



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2022-0157374
(43) 공개일자 2022년11월29일

- (51) 국제특허분류(Int. Cl.)
G11C 11/4093 (2021.01) G11C 11/408 (2006.01)
G11C 11/4097 (2006.01) G11C 5/02 (2006.01)
G11C 7/10 (2021.01) H01L 27/108 (2006.01)
- (52) CPC특허분류
G11C 11/4093 (2018.05)
G11C 11/4085 (2013.01)
- (21) 출원번호 10-2022-7028869
- (22) 출원일자(국제) 2022년12월23일
심사청구일자 없음
- (85) 번역문제출일자 2022년08월22일
- (86) 국제출원번호 PCT/US2020/066985
- (87) 국제공개번호 WO 2021/194597
국제공개일자 2021년09월30일
- (30) 우선권주장
16/827,542 2020년03월23일 미국(US)
- (71) 출원인
인텔 코포레이션
미합중국 캘리포니아 95054 산타클라라 미션 칼리지 블러바드 2200
- (72) 발명자
샤르마 아비셰크
미국 오리건주 97124 힐스보로 아파트먼트 541 노스이스트 오렌코 스테이션 루프 967
- 사토 노리유키
미국 오리건주 97124 힐스보로 노스이스트 오렌코 스테이션 파크웨이 에이318 1299
(뒷면에 계속)
- (74) 대리인
제일특허법인(유)

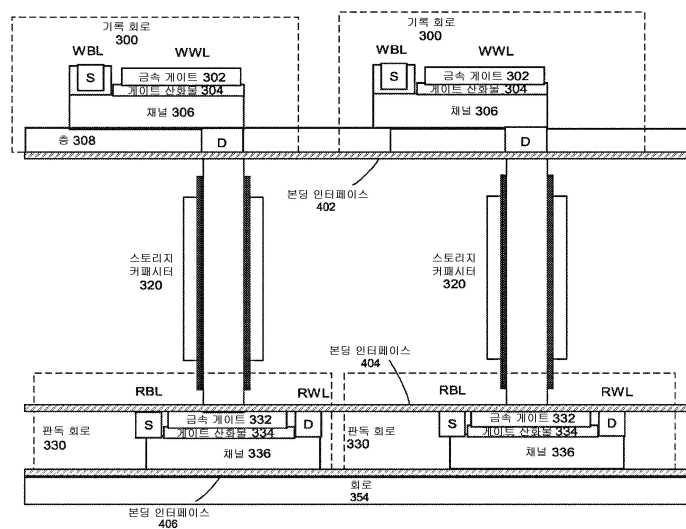
전체 청구항 수 : 총 39 항

(54) 발명의 명칭 백엔드 메모리를 이용한 니어 메모리 컴퓨터

(57) 요약

본 명세서의 예는, 이득 셀 eDRAM 메모리 셀을 포함하는 메모리 디바이스 - 이득 셀 eDRAM 메모리 셀은, 스토리지 셀 위에 적어도 부분적으로 형성된 기록 회로와, 스토리지 셀 아래에 적어도 부분적으로 형성된 판독 회로를 포함할 수 있음 - 와, 메모리 디바이스에 본딩된 컴퓨터 니어 메모리 디바이스와, 프로세서와, 메모리 디바이스로부터 프로세서로의 인터페이스를 포함하는 시스템에 관한 것이다. 일부 예에서, 메모리 디바이스의 출력을 제곱해서 SRAM 메모리 디바이스의 출력 판독 레이트를 애플레이트하는 회로가 포함되고, 이 회로는 컨트롤러, 멀티플렉서, 또는 레지스터 중 하나 이상을 포함한다. 메모리 디바이스의 표면은 메모리 디바이스 혹은 다른 회로에 본딩될 수 있다. 일부 예에서, 판독 회로를 가진 층이 스토리지 셀을 가진 층에 본딩될 수 있다. 본 명세서에 기술된 기술을 사용하여 임의의 층을 함께 본딩할 수 있다.

대표도



(52) CPC특허분류

G11C 11/4097 (2013.01)
G11C 5/025 (2013.01)
G11C 7/1078 (2013.01)
H01L 27/10808 (2013.01)
H01L 27/10897 (2013.01)
G11C 2211/4066 (2013.01)

(72) 발명자

아타나소프 사라

미국 오리건주 97003 비버턴 사우스웨스트 프랜시스 스트리트 18045

썸블 후세인 에킨

미국 오리건주 97205 포틀랜드 아파트먼트 417 사우스웨스트 비스타 애비뉴 1000

첸 그레고리 케이

미국 오리건주 97229 포틀랜드 노스웨스트 125번 애비뉴 4350

크나그 필

미국 오리건주 97124 힐스보로 아파트먼트 지211 노스이스트 오렌코 스테이션 파크웨이 1241

크리슈나무르시 람

미국 오리건주 97229 포틀랜드 노스웨스트 베르타니 스트리트 13042

유 희 재

미국 오리건주 97124 힐스보로 알에이3-252 엘람영 파크웨이 5200

레 반 에이치

미국 오리건주 97007 비버턴 사우스웨스트 라빈 드라이브 8779

명세서

청구범위

청구항 1

장치로서,

이득 셀 임베디드 동적 랜덤 액세스 메모리(eDRAM) 디바이스를 포함하고, 상기 이득 셀 eDRAM 디바이스는,
기록 회로와,

스토리지 디바이스와,

판독 회로

를 포함하며,

상기 판독 회로는 적어도 부분적으로 상기 기록 회로 아래에 상기 기록 회로의 평면과는 다른 평면에 위치되고,
상기 스토리지 디바이스는 상기 기록 회로를 상기 판독 회로에 연결시키며, 상기 기록 회로와 상기 판독 회로
중 하나 이상이 상기 스토리지 디바이스에 본딩되는,

장치.

청구항 2

제 1 항에 있어서,

상기 기록 회로는,

드레인 영역을 가진 층과,

채널 층과,

게이트 산화물 층과,

기록 워드 라인 인터페이스를 가진 게이트 층과,

기록 비트 라인 인터페이스를 가진 소스 영역

을 포함하는,

장치.

청구항 3

제 1 항에 있어서,

상기 판독 회로는,

채널 영역과,

게이트 산화물 층과,

게이트 층과,

판독 비트 라인 인터페이스를 가진 소스 영역과,

판독 워드 라인 인터페이스를 가진 드레인 영역

을 포함하는,

장치.

청구항 4

제 1 항에 있어서,
상기 기록 회로는,
층과,
드레인 영역과,
상기 층을 통해 상기 드레인 영역을 상기 스토리지 디바이스에 연결하는 도전성 영역과,
채널 층과,
게이트 산화물 층과,
게이트 층과,
기록 비트 라인 인터페이스를 가진 소스 영역
을 포함하는,
장치.

청구항 5

제 1 항에 있어서,
상기 판독 회로는,
게이트 영역과,
상기 게이트 영역 위에 적어도 부분적으로 형성된 게이트 산화물 영역과,
상기 게이트 산화물 영역 위에 적어도 부분적으로 형성된 채널 영역과,
상기 채널 영역의 일부 위에 형성된 소스 영역과,
상기 채널 영역의 일부 위에 형성된 드레인 영역
을 포함하는,
장치.

청구항 6

제 1 항에 있어서,
상기 기록 회로는,
채널 영역과,
상기 채널 영역 위에 적어도 부분적으로 형성된 게이트 산화물 영역과,
상기 게이트 산화물 영역 위에 적어도 부분적으로 형성된 게이트 영역과,
상기 채널 영역의 일부 위에 형성된 소스 영역과,
상기 채널 영역의 일부 위에 형성된 드레인 영역
을 포함하는,

장치.

청구항 7

제 1 항에 있어서,
상기 기록 회로는,
채널 영역과,
상기 채널 영역의 일부 아래에 형성된 소스 영역과,
상기 채널 영역의 일부 아래에 형성된 드레인 영역과,
상기 채널 영역 위에 적어도 부분적으로 형성된 게이트 산화물 영역과,
상기 게이트 산화물 영역 위에 적어도 부분적으로 형성된 게이트 영역
을 포함하는,
장치.

청구항 8

제 1 항에 있어서,
상기 스토리지 디바이스는, 적어도 하나의 커패시터 및 상기 적어도 하나의 커패시터와 컨택트하는 유전체를 포
함하는,
장치.

청구항 9

제 1 항에 있어서,
상기 스토리지 디바이스는 COB(Capacitor Over Bitline)를 포함하는,
장치.

청구항 10

제 1 항에 있어서,
상기 기록 회로는, 폴리-Si, Si, Ge, 폴리-Ge, III-V, GaN, MoS₂, WSe₂, MoSe₂, WSe₂, InS, HfS, ZnS, ZnSe,
In₂O₃, ZnO, AZO, IGZO 또는 IZO 중 하나 이상을 포함하는,
장치.

청구항 11

제 1 항에 있어서,
상기 기록 회로는, 본딩 재료를 사용해서 상기 스토리지 디바이스에 본딩되는,
장치.

청구항 12

제 11 항에 있어서,

상기 본딩 재료는, 이산화 실리콘(SiO₂), 산질화 실리콘(SiON), 탄소 도핑된 산질화 실리콘(SiOCN), 실리콘 탄소 질화물(SiCN), 또는 산탄화 실리콘(SiOC) 중 하나 이상을 포함하는,

장치.

청구항 13

제 1 항에 있어서,

상기 관독 회로는 본딩 재료를 사용해서 상기 스토리지 디바이스에 본딩되는,

장치.

청구항 14

제 13 항에 있어서,

상기 본딩 재료는, 이산화 실리콘(SiO₂), 산질화 실리콘(SiON), 탄소 도핑된 산질화 실리콘(SiOCN), 실리콘 탄소 질화물(SiCN), 또는 산탄화 실리콘(SiOC) 중 하나 이상을 포함하는,

장치.

청구항 15

제 1 항에 있어서,

상기 장치는, 컴퓨트 니어 메모리(compute near memory; CNM) 회로를 포함하고,

상기 CNM 회로는 상기 관독 회로 또는 상기 기록 회로에 본딩되는,

장치.

청구항 16

제 1 항에 있어서,

상기 이득 셀 eDRAM 디바이스로부터의 출력을 상이한 유효 출력 레이트로 제공하기 위한 메모리 에물레이션 회로를 포함하는,

장치.

청구항 17

이득 셀 메모리 디바이스를 형성하는 방법으로서,

제 1 기판 상에, 워드 라인 드라이버, 관독 회로 및 활성 핀(active fin)을 포함하는 제 1 영역을 형성하는 단계와,

상기 제 1 영역 위에 유전체 영역을 형성하는 단계와,

상기 유전체 영역에 적어도 하나의 저장 영역을 형성하는 단계와,

상기 유전체 영역 상에 비실리콘 기반(non-silicon-based) 기록 회로의 층을 형성하는 단계와,

상기 비실리콘 기반 기록 회로의 층 상에 제 2 기판을 본딩하는 단계와,
상기 제 1 기판을 제거해서 상기 활성 핀을 노출시키는 단계
를 포함하는,
방법.

청구항 18

제 17 항에 있어서,
상기 비실리콘 기반 기록 회로의 층 상에 제 2 기판을 본딩하는 단계는, 상기 제 2 기판과 상기 비실리콘 기반 기록 회로의 층의 대향하는 표면에 본딩 재료를 도포하는 단계를 포함하고,
상기 본딩 재료는, 이산화 실리콘(SiO_2), 산질화 실리콘(SiON), 탄소 도핑된 산질화 실리콘(SiOCN), 실리콘 탄소 질화물(SiCN), 또는 산탄화 실리콘(SiOC) 중 하나 이상을 포함하는,
방법.

청구항 19

제 18 항에 있어서,
상기 비실리콘 기반 기록 회로의 층 상에 제 2 기판을 본딩하는 단계는,
이산화 실리콘(SiO_2), 산질화 실리콘(SiON), 탄소 도핑된 산질화 실리콘(SiOCN), 실리콘 탄소 질화물(SiCN), 또는 산탄화 실리콘(SiOC) 중 하나 이상을 사용해서 상기 본딩 재료를 도포하는 단계와,
대향하는 표면의 본딩 재료를 서로 가압하는 단계
를 포함하는,
방법.

청구항 20

제 17 항에 있어서,
상기 유전체 영역 상에 기록 회로의 층을 형성하는 단계는, 상기 유전체 영역의 표면에 기록 회로의 층을 본딩 하는 단계를 포함하는,
방법.

청구항 21

제 17 항에 있어서,
상기 제 2 기판에 회로를 형성하는 단계
를 더 포함하는 방법.

청구항 22

시스템으로서,
적어도 하나의 프로세서와,

상기 적어도 하나의 프로세서에 연결된 적어도 하나의 메모리를 포함하되,

상기 적어도 하나의 메모리는, 이득 셀 임베디드 동적 랜덤 액세스 메모리(eDRAM) 디바이스를 포함하며,

상기 이득 셀 eDRAM 디바이스는,

기록 회로와,

스토리지 디바이스와,

판독 회로

를 포함하며,

상기 판독 회로는 적어도 부분적으로 상기 기록 회로 아래에 상기 기록 회로의 평면과는 다른 평면에 위치되고, 상기 스토리지 디바이스는 상기 기록 회로를 상기 판독 회로에 연결시키며, 상기 기록 회로와 상기 판독 회로 중 하나 이상이 상기 스토리지 디바이스에 본딩되는,

시스템.

청구항 23

제 22 항에 있어서,

상기 스토리지 디바이스는, 적어도 하나의 커패시터 및 상기 적어도 하나의 커패시터와 컨택트하는 유전체를 포함하는,

시스템.

청구항 24

제 22 항에 있어서,

상기 기록 회로는, 폴리-Si, Si, Ge, 폴리-Ge, III-V, GaN, MoS₂, WSe₂, MoSe₂, WSe₂, InS, HfS, ZnS, ZnSe, In₂O₃, ZnO, AZO, IGZO 또는 IZO 중 하나 이상을 포함하는,

시스템.

청구항 25

제 22 항에 있어서,

상기 판독 회로는, 본딩 재료를 사용해서 상기 스토리지 디바이스에 본딩되고,

상기 본딩 재료는, 이산화 실리콘(SiO₂), 산질화 실리콘(SiON), 탄소 도핑된 산질화 실리콘(SiOCN), 실리콘 탄소 질화물(SiCN), 또는 산탄화 실리콘(SiOC) 중 하나 이상을 포함하는,

시스템.

청구항 26

장치로서,

이득 셀 임베디드 동적 랜덤 액세스 메모리(eDRAM) 메모리 셀을 포함하는 메모리 디바이스 - 상기 이득 셀 eDRAM 메모리 셀은, 스토리지 셀 위에 적어도 부분적으로 형성된 기록 회로와, 상기 스토리지 셀 아래에 적어도 부분적으로 형성된 판독 회로를 포함함 - 와,

상기 메모리 디바이스에 본딩된 컴퓨터 니어 메모리 디바이스
를 포함하는,
장치.

청구항 27

제 26 항에 있어서,
상기 컴퓨터 니어 메모리 디바이스는, 벡터-벡터 곱셈, 곱셈 및 누산(multiplication and accumulation(MAC)), AND, OR, XOR, NOT, 신경망 활성화 연산(neural network activation operations), soft-max의 멱승(exponentiation for soft-max), 데이터 압축, 암호화, 미디어 가속기, 문자열 매칭 또는 정규 표현식 가속기, 또는 포인터 추적(pointer chasing) 중 하나 이상을 수행하는 것인,
장치.

청구항 28

제 26 항에 있어서,
상기 메모리 디바이스의 출력을 제공해서 제 2 메모리 디바이스의 출력을 에플레이트하는 에플레이션 회로를 더 포함하는 장치.

청구항 29

제 28 항에 있어서,
상기 에플레이션 회로는 컨트롤러, 멀티플렉서, 또는 레지스터 중 하나 이상을 포함하는,
장치.

청구항 30

제 26 항에 있어서,
상기 메모리 디바이스 및 상기 컴퓨터 니어 메모리 디바이스에 연결된 프로세서
를 더 포함하는 장치.

청구항 31

제 26 항에 있어서,
상기 컴퓨터 니어 메모리 디바이스는 본딩 재료를 사용해서 상기 메모리 디바이스에 본딩되고,
상기 본딩 재료는, 이산화 실리콘(SiO₂), 산질화 실리콘(SiON), 탄소 도핑된 산질화 실리콘(SiOCN), 실리콘 탄소 질화물(SiCN), 또는 산탄화 실리콘(SiOC) 중 하나 이상을 포함하는,
장치.

청구항 32

시스템으로서,

이득 셀 임베디드 동적 랜덤 액세스 메모리(eDRAM) 메모리 셀을 포함하는 메모리 디바이스 - 상기 이득 셀 eDRAM 메모리 셀은, 스토리지 셀 위에 적어도 부분적으로 형성된 기록 회로와, 상기 스토리지 셀 아래에 적어도 부분적으로 형성된 판독 회로를 포함함 - 와,

상기 메모리 디바이스에 본딩된 컴퓨터 니어 메모리 디바이스와,

프로세서와,

상기 메모리 디바이스로부터 상기 프로세서로의 인터페이스

를 포함하는 시스템.

청구항 33

제 32 항에 있어서,

상기 컴퓨터 니어 메모리 디바이스는, 벡터-벡터 곱셈, 곱셈 및 누산(MAC), AND, OR, XOR, NOT, 신경망 활성화 연산, soft-max의 역승, 데이터 압축, 암호화, 미디어 가속기, 문자열 매칭 또는 정규 표현식 가속기, 또는 포 인터 추적 중 하나 이상을 수행하는 것인,

시스템.

청구항 34

제 32 항에 있어서,

상기 메모리 디바이스의 출력을 제공해서 SRAM 메모리 디바이스의 출력 판독 레이트를 에플레이트하는 에플레이션 회로

를 더 포함하는 시스템.

청구항 35

제 34 항에 있어서,

상기 에플레이션 회로는 컨트롤러, 멀티플렉서, 또는 레지스터 중 하나 이상을 포함하는,

시스템.

청구항 36

제 35 항에 있어서,

상기 컨트롤러는, 상기 메모리 디바이스의 다수의 뱅크로의 메모리 액세스 요청의 전달을 제어하고,

상기 멀티플렉서는, 상기 컨트롤러에 의한 선택에 따라서, 상기 메모리 디바이스로부터 상기 컴퓨터 니어 메모리로 데이터를 송신하거나 혹은 상기 멀티플렉서는 상기 컴퓨터 니어 메모리로부터 데이터를 수신해서 상기 메모리 디바이스에 제공하는,

시스템.

청구항 37

제 32 항에 있어서,

상기 레지스터는, 상기 메모리 디바이스로부터 출력되는 데이터를, 상기 컴퓨터 니어 메모리 디바이스에 출력하

기 이전에, 버퍼링하는,
시스템.

청구항 38

제 32 항에 있어서,

상기 컴퓨터 니어 메모리 디바이스는, 이산화 실리콘(SiO2), 산질화 실리콘(SiON), 탄소 도핑된 산질화 실리콘(SiOCN), 실리콘 탄소 질화물(SiCN), 또는 산탄화 실리콘(SiOC) 중 하나 이상을 포함하는 본딩 재료를 사용해서 상기 메모리 디바이스에 본딩되는,

시스템.

청구항 39

제 32 항에 있어서,

상기 프로세서에 통신 가능하게 연결된 네트워크 인터페이스
를 더 포함하는 시스템.

발명의 설명

기술 분야

[0001] 우선권 주장

[0002] 본 출원은 35 U.S.C. § 365(c) 하에서 2020년 3월 23일에 출원된 "COMPUTE NEAR MEMORY WITH BACKEND MEMORY"라는 명칭의 미국 출원 제 16/827,542호의 우선권을 주장하며, 그 전체 내용은 본 명세서에 참조로서 포함된다.

[0003] 인공 지능(AI) 워크로드, 머신 러닝, 딥 러닝, 그래프 처리 등과 같은 최신 데이터 집약적 애플리케이션은 많은 양의 데이터를 처리한다. 예를 들어, 딥 러닝(DL) 워크로드는 트레이닝 태스크와 추론 태스크 모두에 대해 많은 양의 NN(Neural Network) 파라미터에 대한 행렬 연산을 수행하는 것을 포함할 수 있다. 이들 데이터-집약적 애플리케이션(data-intensive application)은, 엄청난 양의 데이터를 보관할 수 있는 대용량 스토리지를 필요로 하고, 데이터를 빠르고 효율적으로 처리할 수 있는 엄청난 컴퓨터 기능을 필요로 한다. 데이터-집약적 워크로드를 위한 스토리지와 컴퓨터 효율성의 측면을 모두 해결하는 것은 어려운 일이다.

도면의 간단한 설명

[0004] 도 1은 메모리 디바이스의 예를 도시한다.

도 2는 실시예의 예를 도시한다.

도 3a 내지 도 3d는 이득 셀 eDRAM 구조의 단면을 도시한다.

도 4는 공유 층을 가진 병렬식(side-by-side) 이득 셀 디바이스의 예를 도시한다.

도 5는, 도 4의 구조의 평면도의 예를 도시한다.

도 6은 일부 실시예에 따른 이득 셀 eDRAM 구조의 단면을 도시한다.

도 7a 내지 도 7c는 상부 게이트 구조 및 하부 게이트 구조의 다양한 예를 도시한다.

도 8은 다중 이득 셀 eDRAM 디바이스를 갖는 메모리 디바이스의 단면도를 도시한다.

도 9는 이득 셀 트랜지스터의 일부 실시예의 회로 표현을 도시한다.

도 10a 내지 도 10f는 메모리 디바이스를 형성하는 예를 도시한다.

도 11은 프로세스를 도시한다.

도 12a 내지 도 12c는 이득 셀 메모리 디바이스의 단면도를 도시한다.

도 13은 이득 셀 메모리 디바이스의 다른 단면도를 도시한다.

도 14는 다양한 디바이스의 평면도의 예를 도시한다.

도 15는 니어 메모리(near-memory) 및 니어 메모리 컴퓨터 블록을 가진 시스템의 예시적인 논리 블록도를 도시한다.

도 16은 제 1 메모리 타입의 메모리가 다른 메모리 타입의 출력을 에뮬레이트하는데 사용될 수 있는 회로를 가진 예시적인 시스템을 도시한다.

도 17은 다수의 뱅크에 액세스하기 위해 전달될 수 있는 신호의 예를 도시한다.

도 18은, 다수의 SRAM 뱅크가 액세스되고, 뱅크로부터의 데이터가 연결되거나 혹은 결합되어서 데이터를 컴퓨팅 엔진에 제공하는 예시적인 시스템을 도시한다.

도 19는 임베디드 DRAM 메모리로부터 출력된 데이터가 분할되어서, 다수의 프로세싱 요소(PE)에 제공되는 임베디드 DRAM 메모리의 예를 도시한다.

도 20은 더 빠른 메모리 디바이스를 에뮬레이트하는 속도로 메모리 디바이스로부터의 데이터를 판독하기 위한 예시적인 프로세스를 도시한다.

도 21은 다양한 시스템의 예를 도시한다.

도 22는 시스템을 도시한다.

도 23은 환경을 도시한다.

발명을 실시하기 위한 구체적인 내용

[0005] 컴퓨터 시스템은, 데이터에 대한 신속한 액세스를 제공하기 위해서, 동일한 다이에 또는 MCM(multi-chip module)에 통합된 RAM(random access memory)를 포함한다. 컴퓨터 시스템에서 랜덤 액세스 메모리로서 일반적으로 DRAM(Dynamic Random Access Memory)이 사용된다. 그러나, DRAM이 커패시터를 사용해서 데이터를 저장하는 경우, DRAM을 사용하려면 저장된 데이터와 대응하는 회로 및 타이밍 디바이스를 리프레시해야 할 수 있다. DRAM 기반 메모리는, 비트 라인을 충전하는 데 시간이 걸리기 때문에, 판독 및 기록 대기 시간이 발생할 수 있다. 예를 들어, COB(Capacitor Over Bitline)는 비트 라인에 전하를 공급하는 커패시터이다. 경우에 따라서, 판독 동작의 경우, 커패시터를 사용함으로써 인한 RC(resistance capacitance) 시정수 지연이 주어지면, 비트 라인을 충전하는 시간이 원하는 것보다 길어질 수 있다.

[0006] SRAM(Static Random Access Memory)은 DRAM의 대안이다. SRAM은 데이터를 저장하는데 커패시터를 사용하지 않으며 리프레시할 필요가 없다. SRAM은 DRAM보다 판독 및 기록 시간을 더 빠르게 할 수 있다. SRAM은 일반적으로 고속 레지스터 또는 캐시에 사용된다.

[0007] 컴퓨터-니어-메모리(compute-near-memory) 애플리케이션에서, 고밀도 CNM(compute-near-memory) 회로를 가능하게 하기 위해서 임베디드 DRAM(eDRAM) 메모리가 이용될 수 있다. 백엔드 eDRAM 디바이스가 백엔드 금속 층에 구현되는 경우에, CNM은 활성 실리콘 상의 백엔드 eDRAM 어레이 아래에 기회적으로(opportunistically) 배치된 컴퓨터 회로를 제공한다. 이러한 방식으로, CNM 회로는 eDRAM 어레이 바로 아래에 계산부(computation)를 위치시켜서 데이터 이동 시에 낭비되는 에너지 및 대기 시간을 최소화시키고, 중요한 계산에 사용 가능한 다이 상 메모리(on-die memory)의 대역폭을 최대화한다. eDRAM 기반 CNM 회로는 SRAM 기반 CNM 구현예에 비해서 더 높은 메모리 밀도(예를 들어, 단위 영역에 더 많은 데이터 저장)를 제공한다. 예를 들어 eDRAM 기반 CNM 회로는, 인공 지능(AI) 태스크, 머신 러닝, 딥 러닝, 그래프 처리, 암호 해독, 암호화, 압축 해제 또는 압축과 같은, 데이터 집약적 워크로드에, 높은 저장 용량 및 허용 가능한 컴퓨팅 효율성을 제공한다.

