의 상세한 설명이 아닌 특허청구 범위에 나타나며, 그와 동등한 범위 내에 있는 모든 차이점은 본 발명의 범위에 포함되는 것으로 해석되어야 한다.

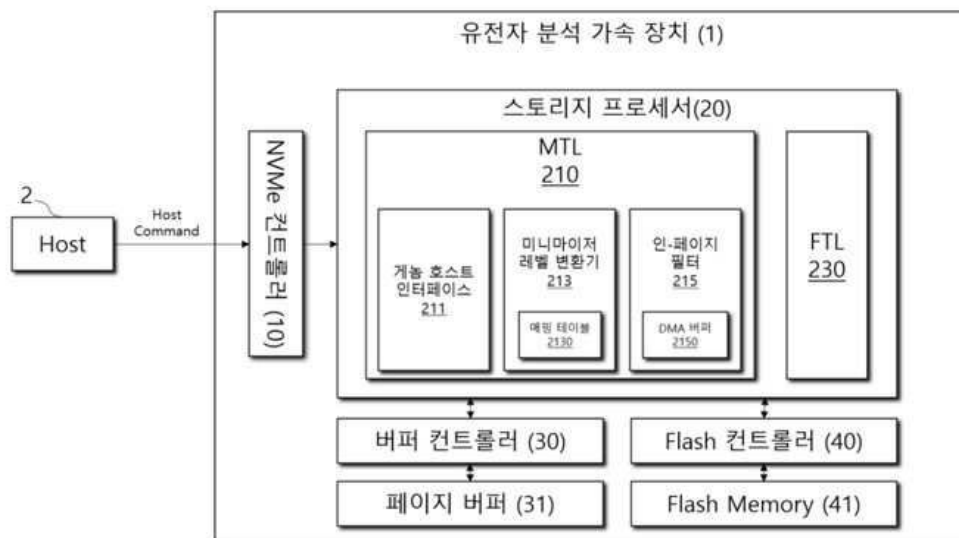
부호의 설명

- [0100] 1: 유전자 분석 가속 장치
2: 호스트 시스템

- 10: NVMe 컨트롤러
- 20: 스토리지 프로세서
- 30: 버퍼 컨트롤러
- 31: 페이지 버퍼
- 40: 플래시 컨트롤러
- 41: 플래시 메모리
- 210: MTL
- 211: 계층 호스트 인터페이스
- 213: 미니마이저 레벨 변환기
- 215: 인-페이지 필터
- 230: FTL
- 2130: 매핑 테이블
- 2150: DMA 버퍼

도면

도면1



도면2

41

page number	0	1	2	3	4	5	6	7	8
P1	ACG	ACG	ACT	AGA	AGA	AGA	{AGA, 3}	{ACT, 1}	{ACG, 2}
P2	AGT	AGT	AGT	AGT	AGT	AGT	AGT	AGT	{AGT, 8}
P3	AGT	AGT	AGT	AGG				{AGG, 1}	{AGT, 3}

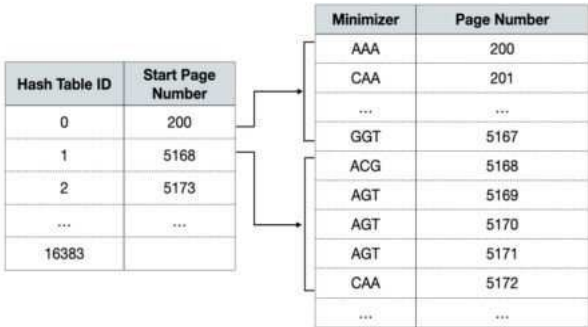
도면3

2130

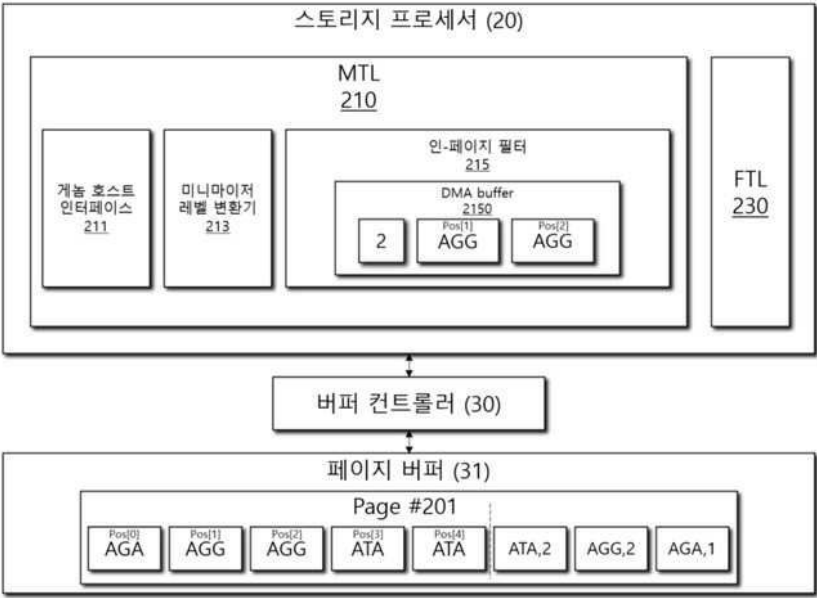
Minimizer	Page Number
AAA	200
CAA	201
...	...
GGT	5167
ACG	5168
AGT	5169
AGT	5170
AGT	5171
CAA	5172
...	...

도면4

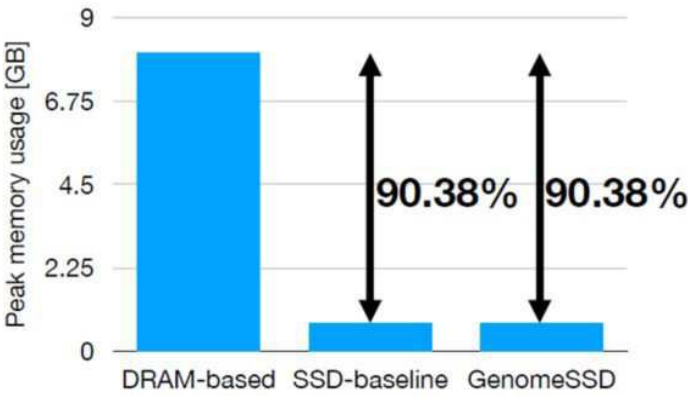
2130



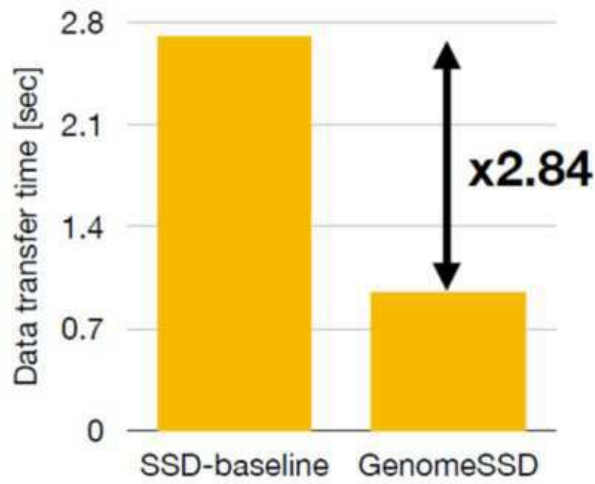
도면5



도면6a



도면6b



도면7