[0008] 이득 셀 임베디드 DRAM(eDRAM) 어레이는, 1 트랜지스터-1 커패시터(1T1C) 셀을 갖는 DRAM 또는 eDRAM보다 더 빠른 판독 및 기록 시간을 제공할 수 있다. 이득 셀 eDRAM은 경우에 따라 SRAM의 대안으로 간주된다. 일부 이득 셀 eDRAM은 여러 개의 트랜지스터(예를 들어, 5개 이하) 및 동적 저장 노드를 사용하는 비트 셀을 사용한다. 예를 들어, 2 트랜지스터(2T) 이득 셀 토폴로지는 기록 트랜지스터(MW) 및 결합형 스토리지 및 판독 트랜지스터(combined storage and read transistor)(MR)를 포함한다. 경우에 따라, 이득 셀 eDRAM 기술은 판독 트랜지스터(MR)를 사용해서 낮은 (row address to column address delay) 및 빠른 판독 시간을 제공할 수 있다. 이득

셀 eDRAM에 대한 설명은 P. Meinerzhagen 등의 "Gain-Cell Embedded DRAMs for Low-Power VLSI Systems-on-Chip"(2018)을 참조한다.

- [0009] 판독 시간의 한가지 척도는 tRCD/tRC의 비율이다. tRC 및 tRCD는 DRAM에서 두 개의 후속하는 행 액세스와 관련된 타이밍들을 나타낸다. tRC는 동일한 뱅크로의 연속 활성화 커맨드들 사이의 행 사이클 시간(Row Cycle Time) 및 최소 시간 간격을 나타낼 수 있다. tRCD는 행 주소 대 열 주소 지연을 나타낼 수 있으며, 이는 활성화 커맨드의 발행과 판독/기록 커맨드의 발행 사이에 걸리는 클럭 사이클 수이다. 어떤 경우에는, 이득 셀 eDRAM이 eDRAM 1T1C 기술보다 훨씬 낮은 tRCD를 제공한다.
- [0010] 도 1은 메모리 디바이스의 예시적인 하이 레벨 블록도를 도시한다. 메모리 디바이스(100)는 SRAM 메모리 어레이 IP(intellectual property)(106) 및 미드로직(midlogic)과 입/출력 접속을 가진 주변기기(108)를 포함할 수 있다. 메모리 어레이(106) 및 미드로직과 입/출력 접속부를 가진 주변기기(108)는 실리콘(Si) 기판(104) 상에 형성될 수 있다. 백엔드 상호 접속부(102)는 메모리 어레이(106) 및 미드로직 및 입/출력(I/O)(108)을 가진 주변기기 위에 형성될 수 있다. 반대로, 메모리 디바이스(100)에서 SRAM를 사용하는 것에 비해서, DRAM 기술을 사용하면 훨씬 더 높은 밀도를 제공한다(예를 들어, 영역당 더 많은 데이터를 저장). 그러나, 모든 판독에는 라이트 백(writeback) 동작이 있기 때문에, 판독 집약적인 애플리케이션의 경우 DRAM은 SRAM보다 상당히 느릴 수 있다. 따라서, DRAM은 느리고 긴 판독 시간으로 인해서 SRAM에 대한 매력적인 대안이 될 수는 없다. DRAM이 SRAM과 같은 고속 메모리와 결합할 수 있도록 하기 위해서, DRAM 메모리 셀의 COB(Capacitor Over Bitline) 충전 시간을 단축시킴으로써 판독 시간을 감소시킬 수 있다.
- [0011] 도 2는 일 실시예의 상위 레벨 개요 예를 도시한다. 메모리 디바이스(200)는 백엔드 상호 접속부 중 적층형 메모리의 이득 셀 eDRAM 메모리 어레이 기술을 포함할 수 있다. 메모리 어레이(202)는 기록 트랜지스터(MW)의 어레이 및 커패시터 스토리지를 포함할 수 있다. 판독 회로(204)는 커패시터 스토리지의 콘텐츠를 판독하는 데 사용되는 하나 이상의 트랜지스터를 포함할 수 있다. 커패시터 스토리지와 관련된 느린 판독 동작과 연관된 문제를 완화시키기 위해서, 이득 셀 eDRAM 구성은 고속 실리콘과 함께 사용될 수 있다. 메모리 어레이(202)와 관련된 판독 동작의 속도를 더 개선하기 위해, 층 전달 실리콘(layer transferred silicon)을 사용해서 더 빠른 판독 트랜지스터 기술을 제공하는 동시에, 메모리 어레이(202) 아래의 실리콘 기판(212)을 다른 용도로 자유롭게 사용할 수 있게 한다. 주변기기/미드로직/I/O(206)는 감지 증폭기(sense amplifier; SA)로의 비트라인 신호 라우팅, 워드라인 드라이버, 판독을 위해 판독 회로(204) 및 메모리 어레이(202)와의 접속에 사용되는 실리콘 트랜지스터, 기록 회로, 멀티플렉스 회로, 디코딩 회로 등을 포함할 수 있다.
- [0012] 일부 실시예에 따르면, 재료들은, 이들을 서로 접촉하는 인터페이스 층을 사용해서 본딩될 수 있다. 금속-금속 본딩은, 반대로 예칭된 금속들의 직접적인 접촉을 제공하도록 형성될 수 있다. 예를 들어, 영역들이 함께 접합될 수 있는데, 여기서 영역들은 서로 다른 기능(예를 들어, 기록 회로, 저장, 판독 회로, I/O 회로, 감지 증폭기 등)을 수행하거나, 서로 다른 재료를 사용해서 형성되거나, 서로 다른 온도에서 형성되는 것이다. 이하 더 자세히 설명하는 바와 같이, 디바이스의 2개 이상의 층들이 함께 본딩될 수 있다.
- [0013] DRAM 메모리 디바이스용 이득 셀에 대한 일부 구현예에서, 다수의 트랜지스터가 사용되며, 트랜지스터는 동일한 평면(예를 들어, X-Y 평면)에 배열된다. 동일한 평면에 트랜지스터를 배열하면 다이 공간을 사용할 수 있고, 다른 용도에 사용할 수 있는 다이 공간을 감소시킬 수 있다. 다양한 실시예는 본딩 프로세스 및 순차적인 3차원 채널 프로세싱을 사용해서(예를 들어, Z-방향을 따라), 판독 디바이스 상의 최상층에 형성된 기록 디바이스를 제공한다. 따라서 DRAM-레벨의 데이터 저장 밀도가, 기록 또는 판독 회로 이외의 다른 회로를 위한 X-Y 평면의 영역을 보존하면서도, SRAM 레벨 판독 속도로 제공될 수 있다.
- [0014] 도 3a는 일부 실시예에 따른 이득 셀 eDRAM 구조의 단면을 도시한다. 다양한 실시예는 기록 회로(300), 스토리지 커패시터(320), 유전체(322), 스토리지 커패시터(324), 및 판독 회로(330)를 포함하는 메모리 디바이스를 제공한다. Z-방향을 따라서, 기록 회로(300)는 판독 회로(330)의 평면과는 다른 평면에서 판독 회로(330)의 상부에 적어도 부분적으로 위치될 수 있다. 기록 회로(300)는 스토리지 커패시터(320)에 도전적으로 연결될 수 있고, 판독 회로(330)는 스토리지 커패시터(320)에 도전적으로 연결될 수 있다. 일부 실시예에서, 층(308)(기록 회로(300)과 콘택트)이 본 명세서에 설명된 방식으로 스토리지 커패시터(320)에 본딩될 수 있다.
- [0015] 일부 예에서, 기록 회로(300)는 금속 게이트(302), 게이트 산화물(304), 채널(306), 및 소스(S) 영역과 드레인(D) 영역을 포함할 수 있다. 금속 게이트(302)는 질화 티타늄(TiN), 텅스텐, 구리, 질화 탄탈(TaN), 루테튬(Ru), 산화 구리, 니켈, 탄소, 티타늄, 탄탈, 폴리-Si, 폴리-Ge, 이리듐, 산화 이리듐, 질화 하프늄 등 중 하나 이상을 포함할 수 있다. 게이트 산화물(304)은 산화 하프늄, 산화 실리콘, 산화 알루미늄, 질화 실리콘, 탄화

실리콘, h-BN, 탄소, 산화 란탄, 이들의 임의의 조합 또는 이들의 다중 층 중 하나 이상을 포함할 수 있다. 채널(306)은 폴리-Si/Ge/III-V/GaN, 단결정 층 전달 Si/Ge/III-V/GaN, IGZO(인듐 갈륨 아연 산화물), 산화 인듐, 산화 아연, WSe₂, WS₂, MoSe₂, 흑린, SnO, HfSnO, Cu₂O, CoO, IZO, AZO, 인듐 텅스텐 산화물, 인듐 주석 산화물(ITO) 또는 이들의 조합 또는 SiO₂, HfO₂, Al₂O₃ 또는 기타 전기 절연체로 도핑된 것을 포함할 수 있다.

[0016] 소스 영역 및 드레인 영역은 각각의 트랜지스터의 게이트 스택에 인접해서 기판 내에 형성될 수 있다. 소스 영역 및 드레인 영역은 주입 또는 확산 프로세스 또는 에칭 또는 증착 프로세스를 사용해서 형성될 수 있다. 주입 또는 확산을 위해서, 붕소, 알루미늄, 안티몬, 인 또는 비소와 같은 도펀트가 기판에 이온 주입되어서 소스 영역 및 드레인 영역을 형성할 수 있다. 이온 주입 프로세스에 후속해서, 일반적으로, 도펀트를 활성화하고 이들이 기판으로 더 확산되게 하는 어닐링 프로세스가 행해진다. 에칭 또는 증착 프로세스에서, 기판은 소스 영역 및 드레인 영역의 위치에 리세스를 형성하도록 먼저 에칭될 수 있다. 이후, 에피택셜 증착 프로세스가 수행되어서, 이 리세스를, 소스 영역 및 드레인 영역을 제조하는 데 사용되는 재료로 충전할 수 있다. 일부 구현예에서, 소스 영역 및 드레인 영역은 실리콘 게르마늄 또는 탄화 실리콘과 같은 실리콘 합금을 사용해서 제조될 수 있다. 일부 구현예에서, 에피택셜 증착된 실리콘 합금은 붕소, 비소 또는 인과 같은 도펀트로 인시츄 도핑될 수 있다. 추가 실시예에서, 소스 영역 및 드레인 영역은 게르마늄 또는 III-V족 재료 또는 합금과 같은 하나 이상의 대안의 반도체 재료를 사용해서 형성될 수도 있다. 그리고 추가 실시예에서, 금속 및/또는 금속 합금의 하나 이상의 층이 소스 영역 및 드레인 영역을 형성하는데 사용될 수 있다.

[0017] 일부 예에서, 금속 게이트(302)가 게이트 산화물(304)을 완전히 덮을 수도 있고, X-Y 평면을 따라 게이트 산화물(304)의 일부를 덮을 수도 있다. 일부 예에서, 게이트 산화물(304)은 X-Y 평면을 따라 채널(306)의 일부를 덮을 수도 있다. 소스 영역(S)은 채널(306)에 콘택트하도록 형성될 수 있다. 기록 비트 라인(WBL)(도시되지 않음)이 소스 영역(S)에 연결될 수 있다. 드레인 영역(D)은 층(308)을 거쳐 채널 내에 형성될 수 있다. 층(308)은 질화 실리콘(SiN) 에칭 정지, 산화물 층(예를 들어, SiO₂, Al₂O₃), 질화물 층, 또는 탄화물 층 중 하나 이상을 포함할 수 있다. 드레인 영역(D)이 채널(306) 및 스토리지 커패시터(320)에 연결될 수 있다. 기록 회로(300)의 다양한 예시적인 구성이 예를 들어, 도 7a와 관련해서 본 명세서에 설명된다.

[0018] 기록 회로(300) 및 층(308)은, 에칭, 화학 기상 증착(CVD), 폴리싱, 도핑, 전자 주입, 이온 주입, 에피택셜 성장 등 중 하나 이상을 포함하는 임의의 반도체 프로세싱 기술을 사용해서 영역 내에 형성될 수 있다. 영역(307)은 산화물, 필드 산화물, 질화물, 질화 실리콘과 같은 임의의 타입의 절연 재료 또는 반도체 재료일 수 있지만, 이것으로 한정되는 것은 아니다.

[0019] 일부 실시예에 따르면, 본딩 인터페이스(338)를 사용해서 제 1 영역(350)의 층(308)의 바닥면을 영역(326)을 포함하는 제 2 영역(352)의 상부면에 본딩하고, 드레인 영역(D)을 스토리지 커패시터(320)에 도전적으로 연결할 수 있다. 일부 예에서, 본딩 인터페이스(338)는 이산화 실리콘(SiO₂), 산질화 실리콘(SiON), 탄소 도핑된-산질화 실리콘(SiOCN), 실리콘 탄소 질화물(SiCN), 또는 산탄화 실리콘(SiOC) 중 하나 이상을 포함할 수 있다. 표면에 본딩 재료를 도포하는 것은, PVD(Physical Vapor Deposition), CVD(Chemical Vapor Deposition), PECVD(Plasma-enhanced Chemical Vapor Deposition), ALD(Atomic Layer Deposition), PEALD(Plasma Enhanced Atomic Layer Deposition), 스퍼온, 습식 증착 공정 또는 기타 유사한 공정에 의해 수행될 수 있다.

[0020] 제 1 영역(350)의 바닥 표면을 제 2 영역(352)의 상부 표면에 본딩하기 위해서, 본딩 인터페이스(338)가 제 1 영역(350)의 바닥 표면에 도포되고, 제 2 영역(352)의 상부 표면에 도포될 수 있다. 이 본드는 표면을 플라즈마 처리(예를 들어, 질소, 아르곤, 산소, NH₃, Cl, F₂ 등의 고에너지 이온의 노출)함으로써 활성화되어서, 브로큰 본드(broken bond) 또는 땀글링 본드(dangling bond)를 생성한다. 이러한 땀글링 본드는 제 1 영역(350)을 형성하는 데 사용되는 웨이퍼와 제 2 영역(352)을 형성하는 데 사용되는 웨이퍼의 접착성을 향상시킬 수 있다. 웨이퍼는 정렬 마스크를 사용해서 서로 정렬되고, 예를 들어 실온 내지 400°C의 범위의 온도에서 서로 가압된다.

[0021] 스토리지 커패시터 및 유전체(예를 들어, 320, 322, 및 324)는 질화 티타늄, 질화 탄탈륨, W, Ru, Ir, Cu, Ta, Ti, 또는 이들의 조합이나 이들의 다층 구조와 같은 내부 전극 재료(320)를 포함할 수 있다. 유전체(322)는 산화 실리콘, 산화 티타늄, 산화 지르코늄, 산화 하프늄, 산화 하프늄 지르코늄, 산화 알루미늄, STO(예를 들어, SrTiO₃) 재료, BTO(예를 들어, BaTiO₃) 재료, 압전(PZT) 재료를 포함할 수 있다. 유전체(322)는, 내부 전극 재료(320) 주위에 형성되면서(예를 들어, 실린더형 내부 전극 재료(320)의 표면을 덮음) 내부 전극 재료(320)와 외부 전극(324) 사이에 위치한 플레이트를 포함할 수 있다.

[0022] 스토리지 커패시터(320) 아래에 (-Z) 방향으로 판독 회로(330)가 형성될 수 있다. 일부 예에서, 판독 회로(330)는 산화 금속(332), 게이트 산화물(334), 및 채널(336)을 포함할 수 있다. 금속 게이트(332)는 질화 티타늄

(TiN), 텅스텐, 구리, 질화 탄탈(TaN), 루테튬(Ru), 산화 구리, 니켈, 탄소, 티타늄, 탄탈, 폴리-Si, 폴리-Ge, 이리듐, 산화 이리듐, 질화 하프늄 등 중 하나 이상을 포함할 수 있다. 게이트 산화물(334)은 산화 하프늄, 산화 실리콘, 산화 알루미늄, 질화 실리콘, 탄화 실리콘, h-BN, 탄소, 산화 탄탄, 이들의 임의의 조합 또는 이들의 다층 중 하나 이상을 포함할 수 있다. 채널(336)은 폴리-Si/Ge/III-V/GaN, 단결정 층 전달 Si/Ge/III-V/GaN, IGZO(인듐 갈륨 아연 산화물), 산화 인듐, 산화 아연, WSe₂, WS₂, MoSe₂, 흑린, SnO, HfSnO, Cu₂O, CoO, IZO, AZO, 인듐 텅스텐 산화물, 인듐 주석 산화물(ITO) 또는 이들의 조합 또는 SiO₂, HfO₂, Al₂O₃ 또는 기타 전기 절연체로 도핑된 것을 포함할 수 있다.

[0023] 일부 예에서, 금속 게이트(322)는 게이트 산화물(334)을 완전히 덮을 수도 있고, X-Y 평면을 따라 게이트 산화물(334)의 일부를 덮을 수도 있다. 일부 예에서, 게이트 산화물(334)은 X-Y 평면을 따라 채널(336)의 일부를 덮을 수도 있다. 소스 영역(S)은 게이트 산화물(334) 및 채널(306)에 컨택트하도록 형성될 수 있다. 소스 영역(S)은 판독 비트 라인(RBL)(도시되지 않음)으로부터 신호를 수신하도록 연결될 수 있다. 드레인 영역(D)은 채널(336)과 컨택트하고 게이트 산화물(334)과 컨택트하도록 형성될 수 있다. 드레인 영역(D)은 판독 워드 라인(RWL)(도시되지 않음)으로부터 신호를 수신하도록 연결될 수 있다.

[0024] 스토리지 커패시터(320), 유전체(322), 스토리지 커패시터(324) 및 판독 회로(330)는 에칭, 화학 기상 증착(CVD), 폴리싱, 도핑, 전자 주입, 이온 주입, 에피택셜 성장 등 중 하나 이상을 포함하는 임의의 반도체 프로세싱 기술을 사용해서 영역(326) 내에 형성될 수 있다. 영역(326)은 산화물, 필드 산화물, 질화물, 질화 실리콘, 산화 알루미늄과 같은 임의의 타입의 절연 재료 또는 반도체 재료일 수 있지만, 이것으로 한정되는 것은 아니다.

[0025] 일부 실시예에서, 제 1 영역(350)은 기록 회로(300), 층(308), 및 비한정의 예로서 100°C 내지 2000°C와 같은 제 1 온도 또는 온도 범위에서 형성된 영역(307)을 포함할 수 있다. 제 2 영역(352)은 스토리지 커패시터(320), 유전체(322), 스토리지 커패시터(324), 판독 회로(330), 및 비한정의 예로서 100°C 내지 2000°C와 같은 제 2 온도 또는 온도 범위에서 형성된 영역(326)을 포함할 수 있다. 제 2 온도 또는 온도 범위는 제 1 온도 또는 온도 범위와 상이할 수도 있고 또는 적어도 부분적으로 중첩될 수도 있다. 따라서, 제 1 영역(350) 및 제 2 영역(352)은 서로 다른 온도에서 생성될 수도 있고 본딩 인터페이스(338)를 사용해서 함께 본딩될 수 있다.

[0026] 도 3b는 일부 실시예에 따른 이득 셀 eDRAM 구조의 예시적인 단면을 도시한다. 일부 실시예에서, 제 1 영역(360)은 제 1 온도 또는 온도 범위에서 기록 회로(300), 층(308), 및 스토리지 커패시터(320), 유전체(322), 스토리지 커패시터(324), 판독 회로(330), 및 영역(326)과 함께 형성된 영역(307)을 포함할 수 있다. 이 예에서, 도 3a의 예에서와 같이, 층(308) 및 드레인(D) 영역을 영역(326) 및 스토리지 커패시터(320)의 일부에 본딩하는 대신, 층(308) 및 드레인 영역이 영역(326) 및 스토리지 커패시터(320) 위에 형성된다.

[0027] 제 2 영역(362)은 제 2 온도 또는 온도 범위에서 영역(340) 내에 형성된 판독 회로(330)를 포함할 수 있다. 제 2 온도 또는 온도 범위는 제 1 온도 또는 온도 범위와 상이할 수 있고 또는 적어도 부분적으로 중첩될 수 있다. 판독 회로(330)는 에칭, 화학 기상 증착(CVD), 폴리싱, 도핑, 전자 주입, 이온 주입, 에피택셜 성장 등 중 하나 이상을 포함하는 임의의 반도체 처리 기술을 사용해서 영역(340) 내에 형성될 수 있다. 영역(340)은 산화물, 필드 산화물, 질화물, 질화 실리콘, 산화 알루미늄과 같은 임의의 타입의 절연 또는 반도체 재료일 수 있지만, 이것으로 한정되는 것은 아니다.

[0028] 일부 실시예에서, 제 1 영역(360) 및 제 2 영역(362)은 상이한 온도 또는 중첩되는 온도에서 생성될 수 있고, 본딩 인터페이스(342)를 사용하여 함께 본딩될 수 있다. 이 예에서, 본딩 인터페이스(342)를 사용하여, 판독 회로(330)는 영역(326)에 본딩될 수 있고, 금속 게이트(332)는 스토리지(320)에 도전적으로 연결될 수 있다. 이 예에서, 본딩 인터페이스(342)를 형성하기 위해, 제 1 영역(360)의 하부는 본딩 인터페이스(342)의 재료로 덮여질 수 있고, 대향하는 제 2 영역(362)의 부분은 본딩 인터페이스(342)의 재료로 덮여질 수 있다. 본딩 인터페이스(342)는 이산화 실리콘(SiO₂), 산질화 실리콘(SiON), 탄소 도핑된 실리콘 산질화물(SiOCN), 실리콘 탄소 질화물(SiCN), 또는 산탄화 실리콘(SiOC) 중 하나 이상을 포함할 수 있다. 제 1 영역(360)을 제 2 영역(362)에 본딩하기 위해, 본딩 인터페이스(338)와 관련하여 앞서 설명된 기술이 사용될 수 있다.

[0029] 일부 경우에, 스토리지 커패시터 및 유전체(예를 들어, 320, 322, 및 324)를 형성하는 것은, 기록 회로(300) 또는 판독 회로(330)를 제조하는 데 사용되는 온도보다 높은 온도 혹은 이와는 다른 온도에서 이루어지며, 이로써 스토리지 커패시터 및 유전체(예를 들어, 320, 322, 324)를 제조하는 것은, 기록 회로(300) 또는 판독 회로(330)를 제조하는 것과 호환되지 않는다는 점에 주의한다. 본딩은 기록 회로(300) 또는 판독 회로(330)를 스토리지 커패시터 및 유전체(예를 들어, 320, 322, 324)와 결합하는 방식을 제공한다.

- [0030] 도 3c는 일부 실시예에 따른 이득 셀 eDRAM 구조의 예시적인 단면을 도시한다. 이 예에서, 제 1 영역(350)은 본딩 인터페이스(338)를 사용해서 제 2 영역(352)에 본딩된다. 제 2 영역(352)은 본딩 인터페이스(370)를 사용해서 회로(354)에 본딩된다. 이 예에서, 본딩 인터페이스(370)를 형성하기 위해서, 제 2 영역(352)의 하부는 본딩 인터페이스(370)의 재료로 덮여지고, 대향하는 회로(354)의 부분은 본딩 인터페이스(370)의 재료로 덮여질 수 있다. 본딩 인터페이스(338)는 이산화 실리콘(SiO₂), 산질화 실리콘(SiON), 탄소 도핑된-산질화 실리콘(SiOCN), 실리콘 탄소 질화물(SiCN), 또는 산탄화 실리콘(SiOC) 중 하나 이상을 포함할 수 있다. 제 1 영역(350)을 제 2 영역(352)에 본딩하기 위해서, 본딩 인터페이스(370)와 관련하여 앞서 설명된 기술이 사용될 수 있다. 회로(354)는 여러 회로들 중에서 특히, 감지 회로, CNM(compute near memory) 회로, SRAM 에플리케이션 회로(본 명세서에 설명됨), 비아, 비트 라인 또는 워드 라인 드라이버용 V_{cc}, V_{ss} 단자를 포함할 수 있다.
- [0031] 도 3d는 일부 실시예에 따른 이득 셀 eDRAM 구조의 예시적인 단면을 도시한다. 이 예에서, 제 1 영역(360)은 본딩 인터페이스(342)를 사용해서 제 2 영역(362)에 본딩될 수 있다. 제 2 영역(362)은 또한 상기 설명한 것과 유사한 방식으로 본딩 인터페이스(370)를 사용하여 회로(354)에 본딩될 수 있다. 따라서 본딩 인터페이스는 모든 구조를 결합하고 전도성 연결을 제공하는데 사용될 수 있다.
- [0032] 도 4는 공유 층을 가진 병렬식(side-by-side) 이득 셀 디바이스의 예를 도시한다. 공유 층은 층(308)일 수 있다. 상기 설명한 바와 같이, 상이한 구조들 사이의 본딩이 수행될 수 있다. 예를 들어, 본딩 인터페이스(402, 404, 및/또는 406) 중 하나 이상을 사용해서, 필요한 경우 전도성 결합을 가능하게 하면서 서로 다른 구조를 함께 본딩할 수 있다. 본딩 인터페이스(402, 404, 및/또는 406)는 이산화 실리콘(SiO₂), 산질화 실리콘(SiON), 탄소 도핑된 실리콘 산질화물(SiOCN), 실리콘 탄소 질화물(SiCN), 또는 산탄화 실리콘(SiOC) 중 하나 이상을 포함할 수 있다. 본딩 인터페이스(338)와 관련하여 앞서 설명된 기술을 사용하여 서로 다른 영역을 본딩할 수 있다.
- [0033] 도 5는 도 4의 구조의 평면도의 예를 도시한다. 뷰(502)는 일부 실시예에 따른 기록 회로의 평면도를 제공한다. 셀(504)은 일부 실시예에 따른 기록 회로를 포함할 수 있다. 일부 예에서, 셀(504)의 정면 뷰(502)는 셀(504) 위에 형성된 비트 라인(BL)을 도시하고, 비트 라인(BL)은 셀(504)의 적어도 하나의 다른 카피와 콘택트한다. 나아가, 워드 라인(WL)이 셀(504)의 게이트 산화물에 연결되어서, 셀(504)의 적어도 하나의 다른 카피와 콘택트할 수 있다. 용량성 콘택트(Cap contact)는 용량성 스토리지(도시되지 않음)와의 연결을 위해 셀(504) 아래에 위치될 수 있다.
- [0034] 도 6은 일부 실시예에 따른 이득 셀 eDRAM 구조의 단면을 도시한다. 다양한 실시예는 기록 회로(600), 스토리지 커패시터(320), 유전체(322), 스토리지 커패시터(324), 및 관독 회로(330)를 포함하는 메모리 디바이스를 제공한다. 기록 회로(600)는 적어도 부분적으로 Z 방향으로 관독 회로(330) 상부에 위치될 수 있고, X-Y 평면에서는 관독 회로(330)와는 다른 면에 위치될 수 있다. 기록 회로(600)는 스토리지 커패시터(320)에 전도성 연결될 수 있고 관독 회로(330)는 스토리지 커패시터(320)에 도전적으로 연결될 수 있다.
- [0035] 기록 회로(600)는 금속 게이트(602), 게이트 산화물(604), 채널(606), 및 층(608)을 포함할 수 있다. 금속 게이트(602)는 질화 티타늄(TiN), 텅스텐, 구리, 질화 탄탈(TaN), 루테튬(Ru), 산화 구리, 니켈, 탄소, 티타늄, 탄탈륨, 폴리-Si, 폴리-Ge, 이리듐, 산화 이리듐, 질화 하프늄 등 중 하나 이상을 포함할 수 있다. 게이트 산화물(604)은 산화 하프늄, 산화 실리콘, 산화 알루미늄, 질화 실리콘, 탄화 실리콘, h-BN, 탄소, 산화 란탄, 이들의 임의의 조합 또는 이들의 다중 층 중 하나 이상을 포함할 수 있다. 채널(606)은 폴리-Si/Ge/III-V/GaN, 단결정 층 전달 Si/Ge/III-V/GaN, IGZO(인듐 갈륨 아연 산화물), 산화 인듐, 산화 아연, WSe₂, WS₂, MoSe₂, 흑린, SnO, HfSnO, Cu₂O, CoO, IZO, AZO, 인듐 텅스텐 산화물, 인듐 주석 산화물(ITO) 또는 이들의 조합 또는 SiO₂, HfO₂, Al₂O₃ 또는 기타 전기 절연체로 도핑된 것을 포함할 수 있다. 층(608)은 실리콘 아인슈타이늄(SiN Es)을 포함하도록 형성될 수 있다.
- [0036] 일부 예에서, 금속 게이트(602)는 게이트 산화물(604)을 완전히 덮거나 X-Y 평면을 따라 게이트 산화물(604)의 일부를 덮을 수 있다. 게이트 산화물(604)은 채널(606)의 일부를 덮을 수 있다. 소스 영역(S)은 채널(606)과 콘택트하도록 형성될 수 있다. 기록 비트 라인(WBL) 신호는 소스 영역(S)(도시되지 않음)에 연결될 수 있다. 드레인 영역(D)은 채널(606)과 콘택트하도록 형성될 수 있다. 접속부(610)는 드레인 영역(D)을 스토리지 커패시터(320)에 연결시킬 수 있다. 일부 예에서, 접속부(610)는 Cu, Ru, TiN, W, 또는 Co 중 하나 이상으로 형성될 수 있다. 접속부(610)는 층(608)의 일부를 통해 형성될 수 있다. 이 예에서, 접속부(610)는 역 L 형상 구조이지만, 임의의 형상이 사용될 수도 있다.
- [0037] 이 예에서, 본딩 인터페이스(618)와 본딩 인터페이스(620) 중 하나 이상이 도 3a 내지 도 3d의 구조와 관련하여 설명된 것과 유사한 방식으로 형성되어서 사용될 수 있다. 본딩 인터페이스(618 및/또는 620)는 이산화 실리콘

(SiO₂), 산질화 실리콘(SiON), 탄소 도핑된 산질화 실리콘(SiOCN), 질화 탄소(SiCN) 또는 산탄화 실리콘(SiOC) 중 하나 이상을 포함할 수 있다.

[0038] 도 7a는 상부 및 하부 게이트 구조의 다양한 예의 단면도를 도시한다. 이 예에서, 구조(700, 710, 720, 730)가 도시되어 있고, 이는 기록 회로 또는 판독 회로에서 사용될 수 있다. 일부 예에서, 상부 게이트 상부 컨택트의 경우, 박막 트랜지스터 구성이 측벽 게이트 없이 사용될 수 있다. 일부 예에서, 게이트 구조에 대해 적어도 FinFET 구성이 사용될 수 있다. 기록(상단) 트랜지스터 및 판독(하단) 트랜지스터를 형성하는 데 사용되는 재료는 폴리-Si, Si, Ge, 폴리-Ge, III-V, GaN, TMD(transition metal dichalcogenide)(예를 들어, MoS₂, WSe₂, MoSe₂, WSe₂), InS, HfS, ZnS, ZnSe, In₂O₃, ZnO, AZO, IGZO, IZO 및 다른 백엔드 호환 재료 중 하나 이상을 포함할 수 있다. 백엔드는, 벌크 실리콘 트랜지스터의 형성에 후속하는 처리 층을 포함할 수 있다. 백엔드 호환 재료는 금속 및 유전체를 포함할 수 있으며 온도 버젓(temperature budget)은 약 500°C 이하이다. 다양한 실시 예는 본 명세서에 설명된 백엔드 호환 가능한 트랜지스터를 제공한다.

[0039] 도 7b는 일부 실시예의 상부 게이트 상부 컨택트 구조의 단면의 예를 도시한다. 이 예에서, 구조(710A)는 구조(710)의 단면도이다. FinFET(fin field-effect transistor) 구현이 게이트 구조에 사용되고 측벽 게이트가 존재한다. 일부 실시예에 따르면, 핀(fin)(예를 들어, 좁은 채널)은 평면 구현에 비해 향상된 게이트 제어, 향상된 하위 임계값 스윙(subthreshold swing) 및 향상된 누설을 제공한다. 향상된 게이트 제어를 통해서, 음수 또는 0볼트로 낮은 I_{off}(오프 상태 누설)를 유지하면서, 주어진 전압에서 구동 전류(성능)를 향상시킬 수 있다. 판독 및/또는 기록 트랜지스터는 급격하게(sharply) 턴오프될 수 있어서 하위 임계값 스윙(예를 들어, 게이트 전압을 감소시키면서 전류가 감소하는 비율)이 급격하거나 혹은 낮을 수 있다.

[0040] 도 7c는 상부 게이트 교번 컨택트의 예를 도시하며, 게이트는 커패시터 컨택트와 중첩된다. 구조(730)는 Z 방향으로 게이트와 컨택트가 중첩하는 것을 포함하고, 비트라인 옆 게이트와는 중첩할 수도 있고 중첩하지 않을 수도 있다. Z 방향에서의 게이트와 컨택트 사이에 중첩은, 게이트 아래의 채널에 추가 전하를 생성하고(게이트 필드로 인해), 이는 정전 도핑을 유도한다. 이 정전 도핑은 컨택트 저항을 향상시켜서 구동 전류(성능)를 향상시킨다. 드라이버 전류 성능을 개선하면, 스토리지 커패시터를 충전하는 시간을 감소시키고 대기 시간(예를 들어, 판독/기록 시간)을 감소시킬 수 있다. 게이트가 비트라인 컨택트 옆의 컨택트와 중첩되면, 게이트와 컨택트 사이에 형성된 커패시터로서 인해서, 스토리지 커패시터에 저장된 0 또는 1 값을 식별할 때의 신호 대 잡음비가 악화된다. 그러나 커패시터가 커패시터 컨택트면 상에 대신 형성되기 때문에, 신호 대 잡음비가 악영향을 받지 않을 수도 있으며, 커패시터 스토리지의 데이터 보존이 향상될 수 있다.

[0041] 도 8은 다중 이득 셀 eDRAM 디바이스를 갖는 메모리 디바이스의 단면도를 도시한다. 다양한 실시예에 따르면, 메모리 셀(802)은, 백엔드 기록 회로(804)(MW)가 판독 회로(806)가 있는 X-Y 평면에 있지 않고 Z 방향으로 판독 회로(806)(MR)의 상부에 위치되어 있는 다중 비트 셀을 포함한다. 일부 예에서, Z 방향을 따라서, 백엔드 기록 회로(804)는 스토리지 셀(808)에 접속되고 판독 회로(806)는 스토리지 셀에 접속된다. 이 예에서, 기록 회로(804)는, 게이트가 채널 주위를 둘러싸고 있는 FinFET 트랜지스터로서 구현된다. 기록 회로(804)는 스토리지 셀(808)에 비트를 기록할 수 있다. 이 예에서, 스토리지 셀(808)은 COB(Capacitor Over Bitline) 디바이스로서 구현될 수 있다. 판독 회로(806)는 스토리지 셀(808)에 저장된 하나 이상의 비트 및 스토리지 셀(808)의 CSN 상태를 감지하는 데 사용되는 MR 트랜지스터를 사전 증폭할 수 있다. MR 트랜지스터의 어레이는 저장된 C_{SN} 상태를 판독할 수 있다.

[0042] 본 명세서에서 더 상세히 설명되는 바와 같이, 복수의 디바이스 층은 서로의 상부에 본딩 및 적층될 수 있다. 예를 들어, 하나의 층은 N형 MOS(metal-oxide-semiconductor)(NMOS) 트랜지스터만 포함할 수도 있고, P형 MOS 트랜지스터(PMOS) 트랜지스터만 포함할 수도 있으며, 위 또는 아래의 다른 층은 CMOS(complementary metal-oxide-semiconductor)(NMOS 트랜지스터 및 PMOS를 가짐)를 포함할 수 있다.

[0043] 도 9는 이득 셀 트랜지스터의 일부 실시예의 회로 표현을 도시한다. 이하, 적어도 도 3a 내지 3d, 4, 6 및 8의 디바이스에 관한 예시적인 판독 및 기록 동작을 설명한다. 판독 트랜지스터(MR) 및 기록 트랜지스터(MW)는 메모리 회로의 커패시터(C_{SN})와 상호 작용한다. 예를 들어, 모든 NMOS 트랜지스터 판독 트랜지스터(MR) 및 기록 트랜지스터(MW) 구현예의 경우, DRAM 비트 셀은 MW 트랜지스터 및 커패시터 C_{SN}를 포함한다. 기록 동작의 경우 WBL을 켜면 커패시터(C_{SN})에 저장된 전하가 판독 트랜지스터(MR)의 게이트에 연결된 노드(SN)로 전달된다. 논리 1을 기록하기 위해서, 판독 트랜지스터(MR)가 활성화된다. 논리 0을 기록하기 위해서, 판독 트랜지스터(MR)은 개방된다(open circuit).

- [0044] 커패시터(C_{SN})로부터 데이터를 판독하기 위해서, WBL은 오프로 유지되고, 전하가 커패시터(C_{SN})로부터 라인(WBL)으로 방전되지만, RC 시정수 지연이 발생한다. 커패시터(C_{SN})에 1이 저장되면 MR의 게이트는 전압을 가지며, 채널 저항은 낮다. 커패시터(C_{SN})에 0이 저장되면 저항이 높다. MR의 게이트가 특정 전위(0 또는 1)에 있기 때문에, RWL과 RBL 사이의 저항 변화가 발생한다. RWL과 RBL 사이에서 일정한 전압 차를 적용하면 커패시터(C_{SN})의 콘텐츠를 판독할 수 있다. 커패시터(C_{SN})가 1을 저장하면, 판독 트랜지스터를 지나는 전류가 높다. 커패시터(C_{SN})가 0을 저장하면, 판독 트랜지스터가 NMOS 디바이스인 경우 판독 트랜지스터를 지나는 전류 레벨은 낮으며, 판독 트랜지스터가 PMOS 디바이스인 경우에는 순서가 반대이다.
- [0045] 그러나, X-Y 평면에서 MR 트랜지스터가 MW 트랜지스터 옆에 있는 경우, SN 노드를 판독 트랜지스터(MR)의 게이트에 접속하면, 면적을 소모한다. 기록 회로가 백엔드 트랜지스터인 경우, 기록 회로는 프론트 엔드에 있는 판독 회로에 더해서 추가 X-Y 평면 공간을 소비하지 않는다. 백엔드 트랜지스터는 아래에 콘택트를 가질 수도 있다(게이트와 콘택트가 같은 평면에 있지 않음). 본 명세서에 기술된 바와 같이, 다양한 실시예는 MR 트랜지스터 위의 평면에 MW 트랜지스터를 제공한다.
- [0046] 도 10a 내지 도 10f는 메모리 디바이스를 형성하는 예를 도시한다. 예를 들어, 메모리 디바이스는 이득 셀 eDRAM 디바이스일 수 있다. 도 10a는 실리콘 기판(1010)의 단면도를 도시한다. 실시예에서, 기판(1010)은 실리콘 기판, 소다 석회 유리 또는 붕규산염 유리와 같은 유리 기판, 금속 기판, 플라스틱 기판, SiO₂를 포함하는 기판, 또는 다른 적절한 기판일 수 있다. 구조(1020)는 실리콘 기판(1010) 내부에 또는 위에 형성될 수 있다. 구조(1020)는 판독(MR) 트랜지스터, 비아, 재분배, 금속 라우팅, 워드 라인(WL) 드라이버, 전원 공급 디바이스, 메모리 컨트롤러, 메모리 관리 유닛, 행 디코더 및 드라이버, 논리 및 처리 회로를 포함할 수 있다. 구조(1020)는 다양한 재료 증착, 에칭, 폴리싱 등에 의해 실리콘 기판(1010) 상에 또는 실리콘 기판(1010) 내에 형성될 수 있다. 구조(1020)는 평면 또는 FinFET 기술을 사용하는 임의의 타입의 능동 회로를 포함할 수 있다.
- [0047] 도 10b는 구조(1020)의 층간 유전체(ILD) 증착의 예를 도시한다. ILD 층은, 저-k 유전체 재료와 같은, 집적 회로 구조에서의 적용 가능성으로 알려진 유전체 재료를 사용해서 형성될 수 있다. 사용될 수 있는 유전 재료의 예는 이산화 실리콘(SiO₂), CDO(carbon doped oxide), 질화 실리콘, 퍼플루오로사이클로부탄 또는 폴리테트라플루오로에틸렌과 같은 유기 중합체, 플루오로실리케이트 유리(FSG), 및 실세스퀴옥산, 실록산 또는 유기실리케이트 유리를 포함하지만, 이것으로 한정되는 것은 아니다. ILD 층은 유전 상수를 추가로 감소시키기 위해 기공(pores) 또는 에어 갭을 포함할 수 있다. ILD 층은 산화 실리콘(SiO) 막, 질화 실리콘(SiN) 막, 03-테트라에틸 오르토실리케이트(TEOS), 03-헥사메틸디실록산(HMDS), 플라즈마-TEOS 산화물 층, 또는 다른 적절한 재료를 포함할 수 있다.
- [0048] 도 10c는 트랜지스터의 제 2 층의 형성의 예를 도시한다. 트랜지스터의 제 2 층은 구조(1020) 내에 형성될 수 있다. 구조(1020)는 금속 게이트(302), 게이트 산화물(304), 채널(306), 소스(S) 및 드레인(D) 영역과 관련하여 위에 설명한 채널 재료로 형성된 기록(MW) 회로(1032)와 같은 비 실리콘 기판(non-silicon-based) 트랜지스터를 포함하도록 수정될 수 있다. 기록 회로(1032)는 앞서 설명한 바와 같이, 스토리지 셀과 콘택트해서 상부에 형성될 수 있다. 구조(1030)에는 신호 또는 전압 전달 비아(1033)도 또한 형성될 수 있다. 순차적 3차원 채널 프로세싱은 구조(1030) 내에(예를 들어, 저온에서) 트랜지스터 및 비아를 증착하거나 형성하는데 사용될 수 있다.
- [0049] 일부 예에서, 본딩 층(1034)은 구조(1036)를 구조(1030)에 부착하는 데 사용될 수 있다. 본딩 프로세스의 다양한 예가 본 명세서에서 설명된다. 구조(1036)는 워드라인(WL) 스트랩 및 WL 드라이버로의 비아를 포함할 수 있다.
- [0050] 도 10d는, 구조(1036)의 표면(1040)에 제 2 기판(1050)의 표면(1052)을 본딩하는 것의 예를 도시한다. 일부 예에서, 표면(1040)은 본딩 재료로 덮여질 수 있고 표면(1052)은 본딩 재료로 덮여질 수 있다. 표면(1040)을 표면(1052)에 본딩하는 것은 양면(1040, 1052)을 덮는 본딩 재료의 콘택트에 의해, 이루어질 수 있다. 일부 예에서, 본딩 재료는 SiO₂, SiON, SiOCN, SiCN, 또는 SiOC 중 하나 이상을 포함할 수 있다. 본딩 표면의 다양한 예가 본 명세서에 설명되어 있다. 예를 들어, 제 2 기판(1050)은 캐리어 웨이퍼일 수 있고, 제 2 기판(1050)을 표면(1040)에 본딩한 후에, 최종 구조는 뒤집혀서 패터닝될 수 있다.
- [0051] 도 10e는 기판(1010)을 폴리싱해서 표면(1060)을 노출시키는, 특히 후면으로부터 판독 트랜지스터의 활성 핀(예를 들어, 채널 영역)을 노출시키는 구조를 나타낸다. 에칭 또는 폴리싱 또는 그라인딩과 같은, 기판(1010)을 제거하기 위한 다른 기술이 사용될 수도 있다.

- [0052] 도 10f는 구조(1070) 내에서 순차 처리를 사용해서 (본딩하는 대신에) 층들을 나란히 생성함으로써 (예를 들어, 에칭 산화물 및 금속을 통해) 후면 금속을 처리함으로써 생성되는 구조를 도시한다. 구조(1070)를 형성하는 것은 웨이퍼를 뒤집어서 수행될 수 있고, 이 구조(1070)는 상부에 노출된다. 그러나, 일부 예에서, 구조(1070)는 본 명세서에 설명된 기술에 따라 표면(1060)(도 10e)에 본딩될 수 있다.
- [0053] 도 11은 이득 셀 메모리 디바이스를 형성하기 위한 프로세스를 도시한다. 1102에서, 제 1 구조가 형성될 수 있다. 예를 들어, 제 1 구조는 동일한 타입(예를 들어, PMOS, NMOS)의 트랜지스터 또는 CMOS 트랜지스터 등을 포함할 수 있다. 제 1 구조는 특정한 동작을 제공할 수 있다. 예를 들어, 제 1 구조는 기록 회로, 판독 회로, 저장 용량, 감지 회로 및 주변 로직, SRAM 에플리케이션 회로 등 중 하나 이상을 사용해서 동작을 수행할 수 있다. 제 1 구조는 특정 그룹의 재료를 사용해서 특정 제 1 온도 또는 온도 범위에서 형성될 수 있다.
- [0054] 1104에서, 제 2 구조가 형성될 수 있다. 예를 들어, 제 2 구조는 동일한 타입(예를 들어, PMOS, NMOS)의 트랜지스터 또는 CMOS 트랜지스터 등을 포함할 수 있다. 제 2 구조는 제 1 구조와 다른 동작을 제공할 수 있다. 예를 들어, 제 2 구조는 기록 회로, 판독 회로, 저장 커패시턴스, 감지 회로 및 주변 로직, SRAM 에플리케이션 회로 등 중 하나 이상을 사용해서 동작을 수행할 수 있다. 제 2 구조는 특정한 재료 그룹을 사용하여 형성될 수 있고, 특정한 제 2 온도 또는 온도 범위에서 형성될 수 있으며, 제 2 온도 또는 온도 범위는 제 1 온도 또는 온도 범위와는 상이할 수 있다(또는 제 1 온도 범위와 부분적으로 중첩될 수 있음).
- [0055] 1106에서, 제 1 구조의 표면은 본딩 재료로 부분적으로 또는 전체적으로 덮여질 수 있다. 예를 들어, 표면은 제 2 구조에 본딩되는 제 1 구조의 표면일 수 있다. 제 1 구조의 표면에 본딩 재료를 도포하는 것은, PVD(Physical Vapor Deposition), CVD(Chemical Vapor Deposition), PECVD(Plasma-enhanced Chemical Vapor Deposition), ALD(Atomic Layer Deposition), PEALD(Plasma Enhanced Atomic Layer Deposition), 스펀온, 습식 증착 공정 또는 기타 유사한 공정에 의해 수행될 수 있다. 본딩 재료는 이산화 실리콘(SiO₂), 산질화 실리콘(SiON), 탄소 도핑된 실리콘 산질화물(SiOCN), 실리콘 탄소 질화물(SiCN), 또는 산탄화 실리콘(SiOC) 중 하나 이상을 포함할 수 있다.
- [0056] 1108에서, 제 2 구조의 표면은 본딩 재료로 부분적으로 또는 전체적으로 덮여질 수 있다. 예를 들어, 표면은 제 1 구조에 본딩되는 제 2 구조의 표면일 수 있다. 제 2 구조의 표면에 본딩 재료를 도포하는 것은, PVD(Physical Vapor Deposition), CVD(Chemical Vapor Deposition), PECVD(Plasma-enhanced Chemical Vapor Deposition), ALD(Atomic Layer Deposition), PEALD(Plasma Enhanced Atomic Layer Deposition), 스펀온, 습식 증착 공정 또는 기타 유사한 공정에 의해 수행될 수 있다. 본딩 재료는 이산화 실리콘(SiO₂), 산질화 실리콘(SiON), 탄소 도핑된 실리콘 산질화물(SiOCN), 실리콘 탄소 질화물(SiCN), 또는 산탄화 실리콘(SiOC) 중 하나 이상을 포함할 수 있다.
- [0057] 1110에서, 제 1 구조 상의 본딩 재료 부분은 제 2 구조 상의 본딩 재료 부분에 본딩될 수 있다. 예를 들어, 제 1 구조 상의 본딩 재료와 제 2 구조 상의 본딩 재료 사이에 본딩이 형성될 수 있다. 예를 들어, 서로 다른 표면의 결합 재료 사이에 결합을 형성하기 위해서, (1) 본드는, 표면을 플라즈마 처리(예를 들어, 질소, 아르곤, 산소, NH₃, Cl, F₂ 등의 고에너지 이온의 노출)함으로써 활성화되어서 브로큰 본드(broken bond) 또는 땀글링 본드(dangling bond)를 생성하고, (2) 표면을 정렬시키며, (3) 예를 들어 실온 내지 400℃의 범위의 온도에서 서로 가압한다.
- [0058] 도 12a는 이득 셀 메모리 디바이스의 다른 단면도를 나타낸다. 이 예에서, 구조(1202)의 커패시터 스토리지 디바이스를 구조(1204)의 판독 트랜지스터에 본딩해서 도전적으로 연결하는데 본 명세서에 설명된 기술을 사용하는 본딩 인터페이스가 사용된다. 이 예에서, 구조(1202)는 구조(1204)를 형성하는데 사용되는 것보다 상대적으로 높은 온도에서 형성될 수 있지만, 그 반대가 적용될 수도 있다.
- [0059] 도 12b는 이득 셀 메모리 디바이스의 다른 단면도를 도시한다. 이 예에서, 구조(1212)를 구조(1214)에 본딩하는데 본 명세서에 설명된 기술을 사용하는 본딩 인터페이스가 사용된다. 예를 들어, 구조(1212)는 기록 및 판독 회로를 포함하는 이득 셀 eDRAM 메모리 디바이스를 포함할 수 있는 반면, 구조(1214)는 본 명세서에서 설명되는 다양한 회로를 포함할 수 있다. 이 예에서, 구조(1212)는 구조(1214)를 형성하는데 사용되는 것보다 상대적으로 낮은 온도에서 형성될 수 있지만, 그 반대가 적용될 수도 있다.
- [0060] 도 12c는 이득 셀 메모리 디바이스의 다른 단면도를 도시한다. 이 예에서, 구조(1222)를 구조(1224)에 본딩하는데 본 명세서에 설명된 기술을 사용하는 본딩 인터페이스가 사용된다. 예를 들어, 구조(1222)는 이득 셀 eDRAM 메모리 디바이스를 포함할 수 있는 반면, 구조(1224)는 본 명세서에 설명된 판독 회로 및 다양한 회로를 포함할

수 있다. 이 예에서, 구조(1222)는, 구조(1224)를 형성하는데 사용된 온도보다 상대적으로 높은 온도에서 형성될 수 있지만, 그 반대가 적용될 수도 있다.

- [0061] 도 13은 이득 셀 메모리 디바이스의 다른 단면도를 도시한다. eDRAM 메모리는 백엔드 금속화 층에서 구현될 수 있으며, 메모리 액세스 및 제어 로직과 CNM 회로 위에 적층될 수 있다. 메모리 스택 옆에 TSV(through silicon via)가 형성될 수 있다. TSV로부터 메모리 스택을 분리하기 위해서, 메모리 스택 주위에 밀폐형 인클로저가 형성될 수도 있다.
- [0062] 도 14는 다양한 디바이스의 평면도의 예를 도시한다. 시스템(1402)의 경우, CNM에 의해 사용되는 혹은 CNM에 의해 생성된 데이터를 저장하는 데 SRAM 메모리가 사용된다. 시스템(1404)은, 전력 전달을 위해 eDRAM가 처리 요소(PE) 및 TSV와 긴밀하게 연결되어 있는, CNM 평면도의 정면도이다. 시스템(1404)의 경우, CNM으로부터 판독한 혹은 CNM에 의해 기록된 데이터를 저장하는데 SRAM 대신에 백엔드 이득 셀 eDRAM 디바이스가 사용될 수 있다.
- [0063] 일부 예에서, SRAM 디바이스 대신 이득 셀 eDRAM 디바이스를 사용하기 위해, 이득 셀 eDRAM 디바이스가 적어도 SRAM 디바이스의 출력을 에플레이트할 수 있게 하는 회로가 추가된다. 이득 셀 eDRAM 디바이스는 SRAM 디바이스보다 판독 액세스 시간이 훨씬 느리지만, 더 높은 밀도의 메모리 셀을 제공하므로, SRAM 디바이스와 동일한 X-Y 평면 풋프린트에 대해 더 많은 데이터를 저장할 수 있다. 일부 예에서, 이득 셀 eDRAM 디바이스는 복수의 메모리 뱅크를 포함할 수 있고, 복수의 메모리 뱅크에 걸쳐서 데이터가 저장되어서, 데이터의 판독이 연속적인 클럭 사이클에서 개시될 수 있으며, SRAM 디바이스에 비해 eDRAM 디바이스의 판독 시간이 더 느림에도 불구하고, 데이터는 SRAM 디바이스의 스루풋으로 또는 거의 동일한 스루풋으로 사용할 수 있다. 일부 예에서 시간 인터리빙(time-interleaving) 액세스 회로는 "lgc"(로직) 색선으로 표시된 활성 실리콘에 구현된다.
- [0064] eDRAM이 SRAM 디바이스를 에플레이트하게 하는 회로는 X-Y 평면 풋프린트를 절약하기 위해 Z 방향으로 백엔드 메모리 아래에 배치되거나, 혹은 eDRAM 디바이스 옆에 배치되어서 추가 X-Y 평면 풋프린트를 사용할 수 있게 한다.
- [0065] 도 14는 SRAM 디바이스 대신에 또는 SRAM 디바이스에 더해서, 백엔드 임베디드 DRAM 메모리 파티션을 사용하는 CNM 회로 모듈의 예시적인 구현예를 도시한다. 다양한 실시예를 통해서, SRAM과 동일한 OPS(Operations Per Second), 스루풋 및 니어-메모리 대역폭을 제공하는 eDRAM 디바이스를 제공한다. 동일한 스루풋 및 성능(OPS 측면에서)을 매칭시키기 위해서, eDRAM 뱅크들이 SRAM 메모리 뱅크들과 동일한 온다이 대역폭을 처리 엔진에 제공한다. 그러나, eDRAM은 캐패시터를 사용해서 데이터에 액세스하므로 SRAM보다 느리고, DRAM 행-액세스는 SRAM 행_액세스 대기 시간보다 느릴 수 있다. 다양한 실시예를 통해서 시간 다중화된 메모리 뱅크 액세스가 수행되게 된다.
- [0066] 도 15는 니어 메모리 및 컴퓨트 블록을 갖는 시스템의 예시적인 논리 블록도를 도시한다. 이 예에서, 니어 메모리(1502)는 메모리 액세스 커맨드 및 클럭 신호(예를 들어, 1GHz 또는 다른 주파수)를 수신하는 SRAM 디바이스이다. 이 예에서 메모리 액세스 커맨드는 128비트의 데이터를 판독하는 판독 커맨드이지만, 임의의 크기의 데이터가 판독될 수는 있다. SRAM은 클럭 신호의 클럭 에지에서 타이밍된 128비트의 데이터를 출력한다. 컴퓨트 블록(1504)의 처리 요소(PE)(1506)는 SRAM 디바이스로부터 데이터를 수신하고, 입력 벡터를 사용해서 계산을 수행한다. 예를 들어, 데이터는 가중치가 될 수 있고, PE(1506)는 트레이닝 또는 추론을 위한 머신 러닝(ML) 또는 인공 지능(AI) 애플리케이션에서 사용되는 신경망의 일부로서, 행렬 곱의 일부로 입력 벡터와 가중치의 곱을 수행할 수 있다.
- [0067] 일부 경우에, DRAM 뱅크는 동일한 X-Y 평면 풋프린트 영역에 대해 SRAM 디바이스보다 더 많은 데이터 저장을 제공할 수 있다. 다양한 실시예는 DRAM 디바이스에 에플레이터 회로를 제공해서, DRAM 디바이스가 SRAM 디바이스와 동일한 스루풋으로 출력 데이터를 제공할 수 있게 한다.
- [0068] 도 16은 제 1 메모리 타입의 메모리가 다른 메모리 타입의 출력을 에플레이트하는 회로가 사용될 수 있는 예시적인 시스템을 도시한다. 예를 들어, 메모리(1604)는 eDRAM 디바이스의 복수의 뱅크를 포함할 수 있다. eDRAM 디바이스 뱅크는 행과 열로 주소 지정이 가능한 스토리지 디바이스를 포함할 수 있으며, 하나 또는 다수의 디바이스 내에 있을 수 있다. 제어 회로(1602)는, SRAM 실시예와 비교해서 유사한 출력(판독) 대역폭을 달성하기 위해서, eDRAM 뱅크에 대해 행해지는 데이터 액세스 요청을 시간 인터리브(time-interleave)할 수 있다. 제어 회로(1602)는 연속적인 클럭 사이클에 걸쳐서 메모리 뱅크(eDRAM#0 내지 eDRAM#7)로의 메모리 액세스 신호(예를 들어, 판독 또는 기록)의 전달을 제어할 수 있다. 다른 수의 메모리 뱅크가 사용될 수도 있다. 예를 들어, 제 1 클럭 사이클 동안에 메모리 액세스가 eDRAM#0으로 전달될 수 있고; 제 2 클럭 사이클(예를 들어, 다음 또는 후

속) 동안에 메모리 액세스가 eDRAM#1으로 전달될 수 있는 등이다.

[0069] 일부 예에서, 제어 회로(1602)는 메모리 액세스 커맨드를 게이팅 회로(1608)로 전달할 수 있다. 게이팅 회로(1608)는, 상이한 타입의 메모리 디바이스(예를 들어, SRAM)에 제공되었을 것과 동일한 클록 신호를 수신할 수 있고 또한 메모리 액세스 커맨드를 수신할 수 있다. 게이팅 회로(1608)는 클록 신호의 에지(예를 들어, 상승 또는 하강)에서 메모리 액세스 커맨드를 전달할 수 있는데, 메모리 액세스 커맨드가 eDRAM#0 내지 eDRAM#7 중 어느 것이 액세스되어야 하는지 인터리브 방식으로 전달되게 할 수 있다.

[0070] 예를 들어, SRAM에서 주소 0 및 1을 판독하기 위해서, 주소 0에 저장된 데이터가 클록 사이클 0에서 액세스될 수 있고, 주소 1의 데이터는 다음 클록 사이클 1 동안 심지어 동일한 메모리 뱅크로부터 액세스될 수 있다. 그러나 eDRAM은 SRAM에 비해 더 느린 판독 시간(예를 들어, 더 높은 판독 대기 시간)을 제공할 수 있다. eDRAM#0 뱅크의 주소 0 및 주소 1에 데이터를 저장하려면, 두 주소로부터 데이터를 취득하기 위해 다수의 클록 주기가 필요할 수 있다. 일부 예들에서, 연속적으로 액세스되는 데이터가 서로 다른 eDRAM 뱅크에 저장될 수 있다. 따라서, 주소 0의 데이터는 eDRAM#0 뱅크에 저장될 수 있고, 주소 1의 데이터는 eDRAM#1 뱅크에 저장될 수 있다. 예를 들어, 가중치 값이 연속적으로 취득되어야 하는 경우, 가중치 값들은 서로 다른 메모리 뱅크에 저장할 수 있다. 더 빠른 액세스를 위해서, 병렬 액세스를 위해 여러 DRAM에 데이터를 스트라이프할 수 있다. 그러나, 느린 접근이 허용되면, 데이터는 단일 뱅크에 저장되고 연속해서 액세스될 수 있다. 따라서 SRAM 속도에 상응하는 속도로(즉, 매칭되는 스트루트) 데이터에 액세스하기 위해서, eDRAM 뱅크로부터 병렬 클록 주기로 또는 연속적인 클록 주기로 데이터가 액세스될 수 있다. 메모리 뱅크로부터의 판독을 완료하는 데 다수의 클록 사이클이 걸리는 경우, eDRAM#0로부터 주소 0에 액세스하는 것이 eDRAM#1에서 주소 1에 액세스하기 전에 수행될 수 있다. 일부 예들에서, 제어 회로(1602)는 메모리 주소의 최상위 비트를 사용해서, 어느 DRAM 뱅크가 액세스될지를 선택할 수 있다.

[0071] 일부 예들에서, 제어 회로(1602)는, 판독 데이터를 버퍼링하는데 사용되는 메모리 버퍼의 크기를 감소시키기 위해서, 메모리(1604)에 전달되는 요청들을 스테거해서(stagger) 메모리 뱅크들로부터 수신되는 데이터의 버퍼링의 양을 감소시킬 수 있다. 메모리(1604)의 메모리 뱅크는 멀티플렉서(1606)에 데이터를 제공할 수 있고, 제어 회로(1602)는 어느 뱅크로부터의 데이터가 컴퓨터 블록(1504)의 처리 요소(1506)로 전달될지를 제어할 수 있다.

[0072] 도 17은, 다수의 뱅크에 액세스하기 위해서 전달될 수 있는 신호의 예를 도시한다. 이 예에서, 메모리 뱅크 0 내지 7에 액세스해서 이들 뱅크로부터 데이터를 판독한다. 뱅크 0에서 7까지의 데이터를 판독하기 위해, 게이팅 회로(1608)는 클록 사이클 0의 상승 에지에서 메모리 뱅크 0에 판독 인에이블 신호를 전달하고, 클록 사이클 1의 상승 에지에서 메모리 뱅크 1에 판독 인에이블 신호를 전달할 수 있는 등이다. 이 예에서는 모든 메모리 뱅크에 판독 요청을 전달하는 것을 보여주지만, 모든 판독 요청을 위해서, 사용 가능한 모든 메모리 뱅크에 액세스할 수 있는 것은 아니다. 액세스되는 메모리 뱅크는, 어느 메모리 뱅크가 액세스되는 데이터를 저장하고 있는지에 따라 달라진다. 데이터는 다수의 메모리 뱅크에 저장되어서 시간-인터리브 액세스를 허용 가능하게 할 수 있고, 이로써 복수의 판독 동작이 시간적으로 부분적으로 중첩된다.

[0073] 도 16을 참조하면, PE(1506)는 벡터-벡터 곱셈을 수행함으로써 내적 연산을 수행하고, 1GHz 클록 주파수에서 이진 비트 요소 또는 다중 비트 요소를 갖는 니어-메모리로부터 데이터를 수신한다. CNM 회로의 연산의 다른 예는, 곱셈 및 누산(multiplication and accumulation(MAC)), 이진 연산(AND, OR, XOR, NOT)(예를 들어, 이진 산술 연산(binary arithmetic), 비트 직렬 연산, 비트 블리드 또는 초차원 컴퓨팅(hyper-dimensional)), 단일-배치(batch) 또는 다중-배치 행렬-행렬 또는 행렬-벡터 내적 연산, ReLu, tanh, sigmoid와 같은 신경망 활성화 함수, soft-max의 맥승, 데이터 압축, 암호화, 미디어 가속기, 문자열 매칭 또는 정규 표현식 가속기, 또는 포인터 추적 회로를 포함한다. PE(1506)은 임의의 타입의 고정 기능 디바이스나 또는 구성 가능한 기능 디바이스가 될 수 있다.

[0074] 도 18은, 다수의 SRAM 뱅크가 액세스되고, 뱅크로부터의 데이터가 연결되거나 결합되어서 데이터를 컴퓨팅 엔진에 제공하는 예시적인 시스템을 도시한다. 이 예에서, 컴퓨터 엔진은 배치-2 벡터-벡터 내적 연산(batch-2 vector-vector dot-product operations)을 수행한다. 예를 들어, 클록 에지 또는 에지들(예를 들어, 상승 또는 하강)에서, SRAM 뱅크 #0 및 SRAM 뱅크 #1은 연결된 데이터 세그먼트를 제공하기 위해 결합된 데이터를 제공할 수 있다. 데이터의 연결은 메모리 버퍼(미도시)를 사용해서 발생할 수도 있다. 이 예에서 SRAM 뱅크 #0 및 SRAM 뱅크 #1은 모두 32b 데이터를 제공할 수 있으며, 32b 데이터는 결합되어서 64b(64비트) 값으로 제공된다. 64b 값은 비트 너비 64b의 단일 요소(예를 들어, 정수 또는 부동 소수점 표현)이거나, 혹은 1비트(예를 들어, 이진 표현), 2비트(예를 들어, {-1, 0, +1} 값에 대한 삼진 표현), 또는 다수-비트(예를 들어, 다수-비트 정수 또는

부동 소수점 표현) 요소의 벡터일 수 있다. 다른 크기의 데이터가 사용될 수도 있고 결합될 수도 있다. 결합된 값은 병렬 처리를 위해서 하나 이상의 PE(처리 요소)에 제공될 수 있다.

- [0075] 일부 예들에서, SRAM으로부터의 데이터 출력은, eDRAM 뱅크로부터의 데이터와 결합되거나 연결될 수 있다. 예를 들어, 도 18에서, SRAM#0은 SRAM 디바이스로서의 동작을 에뮬레이트하는 회로를 가진 DRAM 뱅크로 대체될 수 있으며, SRAM#1은 SRAM 디바이스가 될 수 있다.
- [0076] 도 19는 임베디드 DRAM 메모리로부터의 데이터 출력이 다중 프로세싱 요소(PE)에 제공되는 임베디드 DRAM 메모리의 예를 도시한다. 예를 들어, 멀티플렉서(1606)에 의해 제공되는 128비트 값은 버퍼(1902)에 W0 및 W1으로서 저장될 수 있다. 이 예에서, W0 및 W1은 모두 크기가 64비트일 수도 있고 다른 크기일 수도 있다. 멀티플렉서(1904)는 클록 에지에서 W0 또는 W1을 PE(1906-0 및 1906-1)에 전달할 수 있다. 예를 들어, 멀티플렉서(1904)로부터의 출력은 액세스된 주소의 최하위 비트에 의해 제어될 수 있다.
- [0077] 일부 예에서, 제어(1602), 멀티플렉서(1606), 회로(1608), 멀티플렉서(1904) 및 신호 전도성 경로 중 하나 이상은 Z-방향으로 eDRAM 메모리 디바이스 아래에 배치되어서, X-Y 풋프린트를 절감할 수 있다. 일부 예에서, 제어 회로(1602), 멀티플렉서(1606), 회로(1608), 멀티플렉서(1904) 및 신호 전도성 경로 중 하나 이상이 X-Y 평면에서 eDRAM 메모리 디바이스의 측면에 배치될 수 있다.
- [0078] 도 20은, 더 빠른 메모리 디바이스를 에뮬레이트하는 속도로 메모리 디바이스로부터 데이터를 판독하기 위한 예시적인 프로세스를 도시한다. 2002에서, 연속적인 판독 액세스로 액세스되는 데이터는, 메모리 디바이스의 상이한 메모리 뱅크에 저장된다. 데이터는 상이한 eDRAM 메모리 뱅크들에 저장될 수도 있으며, 여기서 상이한 eDRAM 메모리 뱅크들은 동일한 혹은 상이한 클록 사이클에서 판독을 위해 액세스할 수 있다.
- [0079] 2004에서, 적어도 부분적으로 시간-중첩 판독 동작을 제공하기 위해서, 메모리 디바이스의 상이한 메모리 뱅크들에 판독 요청이 제공된다. 예를 들어, 판독 요청은 제 1 클록 에지에서 메모리 뱅크로 전파될 수 있고, 판독 요청은 후속 클록 에지에서 다른 메모리 뱅크로 전파될 수 있다.
- [0080] 2006에서, 상이한 메모리 뱅크들로부터의 데이터가 적어도 부분적으로 병렬로 출력될 수 있다. 상이한 메모리 뱅크들로부터의 데이터 출력 속도는 제 2 메모리 디바이스의 데이터 출력 속도를 에뮬레이트할 수 있다. 예를 들어, 데이터는, SRAM의 스트루프와 매칭되도록 프로세서 요소로의 데이터 출력의 속도를 에뮬레이트하는 방식으로(예를 들어, SRAM) eDRAM 메모리 뱅크로부터 컴퓨터 블록으로 출력된 후, 다중화될 수 있다.
- [0081] 다양한 실시예는 임의의 처리 기술에 적용 가능하고, 임의의 수의 메모리 파티션에 적용 가능하다. 예로서 배치-2 내적(batch-2 dot-product) 엔진이 제공되며, 여기에 표시된 백엔드 eDRAM 메모리 액세스 기술은 임의의 타입의 컴퓨터 블록, 비트 너비, 요소의 수 등에 적용될 수 있다. 액세스 회로는 물리적으로 백엔드 eDRAM 어레이 바로 아래에 배치되어서, 면적과 배선 리소스를 절감할 수 있다.
- [0082] 도 21은 다양한 시스템의 예를 도시한다. 시스템(2102)은 접속(2108)을 사용해서 프로세서 또는 가속기(2106)에 접속된 메모리(2104)를 포함한다. 백엔드 DRAM 메모리(2104)는 본 명세서에 설명된 기술을 사용해서 형성될 수 있고, 버스 또는 다른 상호 접속부를 사용해서 프로세서 또는 가속기(2106)에 데이터를 제공하거나 혹은 버스 또는 다른 상호 접속부를 사용해서 프로세서 또는 가속기(2106)로부터 데이터를 수신할 수 있다. 접속(2108)은 EMIB(Embedded Multi-die Interconnect Bridge) 또는 칩-칩 상호 접속부일 수 있다.
- [0083] 시스템(2150)은 또 다른 구성으로, 여기서 백엔드 메모리(2152)는 동일한 실리콘 다이에 형성된 프로세싱 요소를 갖는 메모리 뱅크를 포함할 수 있다. 예를 들어, 메모리 뱅크는 본 명세서에 설명된 기술에 따라 처리 요소에 본딩될 수 있다. 복수의 메모리 뱅크와 처리 요소의 쌍이 백엔드 메모리 디바이스 내에 형성될 수 있고, 이로써 복수의 메모리 뱅크가 본 명세서에 설명된 기술에 따라 다수의 처리 요소에 본딩될 수 있다.
- [0084] 처리 요소는, 메모리 뱅크로부터의 콘텐츠를 처리하고, 프로세서 또는 가속기에 출력을 제공하거나 결과를 메모리에 저장하는 CNM 회로일 수 있다. 일부 예에서, 다수의 DRAM 메모리 시스템은 서비스 체인 처리용 버스, 상호 접속부, 네트워크 또는 패브릭을 통해 접속될 수 있으며, 이로써 하나의 디바이스는 데이터를 처리하고, 다른 디바이스에 데이터를 제공해서 저장 및 처리하게 한다. 메모리 뱅크 및 PE의 다양한 실시예는 본 명세서에 설명된 방식으로 구성된 eDRAM 및 CNM일 수 있다. 메모리 뱅크의 다양한 실시예는, eDRAM이 SRAM 디바이스의 동작을 에뮬레이트하는 것을 가능하게 하는 회로를 가진 eDRAM일 수 있다.
- [0085] 백엔드 메모리(2152)는 접속(2156)을 사용해서 프로세서 또는 가속기 디바이스(2154)에 접속될 수 있다. 접속(2156)은 EMIB(Embedded Multi-die Interconnect Bridge) 또는 칩-대-칩 상호접속일 수 있다. 프로세서 또는

가속기 디바이스(2106)는 범용-마이크로프로세서 또는 특수 목적 마이크로프로세서, CPU, GPU, 디지털 신호 프로세서(DSP), 프로그래머블 컨트롤러, 주문형 집적 회로(ASIC), 프로그래머블 로직 디바이스(PLD)와 같은 임의의 타입의 디바이스 동일 수 있고 또는 이러한 디바이스의 조합일 수 있다.

[0086] 도 22는 시스템을 도시한다. 이 시스템은 본 명세서에 기술된 실시예를 사용할 수 있으며, 이로써 메모리 디바이스는 본 명세서에 기술된 방식으로 형성된 또는 SRAM 에뮬레이션 회로를 가진 CNM 피처를 포함한다. 시스템(2100)은, 시스템(2100)에 필요한 처리, 동작 관리 및 명령어 실행을 제공하는 프로세서(2110)를 포함한다. 프로세서(2210)는 임의의 타입의 마이크로프로세서, 중앙 처리 유닛(central processing unit)(CPU), 그래픽 처리 유닛(graphics processing unit)(GPU), 처리 코어 또는 시스템(2200)에 처리를 제공할 다른 처리 하드웨어, 또는 프로세서의 조합을 포함할 수 있다. 프로세서(2210)는 시스템(2200)의 전반적인 동작을 제어하며, 하나 이상의 프로그램 가능 범용 또는 특수 목적 마이크로프로세서, 디지털 신호 프로세서(digital signal processor)(DSP), 프로그램 가능 컨트롤러, 주문형 집적 회로(application specific Integrated circuit)(ASIC), 프로그램 가능 로직 디바이스(programmable logic device)(PLD) 등 또는 이러한 디바이스의 조합일 수 있거나 이를 포함할 수 있다.

[0087] 일례로, 시스템(2200)은 프로세서(2210)에 연결된 인터페이스(2212)를 포함하며, 이러한 인터페이스는 메모리 서브시스템(2220) 또는 그래픽 인터페이스 컴포넌트(2240) 또는 가속기(2242)와 같은 더 높은 대역폭 연결을 필요로 하는 시스템 컴포넌트를 위한 높은 속도의 인터페이스 또는 높은 처리량의 인터페이스를 나타낼 수 있다. 인터페이스(2212)는, 스탠드얼론 컴포넌트일 수도 있고 또는 프로세서 다이에 통합될 수도 있는 인터페이스 회로를 나타낸다. 그래픽 인터페이스(440)가 존재하는 경우, 이는 시스템(2200)의 사용자에게 시각적 디스플레이를 제공하기 위한 그래픽 컴포넌트에 인터페이스한다. 일례로, 그래픽 인터페이스(2240)는 사용자에게 출력을 제공하는 고화질(high definition)(HD) 디스플레이를 구동할 수 있다. 고화질은 약 100 PPI(인치당 픽셀) 이상의 픽셀 밀도를 갖는 디스플레이를 지칭할 수 있고 풀 HD(예를 들어, 1080p), 망막 디스플레이, 4K(초고화질(ultra-high definition) 또는 UHD) 또는 다른 것과 같은 포맷을 포함할 수 있다. 일례로, 디스플레이는 터치스크린 디스플레이를 포함할 수 있다. 일례로, 그래픽 인터페이스(2240)는 메모리(2230)에 저장된 데이터에 기초해서 또는 프로세서(2210)에 의해 실행된 동작에 기초하여 또는 둘 모두에 기초해서 디스플레이를 생성한다.

[0088] 가속기(2242)는 프로세서(2210)에 의해 액세스되거나 사용될 수 있는 고정 기능의 혹은 프로그래밍 가능한 오프로드 엔진일 수 있다. 예를 들어, 가속기(2242) 중 가속기는 압축(DC) 능력, 공개 키 암호화(public key encryption)(PKE)와 같은 암호화 서비스, 암호, 해시/인증 능력, 암호 해독 또는 다른 능력이나 서비스를 제공할 수 있다. 일부 실시예에서, 추가적으로 또는 대안적으로, 가속기(2242) 중 가속기는 본 명세서에 설명된 바와 같은 필드 선택 컨트롤러 능력을 제공한다. 일부 경우에, 가속기(2242)는 CPU 소켓(예를 들어, CPU를 포함하고 CPU와의 전기적 인터페이스를 제공하는 마더 보드 또는 회로 보드와의 커넥터) 내에 통합될 수 있다. 예를 들어, 가속기(2242)는 단일 또는 다중 코어 프로세서, 그래픽 처리 유닛, 논리적 실행 유닛 단일 또는 다중 레벨 캐시, 프로그램 또는 스레드를 독립적으로 실행할 수 있는 기능 유닛, 주문형 집적 회로(ASIC), 신경망 프로세서(neural network processor)(NNP), 프로그램 가능 제어 로직, 및 필드 프로그램 가능 게이트 어레이(field programmable gate array)(FPGA)와 같은 프로그램 가능 처리 요소를 포함할 수 있다. 가속기(2242)가 다수의 신경망, CPU, 프로세서 코어, 범용 그래픽 처리 유닛을 제공할 수 있거나, 또는 그래픽 처리 유닛이 인공지능(artificial intelligence)(AI) 또는 기계 학습(machine learning)(ML) 모델에 의해 사용하는데 이용할 수 있게 만들어질 수 있다. 예를 들어, AI 모델은 강화 학습 체계, Q-학습 체계, 딥(deep)-Q 학습 또는 A3C(Asynchronous Advantage Actor-Critic), 조합 신경망(combinatorial neural network), 재발 조합 신경망(recurrent combinatorial neural network), 또는 다른 AI 또는 ML 모델 중 임의의 것 또는 그의 조합을 사용하거나 포함할 수 있다. 다수의 신경망, 프로세서 코어 또는 그래픽 처리 유닛은 AI 또는 ML 모델에 의해 사용하는데 이용할 수 있게 만들어질 수 있다.

[0089] 메모리 서브시스템(2220)은 시스템(2200)의 메인 메모리를 나타내고 프로세서(2210)에 의해 실행될 코드 또는 루틴을 실행하는데 사용될 데이터 값을 위한 스토리지를 제공한다. 메모리 서브시스템(2220)은 판독 전용 메모리(read-only memory)(ROM), 플래시 메모리, DRAM과 같은 랜덤 액세스 메모리(random access memory)(RAM)의 하나 이상의 변종 또는 다른 메모리 디바이스, 또는 이러한 디바이스의 조합과 같은 하나 이상의 메모리 디바이스(2230)를 포함할 수 있다. 메모리(2230)는 그 중에서도, 시스템(2200)에서 명령어의 실행을 위한 소프트웨어 플랫폼을 제공하는 운영 체제(OS)(2232)을 저장하고 호스팅한다. 나아가, 애플리케이션(2234)은 메모리(2230)로부터 OS(2232)의 소프트웨어 플랫폼 상에서 실행될 수 있다. 애플리케이션(2234) 및 OS(2232)는 명백하게 메모리 영역이 할당되어 있는 가상 머신 환경 또는 컨테이너 환경 내에서 실행될 수 있다. 애플리케이션(2234)은 하

나 이상의 기능의 실행을 수행할 자체의 작동 로직을 가진 프로그램을 나타낸다. 프로세스(2236)는 OS(2232) 또는 하나 이상의 애플리케이션(2234) 또는 조합에 보조 기능을 제공하는 에이전트 또는 루틴을 나타낸다. OS(2232), 애플리케이션(2234) 및 프로세스(2236)는 시스템(2200)에 기능을 제공하는 소프트웨어 로직을 제공한다. 일례로, 메모리 서브시스템(2220)은 커맨드를 생성하여 메모리(2230)에 발행하는 메모리 컨트롤러인 메모리 컨트롤러(2222)를 포함한다. 메모리 컨트롤러(2222)는 프로세서(2210)의 물리적 부분 또는 인터페이스(2212)의 물리적 부분일 수 있다는 것이 이해될 것이다. 예를 들어, 메모리 컨트롤러(2222)는 프로세서(2210)를 구비한 회로에 일체화된 일체형 메모리 컨트롤러일 수 있다.

[0090] 구체적으로 도시되지 않았지만, 시스템(2200)은, 메모리 버스, 그래픽 버스, 인터페이스 버스 등과 같은, 디바이스 사이의 하나 이상의 버스 또는 버스 시스템을 포함할 수 있다는 것이 이해될 것이다. 버스 또는 다른 신호 라인은 컴포넌트를 통신적으로 또는 전기적으로 함께 연결할 수도 있고, 컴포넌트를 통신적으로와 전기적으로 모두 연결할 수도 있다. 버스는 물리적 통신 라인, 지점 간 접속(point-to-point connection), 브릿지, 어댑터, 컨트롤러 또는 다른 회로 또는 조합을 포함할 수 있다. 버스는 예를 들어 시스템 버스, 주변 컴포넌트 인터커넥트(Peripheral Component Interconnect)(PCI) 버스, 하이퍼 트랜스포트 또는 산업 표준 아키텍처(Industry Standard Architecture)(ISA) 버스, 소형 컴퓨터 시스템 인터페이스(Small Computer System Interface)(SCSI) 버스, 범용 직렬 버스(universal serial bus)(USB) 또는 IEEE(Institute of Electrical and Electronics Engineers) 표준 1394 버스(Firewire) 중 하나 이상을 포함할 수 있다.

[0091] 일례로, 시스템(2200)은 인터페이스(2212)에 연결될 수 있는 인터페이스(2214)를 포함한다. 일례로, 인터페이스(2214)는 스탠드얼론 컴포넌트 및 통합 회로를 포함할 수 있는 인터페이스 회로를 나타낸다. 일례로, 다수의 사용자 인터페이스 컴포넌트 또는 주변 컴포넌트 또는 둘 모두가 인터페이스(2214)에 연결된다. 네트워크 인터페이스(2250)는 하나 이상의 네트워크를 통해 원격 디바이스(예를 들어, 서버 또는 다른 컴퓨팅 디바이스)와 통신하는 능력을 시스템(2200)에 제공한다. 네트워크 인터페이스(2250)는 이더넷 어댑터, 무선 상호연결 컴포넌트, 셀룰러 네트워크 상호연결 컴포넌트, USB(범용 직렬 버스), 또는 다른 유선이나 무선 표준 기반 또는 전용 인터페이스를 포함할 수 있다. 네트워크 인터페이스(2250)는, 메모리에 저장된 데이터를 전송하는 것을 포함할 수 있는, 동일한 데이터 센터 또는 랙에 있는 디바이스 또는 원격 디바이스에 데이터를 송신할 수 있다. 네트워크 인터페이스(2250)는, 수신된 데이터를 메모리에 저장하는 것을 포함할 수 있는, 원격 디바이스로부터 데이터를 수신할 수 있다. 다양한 실시예가 네트워크 인터페이스(2250), 프로세서(2210) 및 메모리 서브시스템(2220)과 관련하여 사용될 수 있다.

[0092] 일례로, 시스템(2200)은 하나 이상의 입력/출력(input/output)(I/O) 인터페이스(들)(2260)를 포함한다. I/O 인터페이스(2260)는 사용자가 시스템(2200)과 상호 작용하는 하나 이상의 인터페이스 컴포넌트(예를 들어, 오디오, 영숫자, 촉각/터치 또는 다른 인터페이스하는 것)를 포함할 수 있다. 주변기기 인터페이스(2270)는 위에서 구체적으로 언급되지 않은 임의의 하드웨어 인터페이스를 포함할 수 있다. 주변기기는 일반적으로 시스템(2200)에 종속적으로 접속되는 디바이스를 말한다. 종속 접속이란 시스템(2200)이 동작이 실행되고 사용자가 상호 작용하는 소프트웨어 플랫폼 또는 하드웨어 플랫폼 또는 둘 모두를 제공하는 접속이다.

[0093] 일례로, 시스템(2200)은 비휘발성 방식으로 데이터를 저장하는 저장 서브시스템(2280)을 포함한다. 일례로, 특정 시스템 구현예에서, 스토리지(2280)의 적어도 특정 컴포넌트는 메모리 서브시스템(2220)의 컴포넌트와 중복될 수 있다. 저장 서브시스템(2280)은 스토리지 디바이스(들)(2284)를 포함하며, 스토리지 디바이스는 하나 이상의 자기, 고체 상태 또는 광학 기반 디스크 또는 조합과 같이 비휘발성 방식으로 대량의 데이터를 저장하기 위한 임의의 통상적 매체이거나 이를 포함할 수 있다. 스토리지(2284)는 코드 또는 명령어 및 데이터(2286)를 지속적인 상태로 보유한다(즉, 시스템(2200)으로의 전력이 중단되어도 값은 유지된다). 스토리지(2284)가 일반적으로 "메모리"인 것으로 간주될 수 있지만, 메모리(2230)는 전형적으로 명령어를 프로세서(2210)에 제공하는 실행 또는 작동 메모리이다. 스토리지(2284)가 비휘발성인 반면에, 메모리(2230)는 휘발성 메모리를 포함할 수 있다(즉, 시스템(2200)으로의 전력이 차단되면 데이터의 값 또는 상태는 확정적이지 않다). 일례로, 저장 서브시스템(2280)은 스토리지(2284)와 인터페이스하는 컨트롤러(2281)를 포함한다. 일례로, 컨트롤러(2281)는 인터페이스(2214) 또는 프로세서(2210)의 물리적 부분이거나 또는 프로세서(2210) 및 인터페이스(2214) 둘 모두에서 회로 또는 로직을 포함할 수 있다.

[0094] 휘발성 메모리는 디바이스로의 전력이 차단되면, 상태(및 따라서 저장된 데이터)가 확정적이지 않은(indeterminate) 메모리이다. 다이나믹 휘발성 메모리는 디바이스에 저장된 데이터를 상태 유지를 위해 리프레시하는 것을 필요로 한다. 다이나믹 휘발성 메모리의 하나의 예는 DRAM(다이나믹 랜덤 액세스 메모리) 또는 동기 DRAM(Synchronous DRAM)(SDRAM)과 같은 몇몇 변종을 포함한다. 본 명세서에 설명된 바와 같은 메모리 서브시

시스템은 DDR3(Double Data Rate 버전 3, 2007년 6월 27일 JEDEC(Joint Electronic Device Engineering Council)에 의해 최초 출시), DDR4(DDR 버전 4, JEDEC에 의해 2012년 9월에 공개된 초기 사양), LPDDR3(Low Power DDR 버전 3, JESD209-3B, JEDEC에 의해 2013년 8월), LPDDR4(LPDDR 버전 4, JESD209-4, JEDEC에 의해 2014년 8월 최초 발표), WIO2(Wide Input/output 버전 2, JESD229-2, JEDEC에 의해 2014년 8월 최초 발표), HBM(High Bandwidth Memory, JESD325, JEDEC에 의해 2013년 10월 최초 발표, LPDDR5(현재 JEDEC에 의해 논의 중), 현재 JEDEC에 의해 논의 중인 HBM2(HBM 버전 2)이나 다른 것 또는 메모리 기술의 조합, 및 이러한 사양의 파생 또는 확장에 기반한 기술과 같은 다수의 메모리 기술과 호환될 수 있다. JEDEC 표준은 www.jedec.org에서 입수할 수 있다.

[0095] 비휘발성 메모리(non-volatile memory)(NVM) 디바이스는 디바이스로의 전력이 차단 되더라도 그의 상태가 확정적인(determinate) 타입의 메모리일 수 있다. 하나의 실시예에서, NVM 디바이스는 NAND 기술과 같은 블록 어드레싱 가능 메모리 디바이스, 또는 보다 구체적으로는 다중 임계 레벨 NAND 플래시 메모리(예를 들어, 단일 레벨 셀("SLC"), 다중 레벨 셀("MLC"), 쿼드 레벨 셀("QLC"), 트리플 레벨 셀("TLC") 또는 몇몇 다른 NAND)를 포함할 수 있다. NVM 디바이스는 또한 바이트-어드레싱 가능 라이트-인-플레이스 3 차원 크로스 포인트 메모리 디바이스(byte-addressable write-in-place three dimensional cross point memory device), 또는 다른 바이트-어드레싱 가능 라이트-인-플레이스 NVM 디바이스(영구 메모리라고도 지칭함), 이를테면 단일 또는 다중 레벨 상변화 메모리(PCM) 또는 스위치가 있는 상변화 메모리(phase change memory with a switch (PCMS)), 칼코게나이드 상변화 재료(예를 들어, 칼코게나이드 유리)를 사용하는 NVM 디바이스, 금속 산화물베이스, 산소 공극 베이스 및 전도성 브리지 랜덤 액세스 메모리(Conductive Bridge Random Access Memory)(CB-RAM)를 포함한 저항성 메모리, 나노와이어 메모리, 강유전체 트랜지스터 랜덤 액세스 메모리(Ferromagnetic Random Access Memory, FRAM), 램비스터 기술을 통합시킨 자기 저항 랜덤 액세스 메모리(MRAM) 메모리, 스핀 전달 토크(STT)-MRAM, 스핀트로닉 자기 접합 메모리(spintronic magnetic junction memory) 기반 디바이스, 자기 터널링 접합(magnetic tunneling junction)(MTJ) 기반 디바이스, 도메인 월(Domain Wall)(DW) 및 스핀 오비탈 전송(Spin Orbit Transfer)(SOT) 기반 디바이스, 사이리스터 기반 메모리 디바이스 또는 임의의 위의 것의 조합, 또는 다른 메모리를 포함할 수 있다.

[0096] 전력 소스(도시되지 않음)는 시스템(2200)의 컴포넌트에 전력을 제공한다. 보다 구체적으로, 전력 소스는 전형적으로 시스템(2200)의 하나 또는 다수의 전력 공급 디바이스와 인터페이스하여 전력을 시스템(2200)의 컴포넌트에 제공한다. 일례로, 전력 공급 디바이스는 벽 콘센트에 플러그되는 AC-DC(교류 대 직류) 어댑터를 포함한다. 이러한 AC 전력은 재생 가능 에너지(예를 들어, 태양광 발전) 전력 소스일 수 있다. 일례로, 전력 소스는 외부 AC-DC 변환기와 같은 DC 전력 소스를 포함한다. 일례로, 전력 소스 또는 전력 공급 디바이스는 충전 필드에 근접해서 충전하는 무선 충전 하드웨어를 포함한다. 일례로, 전력 소스는 내부 배터리, 교류 공급 디바이스, 모션 기반 전력 공급 디바이스, 태양광 전력 공급 디바이스 또는 연료 전지 소스를 포함할 수 있다.

[0097] 일례로, 시스템(2200)은 상호 접속된 프로세서, 메모리, 스토리지, 네트워크 인터페이스 및 다른 컴포넌트를 사용해서 구현될 수 있다. 이더넷(IEEE 802.3), 원격 직접 메모리 액세스(Remote Direct Memory Access)(RDMA), 인피니밴드(InfiniBand), 인텔 광역 RDMA 프로토콜(Internet Wide Area RDMA Protocol)(iWARP), 퀵 UDP 인터넷 커넥션(Quick UDP Internet Connections)(QUIC), RoCE(RDMA over Converged Ethernet), PCIe(Peripheral Component Interconnect express), 인텔 퀵패스 인터커넥트(Intel QuickPath Interconnect)(QPI), 인텔 울트라패스 인터커넥트(Intel Ultra Path Interconnect)(UPI), 인텔 온-칩 시스템 패브릭(Intel On-Chip System Fabric)(IOSF), 옴니패스(Omnipath), 컴퓨트 익스프레스 링크(Compute Express Link)(CXL), 하이퍼트랜스포트(HyperTransport), 고속 패브릭, NVLink, 최신 마이크로컨트롤러 버스 아키텍처(Advanced Microcontroller Bus Architecture)(AMBA) 인터커넥트, OpenCAPI, Gen-Z, CCIX, 3GPP 롱 텀 에볼루션(Long Term Evolution)(LTE)(4G), 3GPP 5G 및 그 변형 등과 같은 고속 인터커넥트가 사용될 수 있다. NVMe-oF(NVMe over Fabrics) 또는 NVMe와 같은 프로토콜을 사용하는 가상화된 저장 노드에 데이터가 저장되고 액세스될 수 있다.

[0098] 본 명세서의 실시예는 스위치, 라우터, 랙, 및 블레이드 서버, 이를테면 데이터 센터 및/또는 서버 팜 환경에서 사용되는 것과 같은 다양한 타입의 컴퓨팅 및 네트워킹 장비에서 구현될 수 있다. 데이터 센터 및 서버 팜에서 사용되는 서버는 랙 기반 서버 또는 블레이드 서버와 같은 어레이형 서버 구성을 포함한다. 이들 서버는 서버 세트를 근거리 네트워크(Local Area Network)(LAN) 사이에 적절한 스위칭 및 라우팅 시설을 갖춘 LAN으로 분할하는 것과 같은, 다양한 네트워크 규정을 통해 통신으로 상호 연결되어 개인 인트라넷을 형성한다. 예를 들어, 클라우드 호스팅 시설은 전형적으로 다수의 서버가 있는 대형 데이터 센터를 사용할 수 있다. 블레이드는 서버-타입 기능, 즉 "서버 온 카드(server on a card)"를 수행하도록 구성된 별도의 컴퓨팅 플랫폼을 포함한다. 따라서, 각각의 블레이드는 적절한 집적 회로(integrated circuit)(IC)와 보드에 실장된 다른 컴포넌트를 결합하

기 위한 내부 배선(즉, 버스)을 제공하는 메인 인쇄 회로 기판(메인 보드)을 비롯한, 통상적인 서버에 공통적인 컴포넌트를 포함한다.

[0099] 메모리 풀, 스토리지 풀 또는 가속기를 포함하고 NVMe-oF를 사용하여 스토리지 또는 메모리 트랜잭션을 스케일 아웃하기 위해 데이터 센터에서 다양한 실시예가 사용될 수 있다. 다양한 실시예는 분산형 리소스(예를 들어, 컴퓨터, 메모리, 스토리지, 가속기, 스토리지)를 사용하는 클라우드 서비스 제공자에 의해 사용될 수 있다. 분산형 리소스는, 기지국, 포그 데이터 센터, 에지 데이터 센터 또는 원격 데이터 센터 중 하나 이상에 위치될 수 있다. 유선 프로토콜 또는 무선 프로토콜(예를 들어, 3GPP LTE(Long Term Evolution)(4G) 또는 3GPP 5G), 온-프레미스 데이터 센터, 오프-프레미스 데이터 센터, 에지 네트워크 요소, 포그 네트워크 요소 및/또는 하이브리드 데이터 센터(예를 들어, 가상화, 클라우드 및 소프트웨어-정의 네트워킹을 사용해서 물리적 데이터 센터 및 분산형 다중 클라우드 환경에서 애플리케이션 워크로드를 제공하는 데이터 센터)를 사용해서 통신을 지원하는 다양한 실시예가, 기지국에서 사용될 수 있다.

[0100] 도 23은 다수의 컴퓨팅 랙(2302)을 포함하는 환경(2300)을 도시하며, 각각의 컴퓨팅 랙은 랙 최상단(Top of Rack)(ToR) 스위치(2304), 포드 관리자(pod manager)(2306) 및 복수의 풀링된 시스템 드로어(pooled system drawer)를 포함한다. 다양한 실시예가 스위치에서 사용될 수 있다. 일반적으로, 풀링된 시스템 드로어는 풀링된 컴퓨터 드로어 및 풀링된 스토리지 드로어를 포함할 수 있다. 선택적으로, 풀링된 시스템 드로어는 또한 풀링된 메모리 드로어 및 풀링된 입력/출력(I/O) 드로어를 포함할 수 있다. 예시된 실시예에서, 풀링된 시스템 드로어는 인텔® XEON® 풀링된 컴퓨터 드로어(2308) 및 인텔® ATOM™ 풀링된 컴퓨터 드로어(2310), 풀링된 스토리지 드로어(2312), 풀링된 메모리 드로어(2314) 및 풀링된 I/O 드로어(2316)를 포함한다. 각각의 풀링된 시스템 드로어는 40 기가비트/초(Gigabit/second)(Gb/s) 또는 100 Gb/s 이더넷 링크 또는 100+ Gb/s 실리콘 광자(Silicon Photonic)(SiPh) 광학 링크와 같은 고속 링크(2318)를 통해 ToR 스위치(2304)에 접속된다.

[0101] 다수의 컴퓨팅 랙(2302)은 네트워크(2320)로의 접속에 의해 도시된 바와 같이, (예를 들어, 포드-레벨 스위치(pod-level switch) 또는 데이터 센터 스위치를 향한) 그들의 ToR 스위치(2304)를 통해 상호 연결될 수 있다. 일부 실시예에서, 컴퓨팅 랙(2302)의 그룹은 포드 관리자(들)(2306)를 통해 별도의 포드로서 관리된다. 하나의 실시예에서, 단일 포드 관리자는 포드 내의 모든 랙을 관리하는데 사용된다. 다른 방안으로, 분산형 포드 관리자가 포드 관리 동작에 사용될 수 있다.

[0102] 환경(2300)은 환경의 다양한 측면을 관리하는데 사용되는 관리 인터페이스(2322)를 더 포함한다. 이것은 랙 구성 데이터(2324)로서 저장된 대응하는 파라미터로 랙 구성을 관리하는 것을 포함한다. 환경(2300)은 컴퓨팅 랙에 사용될 수 있다.

[0103] 다양한 예는 하드웨어 요소, 소프트웨어 요소, 또는 둘 모두의 조합을 사용하여 구현될 수 있다. 일부 예에서, 하드웨어 요소는 디바이스, 컴포넌트, 프로세서, 마이크로프로세서, 회로, 회로 요소(예를 들어, 트랜지스터, 저항기, 커패시터, 인덕터 등), 집적 회로, ASIC, PLD, DSP, FPGA, 메모리 유닛, 로직 게이트, 레지스터, 반도체 디바이스, 칩, 마이크로 칩, 칩 세트 등을 포함할 수 있다. 일부 예에서, 소프트웨어 요소는 소프트웨어 컴포넌트, 프로그램, 애플리케이션, 컴퓨터 프로그램, 애플리케이션 프로그램, 시스템 프로그램, 머신 프로그램, 운영 체제 소프트웨어, 미들웨어, 펌웨어, 소프트웨어 모듈, 루틴, 서브루틴, 함수, 방법, 절차, 소프트웨어 인터페이스, API, 명령어 세트, 컴퓨팅 코드, 컴퓨터 코드, 코드 세그먼트, 컴퓨터 코드 세그먼트, 단어, 값, 기호 또는 이들의 임의의 조합을 포함할 수 있다. 하드웨어 요소 및/또는 소프트웨어 요소를 사용하는 예가 구현되는지를 결정하는 것은 임의의 개수의 인자, 이를테면 주어진 구현에 바람직한 것으로서 원하는 계산 속도, 전력 레벨, 내열성(heat tolerance), 처리 사이클 예산, 입력 데이터 레이트, 출력 데이터 레이트, 메모리 자원, 데이터 버스 속도 및 다른 설계 또는 성능 제약에 따라 변동할 수 있다. 하드웨어, 펌웨어 및/또는 소프트웨어 요소는 본 명세서에서 "모듈", "로직", "회로" 또는 "회로"로 일괄적으로 또는 개별적으로 지칭될 수 있다는 것을 주목하여야 한다. 프로세서는 하드웨어 상태 머신, 디지털 제어 로직, 중앙 처리 유닛 또는 임의의 하드웨어, 펌웨어 및/또는 소프트웨어 요소의 하나 이상의 조합일 수 있다.

[0104] 일부 예는 제조 물품 또는 적어도 하나의 컴퓨터 판독 가능 매체를 사용하여 구현될 수 있거나 또는 적어도 하나의 컴퓨터 판독 가능 매체로서 구현될 수 있다. 컴퓨터 판독 가능 매체는 로직을 저장하는 비일시적 저장 매체를 포함할 수 있다. 일부 예에서, 비일시적 저장 매체는 휘발성 메모리 또는 비휘발성 메모리, 착탈식 또는 비착탈식 메모리, 소거 가능 또는 소거 불가능 메모리, 기입 가능 또는 재기입 가능 메모리 등을 비롯한, 전자 데이터를 저장할 수 있는 하나 이상의 타입의 컴퓨터 판독 가능 저장 매체를 포함할 수 있다. 일부 예에서, 로직은 다양한 소프트웨어 요소, 이를테면 소프트웨어 컴포넌트, 프로그램, 애플리케이션, 컴퓨터 프로그램, 애플

리케이션 프로그램, 시스템 프로그램, 머신 프로그램, 오퍼레이팅 시스템 소프트웨어, 미들웨어, 펌웨어, 소프트웨어 모듈, 루틴, 서브루틴, 함수, 방법, 절차, 소프트웨어 인터페이스, API, 명령어 세트, 컴퓨팅 코드, 컴퓨터 코드, 코드 세그먼트, 컴퓨터 코드 세그먼트, 단어, 값, 기호 또는 이들의 임의의 조합을 포함할 수 있다.

[0105] 일부 예에 따르면, 컴퓨터-관독 가능 매체는 머신, 컴퓨팅 디바이스 또는 시스템에 의해 실행될 때, 머신, 컴퓨팅 디바이스 또는 시스템으로 하여금 설명된 예에 따른 방법 및/또는 동작을 수행하게 하는 명령어를 저장 또는 유지하는 비일시적 저장 매체를 포함할 수 있다. 명령어는 임의의 적합한 타입의 코드, 이를테면, 소스 코드, 컴파일된 코드, 해석된 코드, 실행 코드, 정적 코드, 동적 코드 등을 포함할 수 있다. 명령어는 머신, 컴퓨팅 디바이스 또는 시스템에게 특정 기능을 수행하도록 지시하기 위한 미리 정의된 컴퓨터 언어, 방식 또는 실행에 따라 구현될 수 있다. 명령어는 임의의 적합한 하이-레벨, 로우-레벨, 객체 지향, 시각적, 컴파일된 및/또는 해석된 프로그래밍 언어를 사용하여 구현될 수 있다.

[0106] 적어도 하나의 예의 하나 이상의 양태는 머신, 컴퓨팅 디바이스 또는 시스템에 의해 관독될 때 머신, 컴퓨팅 디바이스 또는 시스템으로 하여금 본 명세서에 설명된 기술을 수행할 로직을 제조하게 하는, 프로세서 내의 다양한 로직을 나타내는 적어도 하나의 머신 관독 가능 매체 상에 저장된 대표적인 명령어에 의해 구현될 수 있다. "IP 코어"로 알려진 이러한 표현은 유형의(tangible) 머신 관독 가능 매체에 저장될 수 있고 다양한 고객 또는 제조 시설에 공급되어 실제로 로직 또는 프로세서를 만드는 제조 머신에 로딩될 수 있다.

[0107] "하나의 예" 또는 "예"라는 문구가 나온다고 해서 반드시 모두가 동일한 예 또는 실시예를 언급하는 것은 아니다. 본 명세서에 설명된 임의의 양태는, 그 양태가 동일한 도면 또는 요소와 관련하여 설명되는지에 관계없이, 본 명세서에 설명된 임의의 다른 양태 또는 유사한 양태와 조합될 수 있다. 첨부 도면에 도시된 블록 기능의 분할, 생략 또는 포함은 이들 기능을 구현하기 위한 하드웨어 컴포넌트, 회로, 소프트웨어 및/또는 요소가 실시예에서 반드시 분할되거나, 생략되거나 또는 포함되는 것으로 추론하는 것은 아니다.

[0108] 일부 예는 "연결된" 및 "접속된"이라는 표현을 그 파생어와 함께 사용하여 설명될 수 있다. 이들 용어는 반드시 서로 동의어로 의도되는 것은 아니다. 예를 들어, "접속된" 및/또는 "연결된"이라는 용어를 사용하는 설명은 둘 이상의 요소가 서로 직접 물리적 또는 전기적으로 컨택트하고 있다는 것을 나타낼 수 있다. 그러나, "연결된"이라는 용어는 두 개 이상의 요소가 서로 직접적으로는 컨택트하고 있지 않지만, 그럼에도 여전히 서로 협력하거나 상호 작용한다는 것을 또한 의미할 수 있다.

[0109] 본 명세서에서 "제 1", "제 2" 등의 용어는 임의의 순서, 수량 또는 중요도를 나타내는 것이 아니라, 한 요소를 다른 요소와 구별하기 위해 사용된다. 본 명세서에서 "한" 및 "하나"라는 용어는 수량의 제한을 나타내는 것이 아니라, 언급된 항목 중 적어도 하나의 존재를 나타낸다. 신호와 관련하여 본 명세서에서 사용되는 "어설트된(asserted)"이라는 용어는 신호가 활성화되는 신호의 상태를 나타내며, 로직 0 또는 로직 1 중 임의의 로직 레벨을 신호에 적용함으로써 달성될 수 있다. "뒤에 이어" 또는 "이후"라는 용어는 어떤 다른 이벤트 또는 이벤트들 직후 또는 이후의 다음이라 말할 수 있다. 대안적인 실시예에 따라 단계의 다른 시퀀스가 수행될 수 있다. 또한, 특정 애플리케이션에 따라 추가 단계가 추가되거나 제거될 수 있다. 변경의 임의의 조합이 사용될 수 있으며, 본 개시 내용의 혜택을 받는 관련 기술분야의 통상의 기술자는 본 개시내용의 많은 변형, 수정 및 대안적인 실시예를 이해할 것이다.

[0110] 달리 구체적으로 언급되지 않는 한, "X, Y 또는 Z 중 적어도 하나"라는 문구와 같은 선언적 언어(disjunctive language)는 맥락 내에서, 항목, 용어 등이 X, Y 또는 Z 이거나 또는 이들의 임의의 조합(예를 들면, X, Y 및/또는 Z)일 수 있다는 것을 제시하는데 일반적으로 사용되는 것처럼 달리 이해된다. 따라서, 이러한 선언적 언어는 일반적으로 특정 실시예가 각각 제시될 X 중 적어도 하나, Y 중 적어도 하나 또는 Z 중 적어도 하나를 필요로 한다는 것을 의도하지 않으며, 암시해서는 안된다. 나아가, 달리 구체적으로 언급되지 않는 한, "X, Y 및 Z 중 적어도 하나"라는 문구와 같은 접속적 언어(conjunctive language)는 또한 "X, Y, 및/또는 Z"를 포함하여, X, Y, Z 또는 이들의 임의의 조합을 의미하는 것으로 이해되어야 한다.

[0111] 본 명세서에 개시된 디바이스, 시스템, 및 방법의 예시적인 예가 아래에 제공된다. 디바이스, 시스템 및 방법의 실시예는 아래에서 설명되는 예 중 임의의 하나 이상의 조합 및 그의 임의의 조합을 포함할 수 있다.

[0112] 예 1은 장치를 포함하며, 이는, 이득 셀 임베디드 동적 랜덤 액세스 메모리(eDRAM) 디바이스를 포함하고, 이득 셀 eDRAM 디바이스는, 기록 회로와, 스토리지 디바이스와, 관독 회로를 포함하며, 관독 회로는 적어도 부분적으로 기록 회로 아래에 기록 회로의 평면과는 다른 평면에 위치되고, 스토리지 디바이스는 기록 회로를 관독 회로에 연결시키며, 기록 회로와 관독 회로 중 하나 이상이 스토리지 디바이스에 본딩된다.

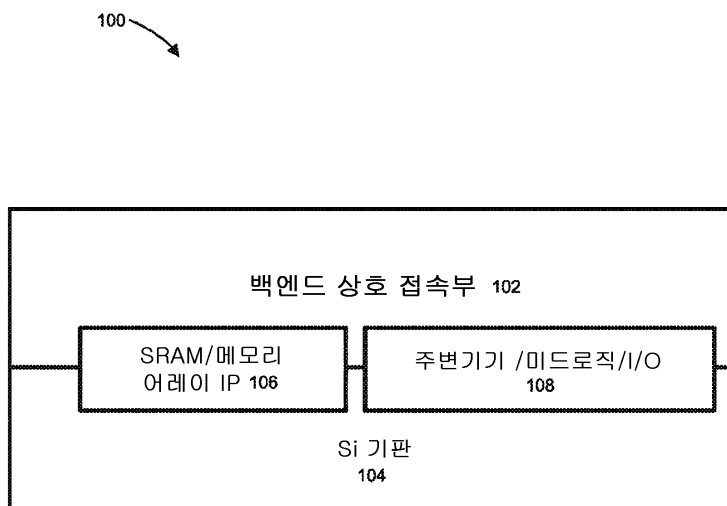
- [0113] 예 2는 임의의 예와 조합될 수 있으며, 기록 회로는, 드레인 영역을 가진 층과, 채널 층과, 게이트 산화물 층과, 기록 워드 라인 인터페이스를 가진 게이트 층과, 기록 비트 라인 인터페이스를 가진 소스 영역을 포함한다.
- [0114] 예 3은 임의의 예와 조합될 수 있으며, 판독 회로는, 채널 영역과, 게이트 산화물 층과, 게이트 층과, 판독 비트 라인 인터페이스를 가진 소스 영역과, 판독 워드 라인 인터페이스를 가진 드레인 영역을 포함한다.
- [0115] 예 4는 임의의 예와 조합될 수 있으며, 기록 회로는, 층과, 드레인 영역과, 층을 통해 드레인 영역을 스토리지 디바이스에 연결하는 도전성 영역과, 채널 층과, 게이트 산화물 층과, 게이트 층과, 기록 비트 라인 인터페이스를 가진 소스 영역을 포함한다.
- [0116] 예 5는 임의의 예와 조합될 수 있으며, 판독 회로는, 게이트 영역과, 게이트 영역 위에 적어도 부분적으로 형성된 게이트 산화물 영역과, 게이트 산화물 영역 위에 적어도 부분적으로 형성된 채널 영역과, 채널 영역의 일부 위에 형성된 소스 영역과, 채널 영역의 일부 위에 형성된 드레인 영역을 포함한다.
- [0117] 예 6은 임의의 예와 조합될 수 있으며, 기록 회로는, 채널 영역과, 채널 영역 위에 적어도 부분적으로 형성된 게이트 산화물 영역과, 게이트 산화물 영역 위에 적어도 부분적으로 형성된 게이트 영역과, 채널 영역의 일부 위에 형성된 소스 영역과, 채널 영역의 일부 위에 형성된 드레인 영역을 포함한다.
- [0118] 예 7은 임의의 예와 조합될 수 있으며, 기록 회로는, 채널 영역과, 채널 영역의 일부 아래에 형성된 소스 영역과, 채널 영역의 일부 아래에 형성된 드레인 영역과, 채널 영역 위에 적어도 부분적으로 형성된 게이트 산화물 영역과, 게이트 산화물 영역 위에 적어도 부분적으로 형성된 게이트 영역을 포함한다.
- [0119] 예 8은 임의의 예와 조합될 수 있으며, 스토리지 디바이스는, 적어도 하나의 커패시터 및 적어도 하나의 커패시터와 컨택트하는 유전체를 포함한다.
- [0120] 예 9는 임의의 예와 조합될 수 있으며, 스토리지 디바이스는 COB(Capacitor Over Bitline)를 포함한다.
- [0121] 예 10은 임의의 예와 조합될 수 있으며, 기록 회로는, 폴리-Si, Si, Ge, 폴리-Ge, III-V, GaN, MoS₂, WSe₂, MoSe₂, WSe₂, InS, HfS, ZnS, ZnSe, In₂O₃, ZnO, AZO, IGZO 또는 IZO 중 하나 이상을 포함한다.
- [0122] 예 11은 임의의 예와 조합될 수 있으며, 기록 회로는, 본딩 재료를 사용해서 스토리지 디바이스에 본딩된다.
- [0123] 예 12는 임의의 예와 조합될 수 있으며, 본딩 재료는, 이산화 실리콘(SiO₂), 산질화 실리콘(SiON), 탄소 도핑된 산질화 실리콘(SiOCN), 실리콘 탄소 질화물(SiCN), 또는 산탄화 실리콘(SiOC) 중 하나 이상을 포함한다.
- [0124] 예 13은 임의의 예와 조합될 수 있으며, 판독 회로는 본딩 재료를 사용해서 스토리지 디바이스에 본딩된다.
- [0125] 예 14는 임의의 예와 조합될 수 있으며, 본딩 재료는, 이산화 실리콘(SiO₂), 산질화 실리콘(SiON), 탄소 도핑된 산질화 실리콘(SiOCN), 실리콘 탄소 질화물(SiCN), 또는 산탄화 실리콘(SiOC) 중 하나 이상을 포함한다.
- [0126] 예 15는 임의의 예와 조합될 수 있으며, 장치는, 컴퓨터 니어 메모리(compute near memory; CNM) 회로를 포함하고, CNM 회로는 판독 회로 또는 기록 회로에 본딩된다.
- [0127] 예 16은 임의의 예와 조합될 수 있으며, 이득 셀 eDRAM 디바이스로부터의 출력을 상이한 유효 출력 레이트로 제공하기 위한 메모리 에물레이션 회로를 포함한다.
- [0128] 예 17은 임의의 예와 조합될 수 있으며, 장치를 포함하고, 이는 이득 셀 임베디드 동적 랜덤 액세스 메모리(eDRAM) 메모리 셀을 포함하는 메모리 디바이스 - 이득 셀 eDRAM 메모리 셀은, 스토리지 셀 위에 적어도 부분적으로 형성된 기록 회로와, 스토리지 셀 아래에 적어도 부분적으로 형성된 판독 회로를 포함함 - 와, 메모리 디바이스에 본딩된 컴퓨터 니어 메모리 디바이스를 포함한다.
- [0129] 예 18은 임의의 예와 조합될 수 있으며, 컴퓨터 니어 메모리 디바이스는, 벡터-벡터 곱셈, 곱셈 및 누산(multiplication and accumulation(MAC)), AND, OR, XOR, NOT, 신경망 활성화 연산(neural network activation operations), soft-max의 멱승(exponentiation for soft-max), 데이터 압축, 암호화, 미디어 가속기, 문자열 매칭 또는 정규 표현식 가속기, 또는 포인터 추적(pointer chasing) 중 하나 이상을 수행하는 것이다.
- [0130] 예 19는 임의의 예와 조합될 수 있으며, 메모리 디바이스의 출력을 제공해서 제 2 메모리 디바이스의 출력을 에물레이트하는 에물레이션 회로를 더 포함한다.

- [0131] 예 20은 임의의 예와 조합될 수 있으며, 에플레이션 회로는 컨트롤러, 멀티플렉서, 또는 레지스터 중 하나 이상을 포함한다.
- [0132] 예 21은 임의의 예와 조합될 수 있으며, 메모리 디바이스 및 컴퓨터 니어 메모리 디바이스에 연결된 프로세서를 더 포함한다.
- [0133] 예 22는 임의의 예와 조합될 수 있으며, 컴퓨터 니어 메모리 디바이스는 본딩 재료를 사용해서 메모리 디바이스에 본딩되고, 본딩 재료는, 이산화 실리콘(SiO₂), 산질화 실리콘(SiON), 탄소 도핑된 산질화 실리콘(SiOCN), 실리콘 탄소 질화물(SiCN), 또는 산탄화 실리콘(SiOC) 중 하나 이상을 포함한다.
- [0134] 예 23은 임의의 예와 조합될 수 있으며, 이득 셀 메모리 디바이스를 형성하는 방법을 포함하고, 이는, 제 1 기판 상에, 워드 라인 드라이버, 판독 회로 및 활성 핀(active fin)을 포함하는 제 1 영역을 형성하는 단계와, 제 1 영역 위에 유전체 영역을 형성하는 단계와, 유전체 영역에 적어도 하나의 저장 영역을 형성하는 단계와, 유전체 영역 상에 비실리콘 기반(non-silicon-based) 기록 회로의 층을 형성하는 단계와, 비실리콘 기반 기록 회로의 층 상에 제 2 기판을 본딩하는 단계와, 제 1 기판을 제거해서 활성 핀을 노출시키는 단계를 포함한다.
- [0135] 예 24는 임의의 예와 조합될 수 있으며, 비실리콘 기반 기록 회로의 층 상에 제 2 기판을 본딩하는 단계는, 제 2 기판과 비실리콘 기반 기록 회로의 층의 대향하는 표면에 본딩 재료를 도포하는 단계를 포함하고, 본딩 재료는, 이산화 실리콘(SiO₂), 산질화 실리콘(SiON), 탄소 도핑된 산질화 실리콘(SiOCN), 실리콘 탄소 질화물(SiCN), 또는 산탄화 실리콘(SiOC) 중 하나 이상을 포함한다.
- [0136] 예 25는 임의의 예와 조합될 수 있으며, 비실리콘 기반 기록 회로의 층 상에 제 2 기판을 본딩하는 단계는, 이산화 실리콘(SiO₂), 산질화 실리콘(SiON), 탄소 도핑된 산질화 실리콘(SiOCN), 실리콘 탄소 질화물(SiCN), 또는 산탄화 실리콘(SiOC) 중 하나 이상을 사용해서 본딩 재료를 도포하는 단계와, 대향하는 표면의 본딩 재료를 서로 가압하는 단계를 포함한다.
- [0137] 예 26은 임의의 예와 조합될 수 있으며, 유전체 영역 상에 기록 회로의 층을 형성하는 단계는, 유전체 영역의 표면에 기록 회로의 층을 본딩하는 단계를 포함한다.
- [0138] 예 27은 임의의 예와 조합될 수 있으며, 제 2 기판에 회로를 형성하는 단계를 더 포함한다.
- [0139] 예 28은 임의의 예와 조합될 수 있으며, 시스템을 포함하고, 이득 셀 임베디드 동적 랜덤 액세스 메모리(eDRAM) 메모리 셀을 포함하는 메모리 디바이스 - 이득 셀 eDRAM 메모리 셀은, 스토리지 셀 위에 적어도 부분적으로 형성된 기록 회로와, 스토리지 셀 아래에 적어도 부분적으로 형성된 판독 회로를 포함함 - 와, 메모리 디바이스에 본딩된 컴퓨터 니어 메모리 디바이스와, 프로세서와, 메모리 디바이스로부터 프로세서로의 인터페이스를 포함한다.
- [0140] 예 29는 임의의 예와 조합될 수 있으며, 컴퓨터 니어 메모리 디바이스는, 벡터-벡터 곱셈, 곱셈 및 누산(MAC), AND, OR, XOR, NOT, 신경망 활성화 연산, soft-max의 역승, 데이터 압축, 암호화, 미디어 가속기, 문자열 매칭 또는 정규 표현식 가속기, 또는 포인터 추적 중 하나 이상을 수행하는 것이다.
- [0141] 예 30은 임의의 예와 조합될 수 있으며, 메모리 디바이스의 출력을 제공해서 SRAM 메모리 디바이스의 출력 판독 레이트를 에플레이트하는 에플레이션 회로를 포함한다.
- [0142] 예 31은 임의의 예와 조합될 수 있으며, 에플레이션 회로는 컨트롤러, 멀티플렉서, 또는 레지스터 중 하나 이상을 포함한다.
- [0143] 예 32는 임의의 예와 조합될 수 있으며, 컨트롤러는, 메모리 디바이스의 다수의 뱅크로의 메모리 액세스 요청의 전달을 제어하고, 멀티플렉서는, 컨트롤러에 의한 선택에 따라서, 메모리 디바이스로부터 컴퓨터 니어 메모리로 데이터를 송신하거나 혹은 멀티플렉서는 컴퓨터 니어 메모리로부터 데이터를 수신해서 메모리 디바이스에 제공한다.
- [0144] 예 33은 임의의 예와 조합될 수 있으며, 레지스터는, 메모리 디바이스로부터 출력되는 데이터를, 컴퓨터 니어 메모리 디바이스에 출력하기 이전에, 버퍼링한다.
- [0145] 예 34는 임의의 예와 조합될 수 있으며, 컴퓨터 니어 메모리 디바이스는, 이산화 실리콘(SiO₂), 산질화 실리콘(SiON), 탄소 도핑된 산질화 실리콘(SiOCN), 실리콘 탄소 질화물(SiCN), 또는 산탄화 실리콘(SiOC) 중 하나 이상을 포함하는 본딩 재료를 사용해서 메모리 디바이스에 본딩된다.

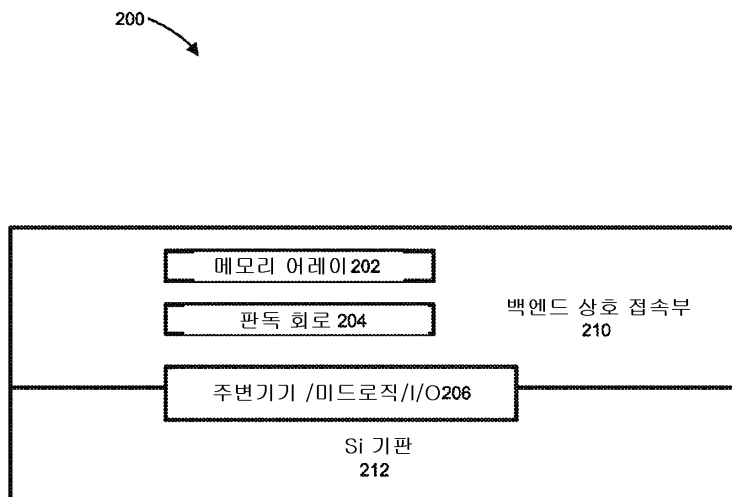
- [0146] 예 35는 임의의 예와 조합될 수 있으며, 프로세서에 통신 가능하게 연결된 네트워크 인터페이스를 포함한다.
- [0147] 예 36은 임의의 예와 조합될 수 있으며, 시스템을 포함하되, 적어도 하나의 프로세서와, 적어도 하나의 프로세서에 연결된 적어도 하나의 메모리를 포함하되, 적어도 하나의 메모리는, 이득 셀 임베디드 동적 랜덤 액세스 메모리(eDRAM) 디바이스를 포함하며, 이득 셀 eDRAM 디바이스는, 기록 회로와, 스토리지 디바이스와, 판독 회로를 포함하며, 판독 회로는 적어도 부분적으로 기록 회로 아래에 기록 회로의 평면과는 다른 평면에 위치되고, 스토리지 디바이스는 기록 회로를 판독 회로에 연결시키며, 기록 회로와 판독 회로 중 하나 이상이 스토리지 디바이스에 본딩된다.
- [0148] 예 37은 임의의 예와 조합될 수 있으며, 스토리지 디바이스는, 적어도 하나의 커패시터 및 적어도 하나의 커패시터와 접촉하는 유전체를 포함한다.
- [0149] 예 38은 임의의 예와 조합될 수 있으며, 기록 회로는, 폴리-Si, Si, Ge, 폴리-Ge, III-V, GaN, MoS2, WSe2, MoSe2, WSe2, InS, HfS, ZnS, ZnSe, In2O3, ZnO, AZO, IGZO 또는 IZO 중 하나 이상을 포함한다.
- [0150] 예 39는 임의의 예와 조합될 수 있으며, 판독 회로는, 본딩 재료를 사용해서 스토리지 디바이스에 본딩되고, 본딩 재료는, 이산화 실리콘(SiO2), 산질화 실리콘(SiON), 탄소 도핑된 산질화 실리콘(SiOCN), 실리콘 탄소 질화물(SiCN), 또는 산탄화 실리콘(SiOC) 중 하나 이상을 포함한다.

도면

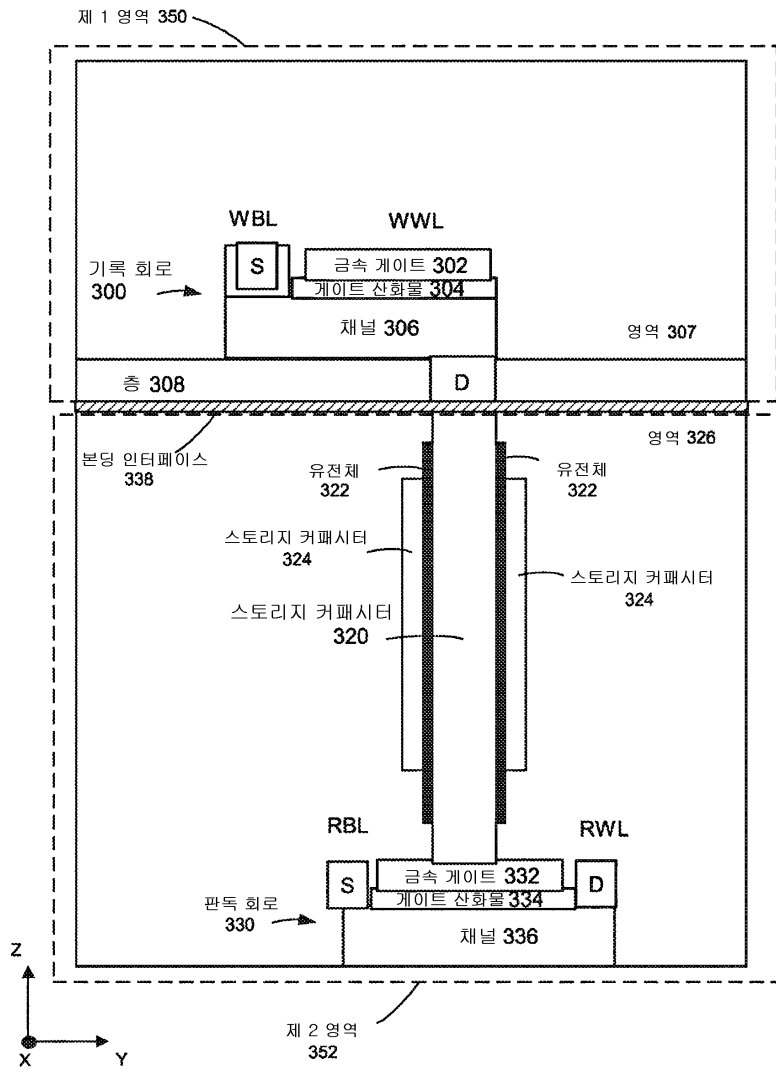
도면1



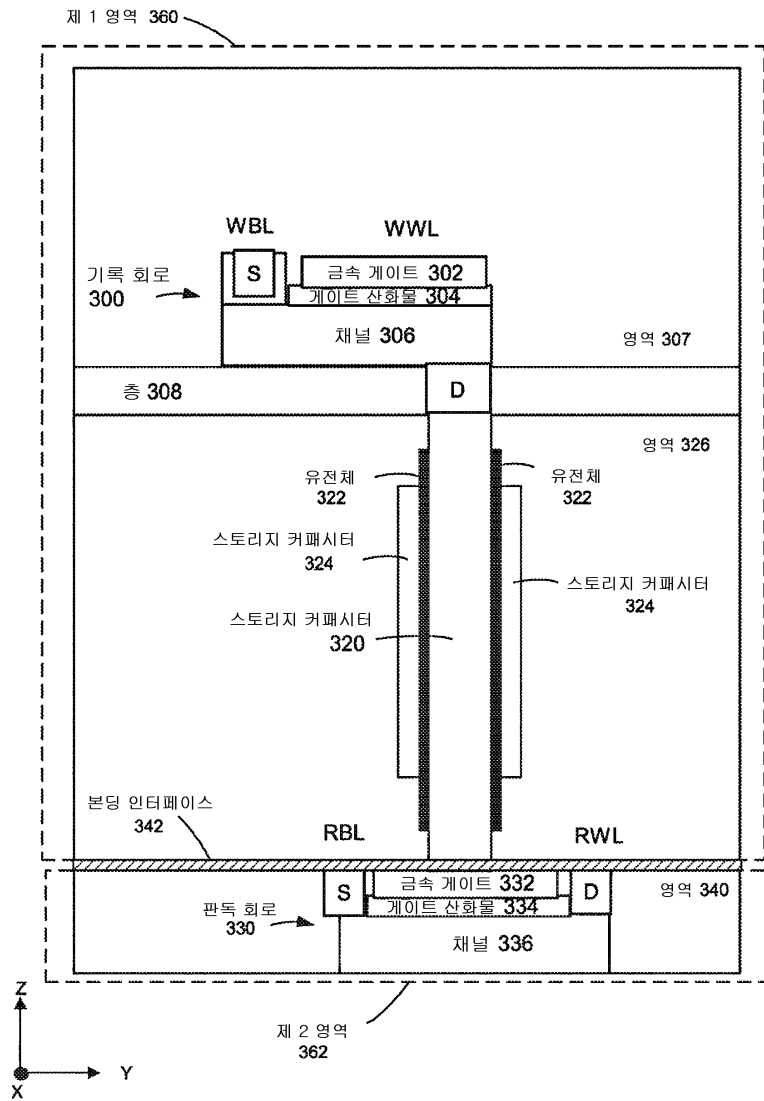
도면2



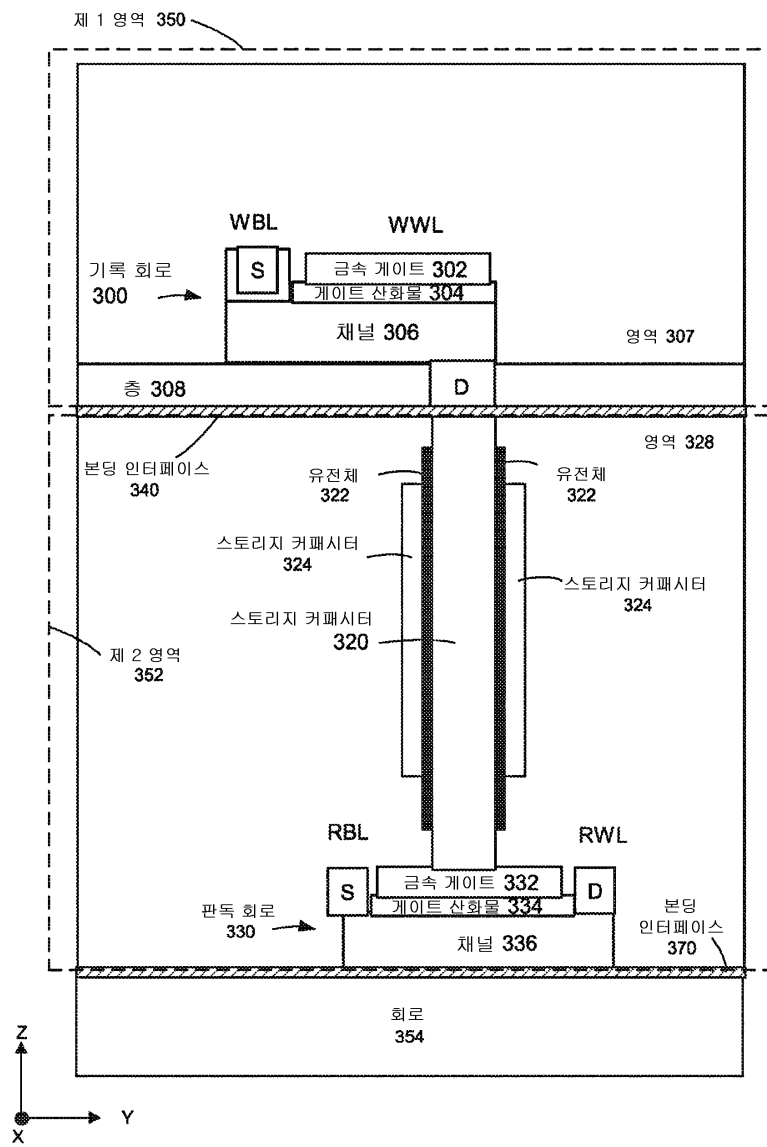
도면3a



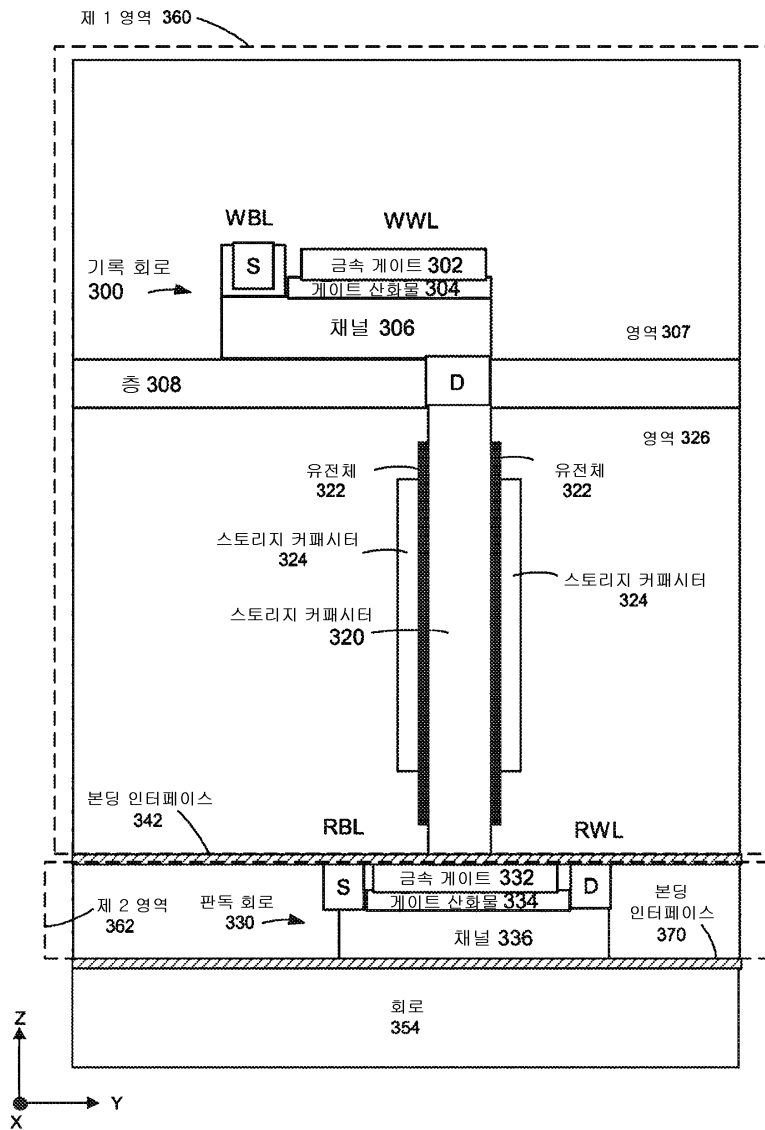
도면 3b



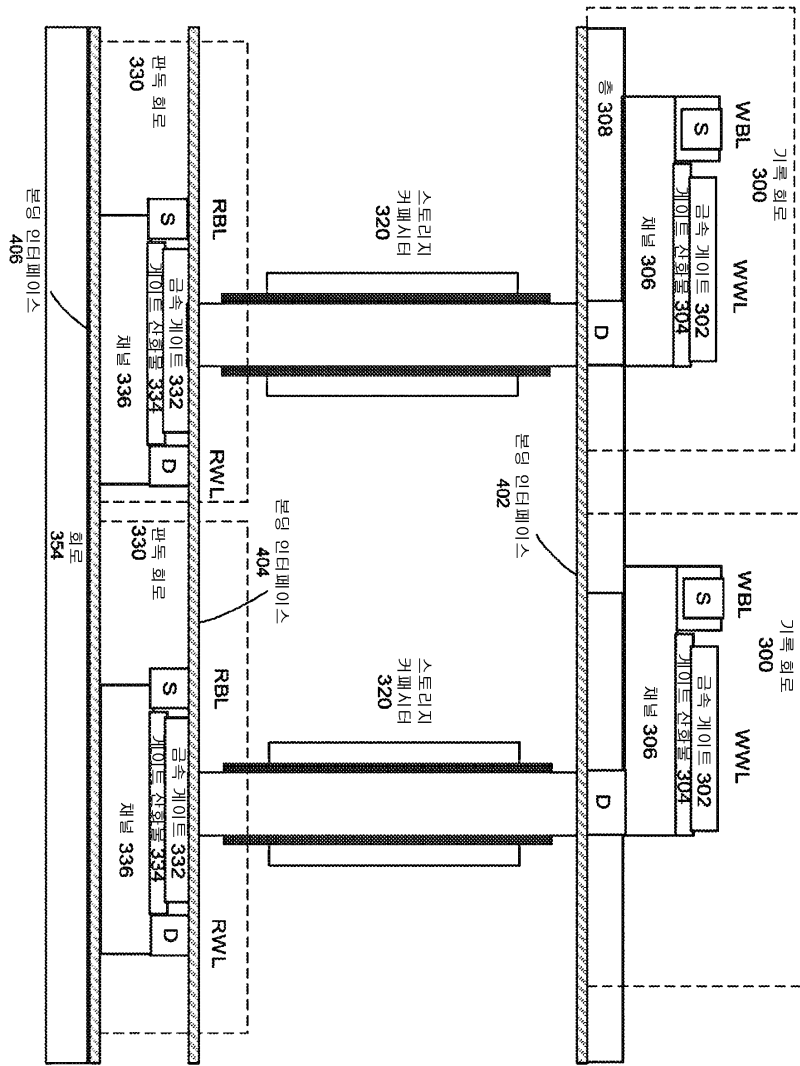
도면3c



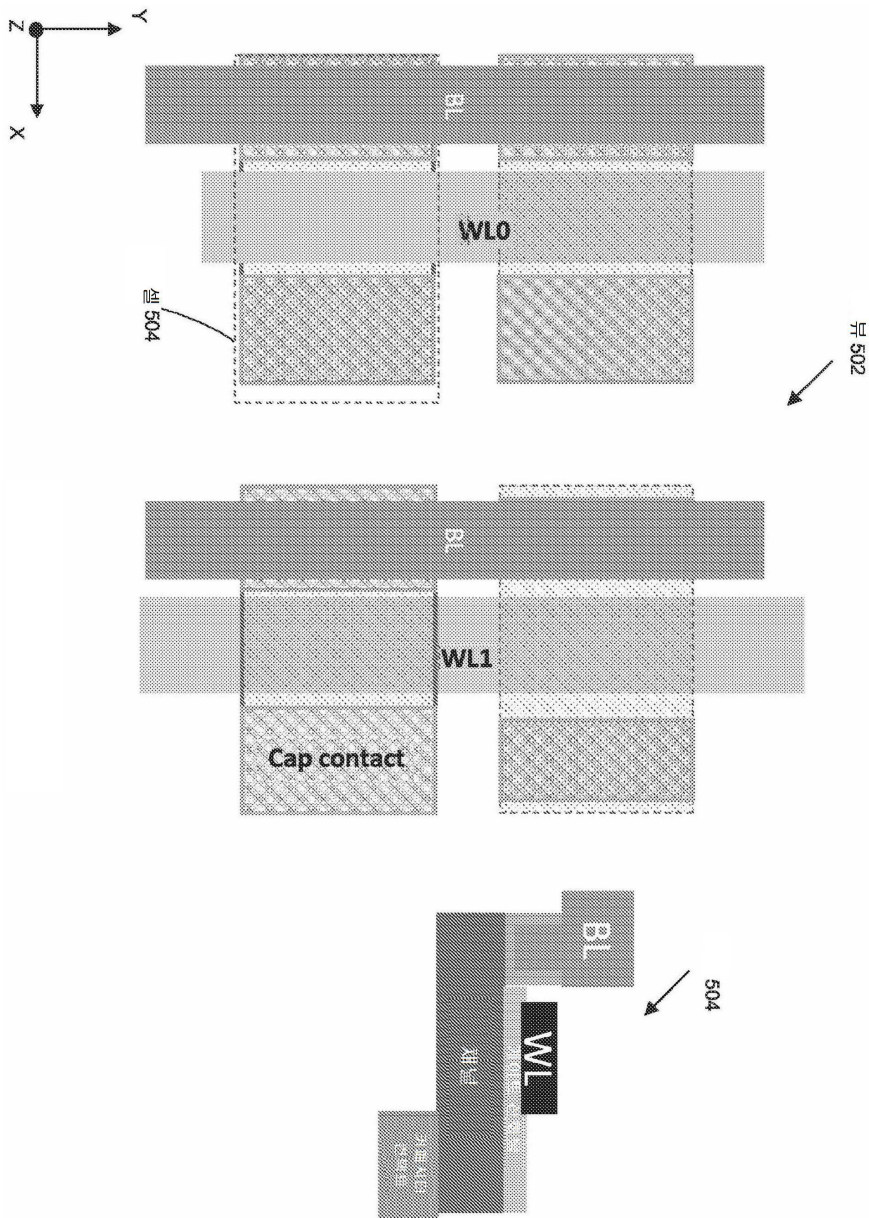
도면3d



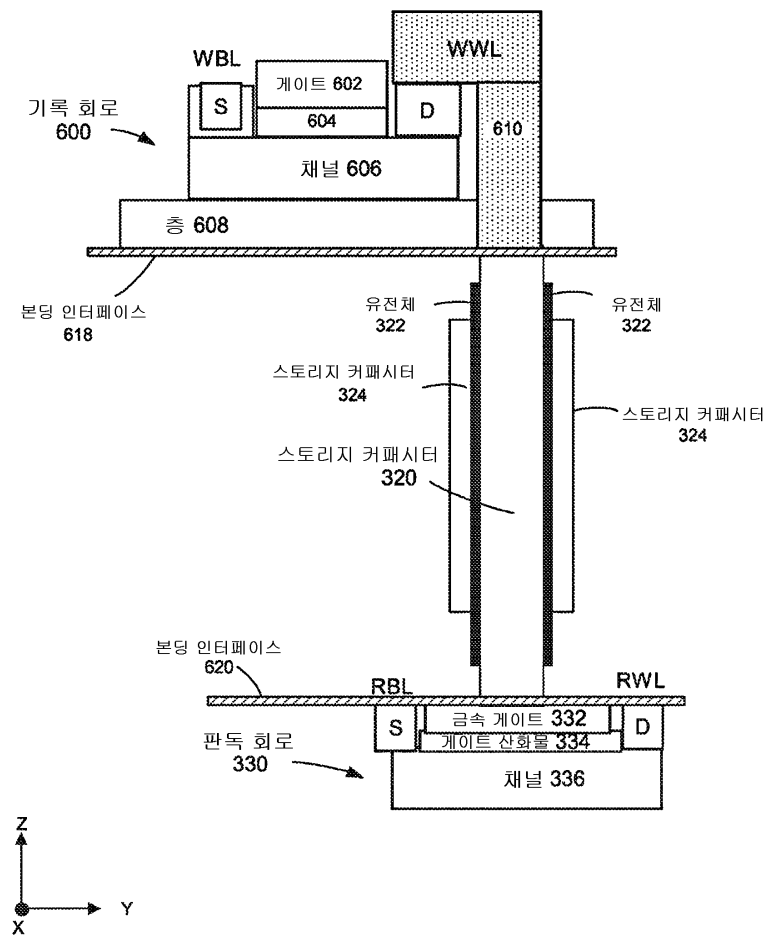
도면4



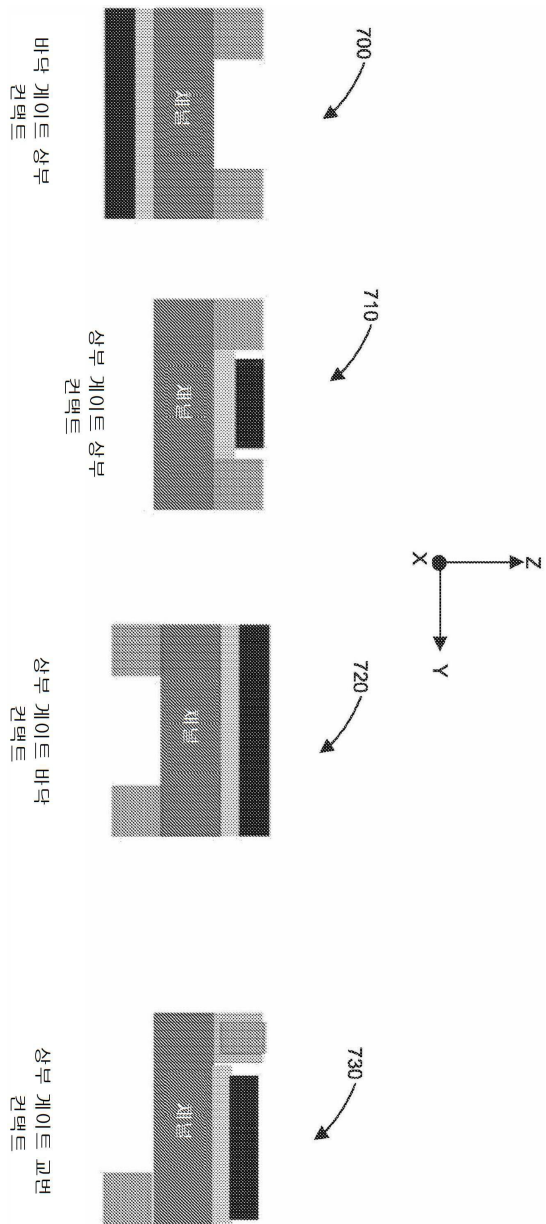
도면5



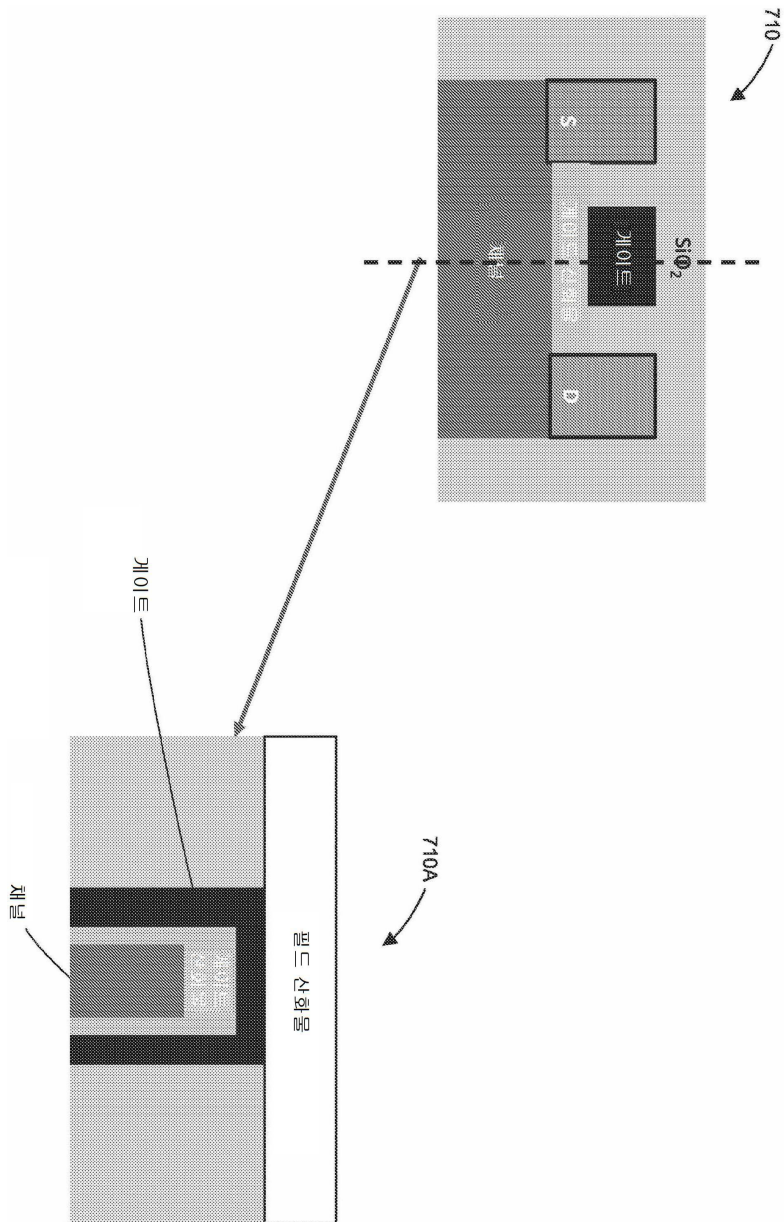
도면6



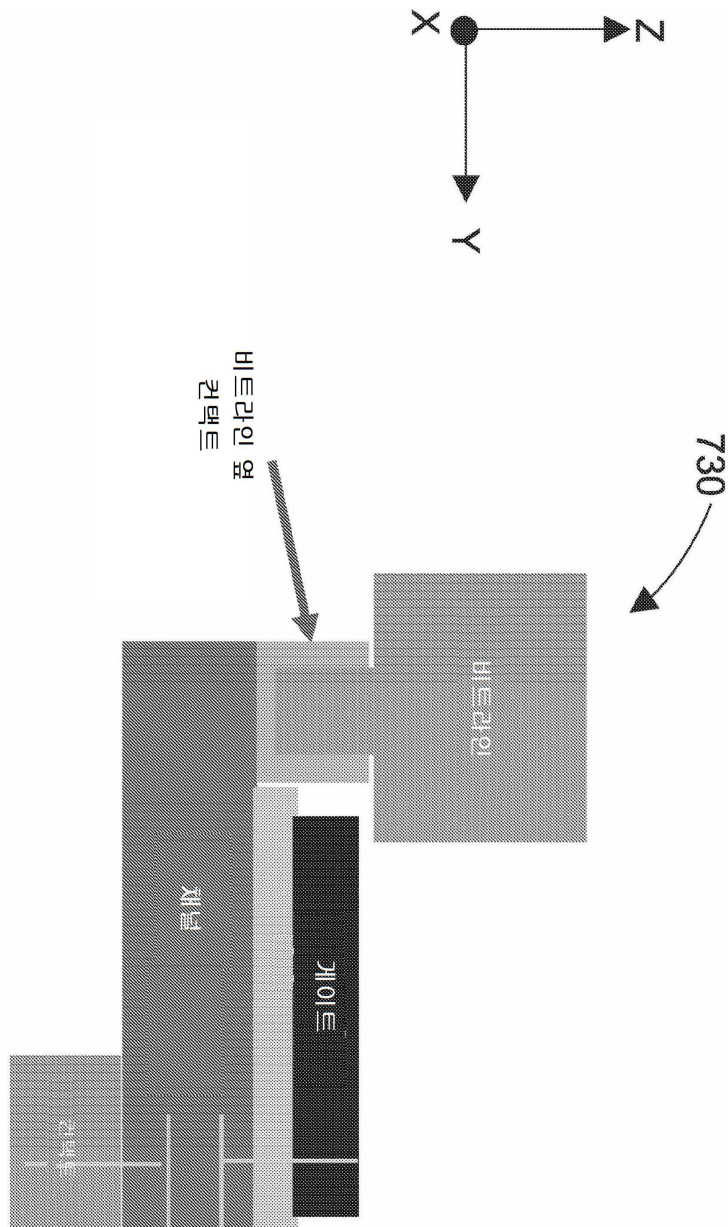
도면7a



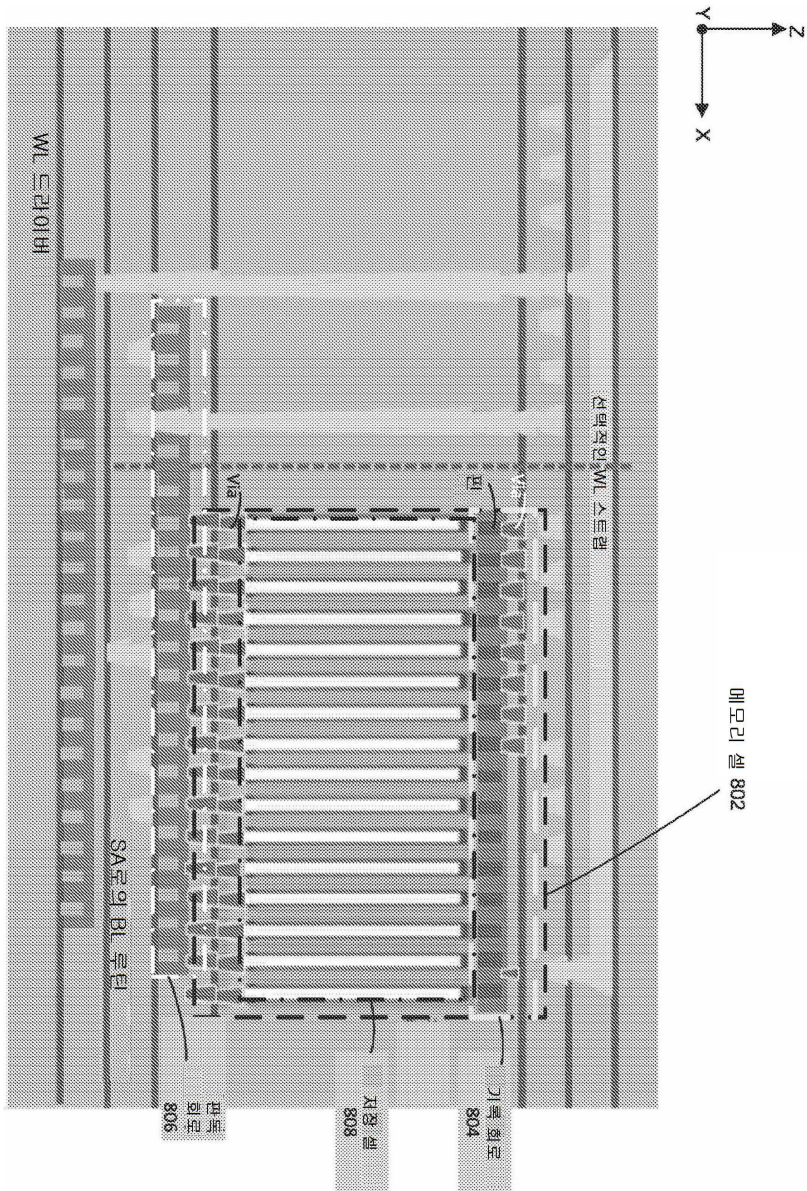
도면7b



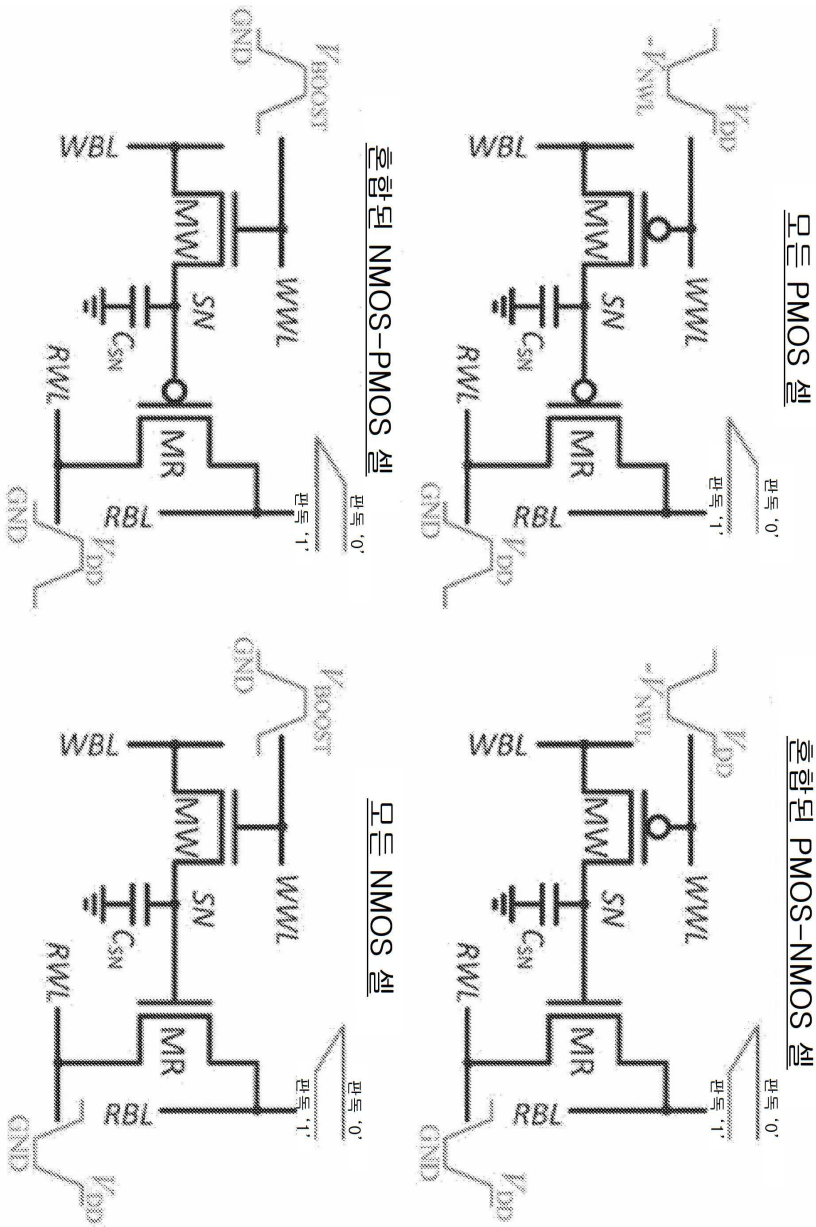
도면7c



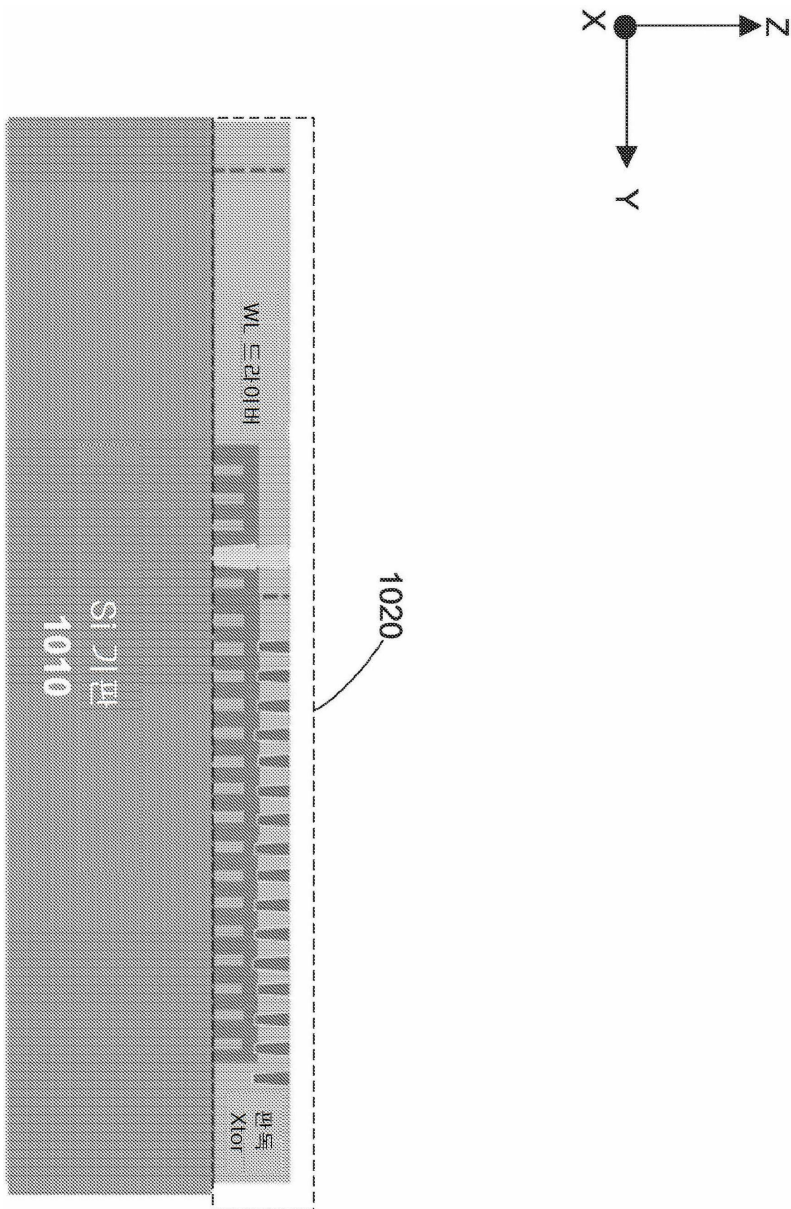
도면8



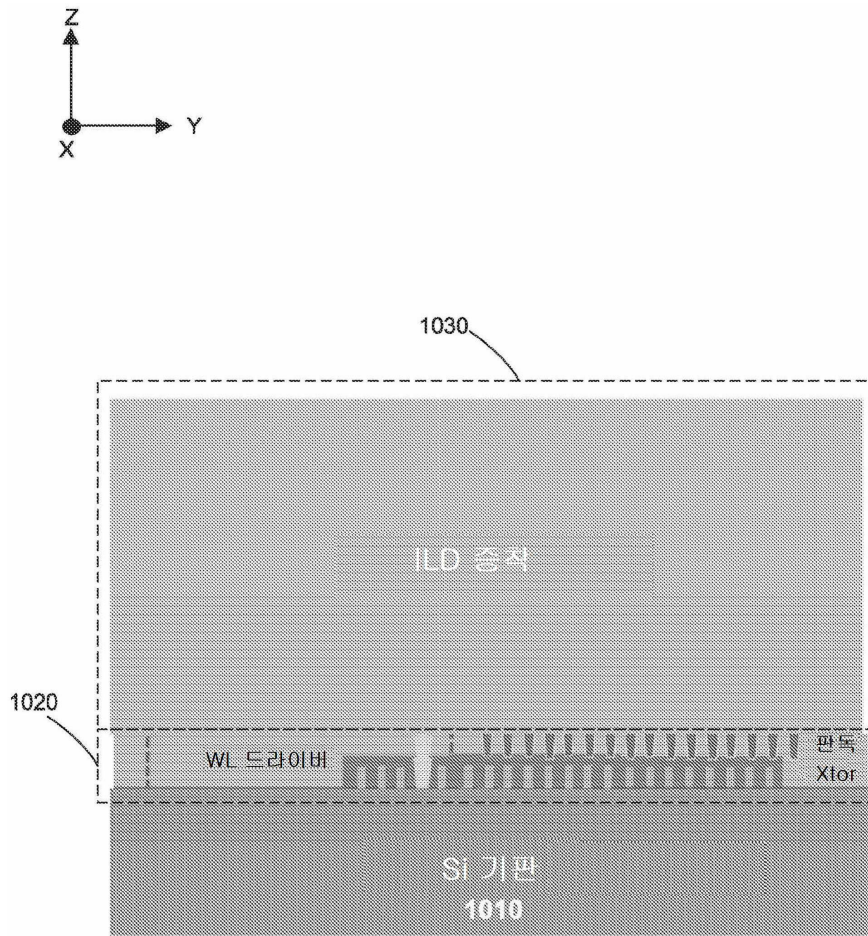
도면9



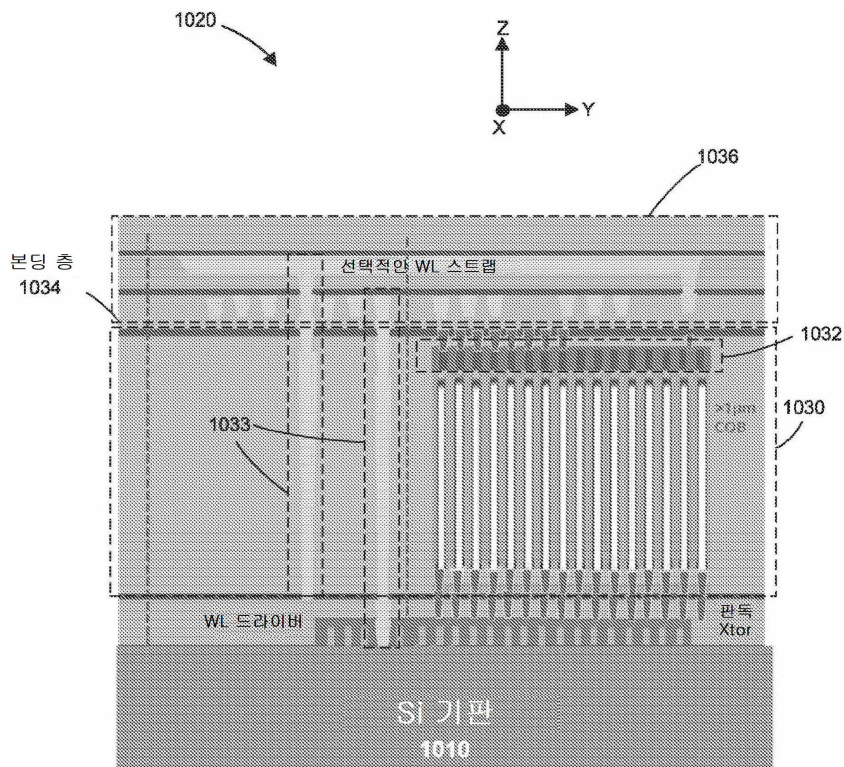
도면10a



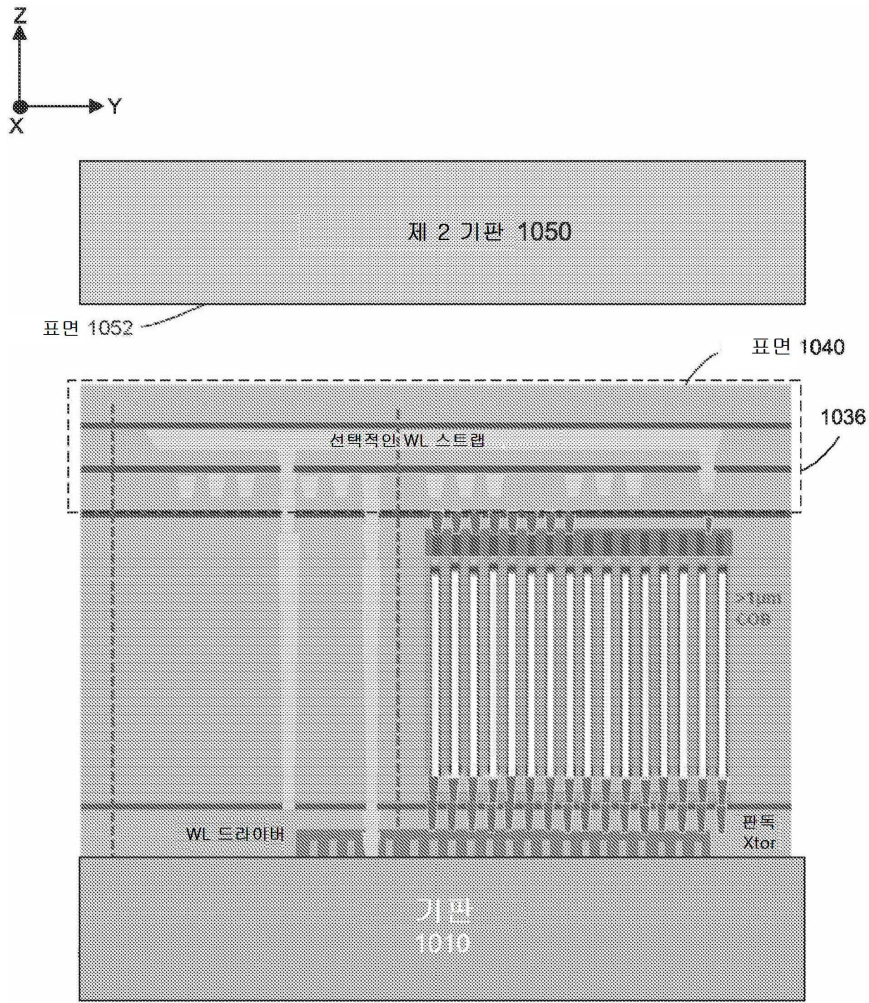
도면10b



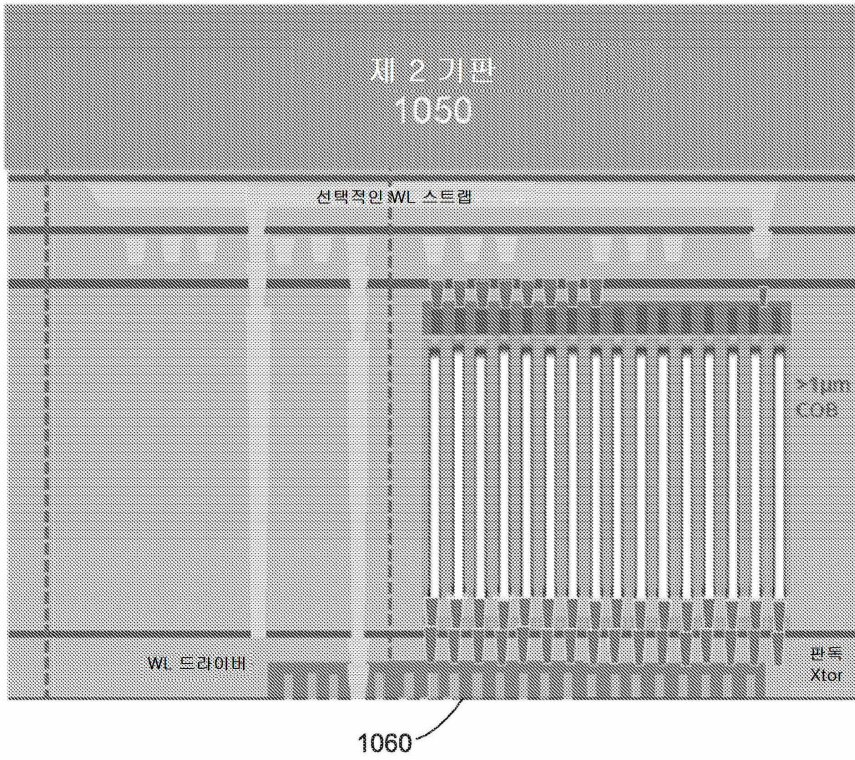
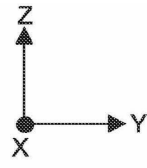
도면10c



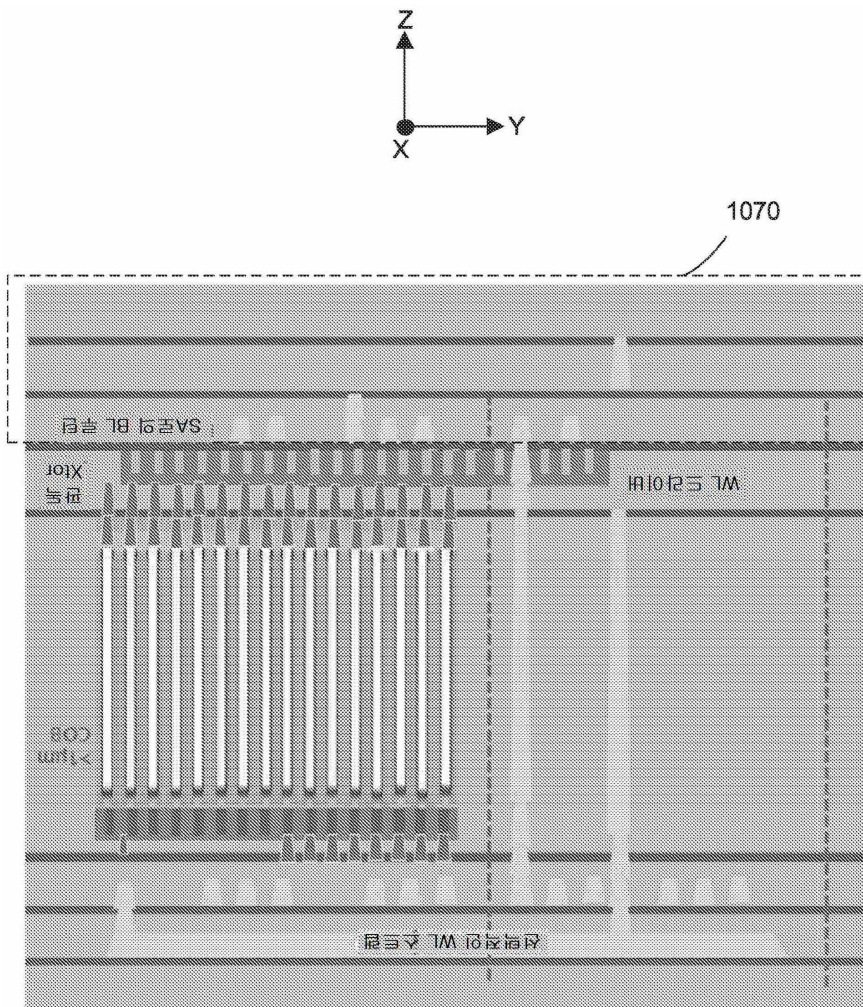
도면10d



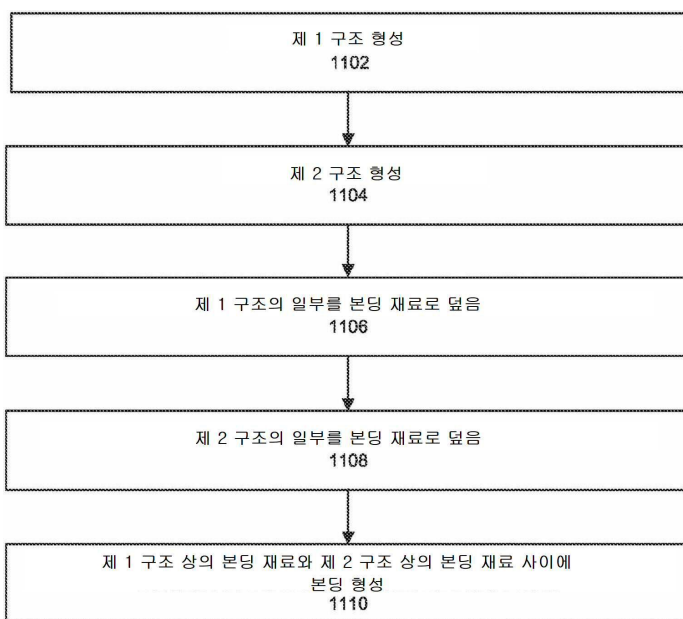
도면10e



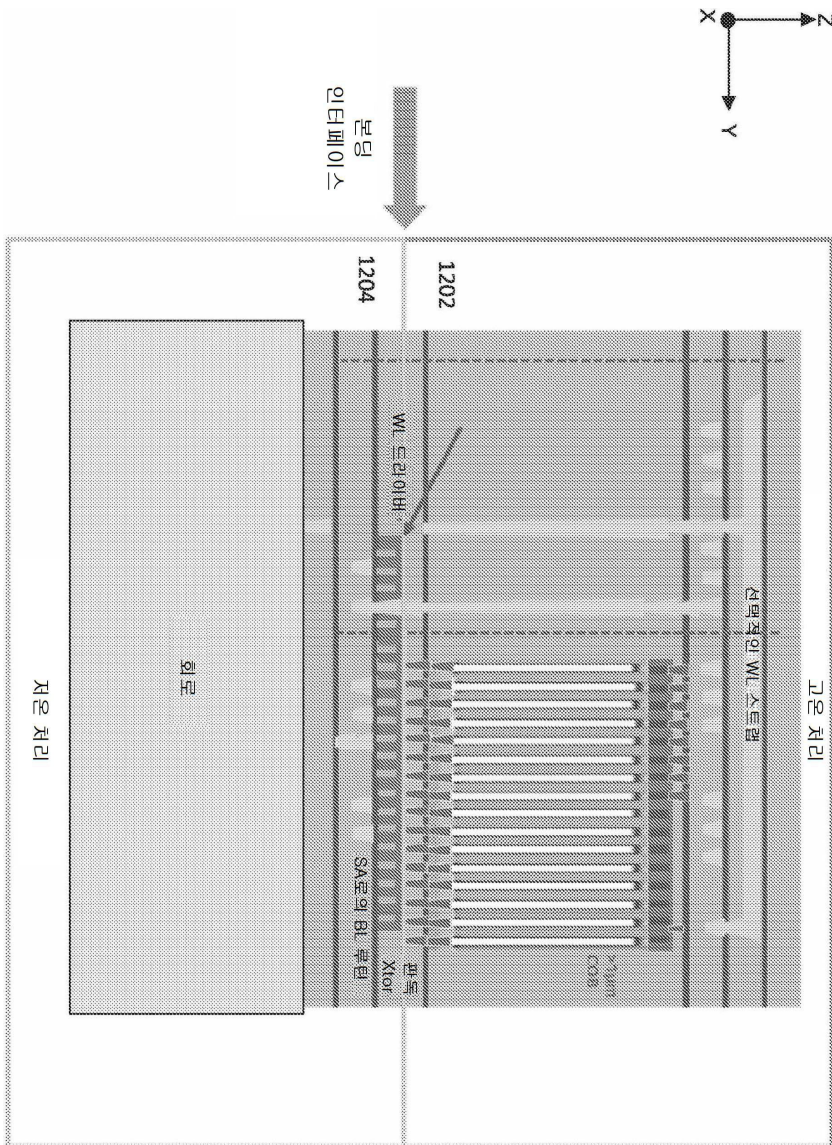
도면10f



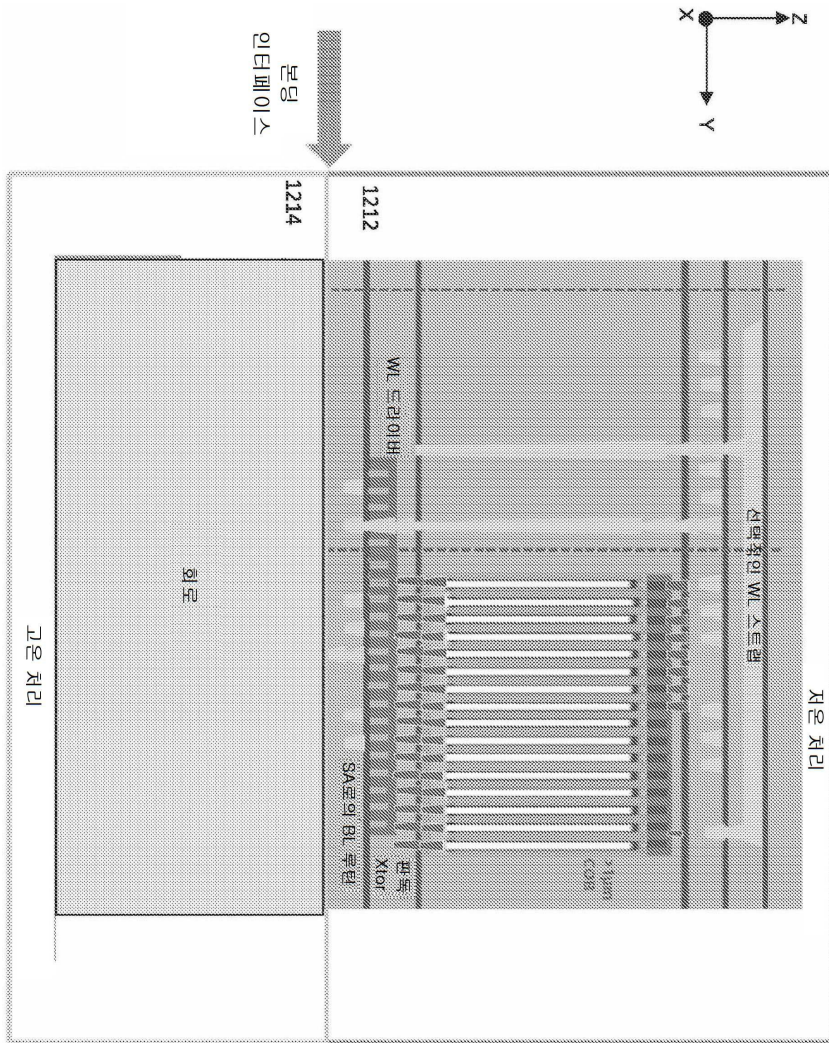
도면11



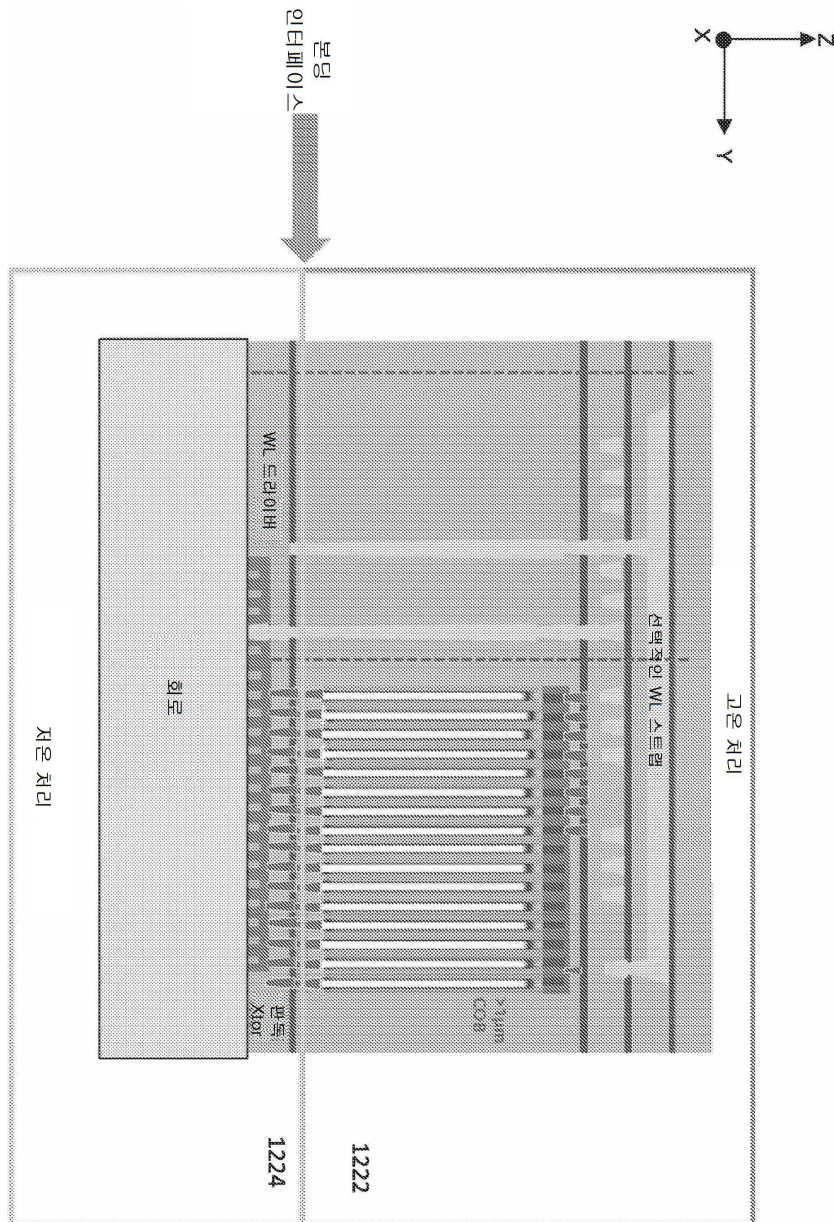
도면12a



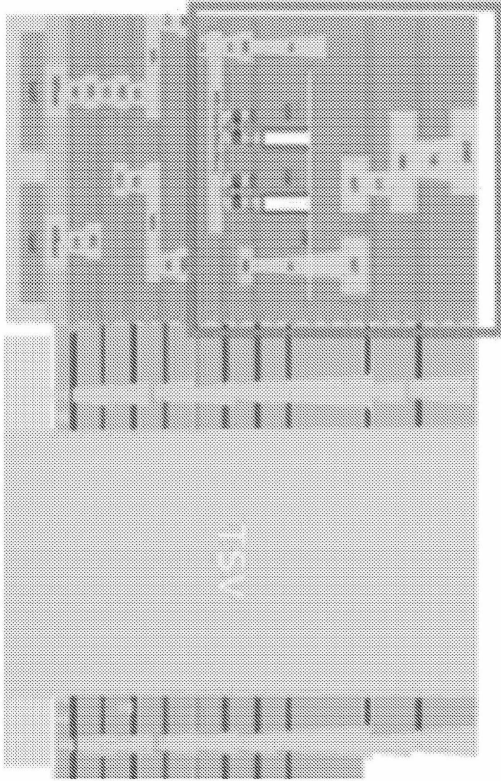
도면12b



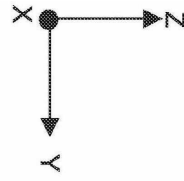
도면12c



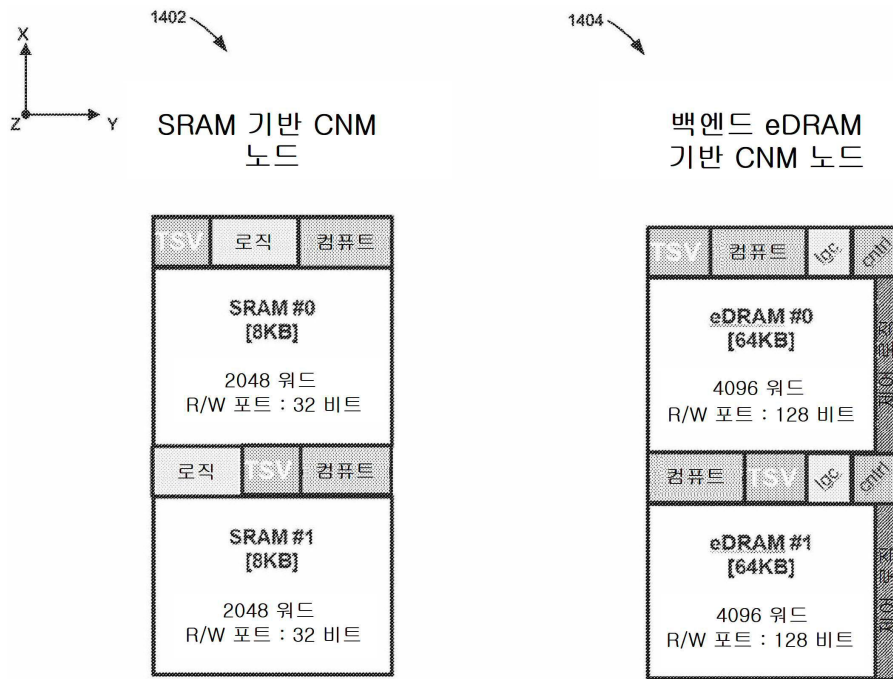
도면13



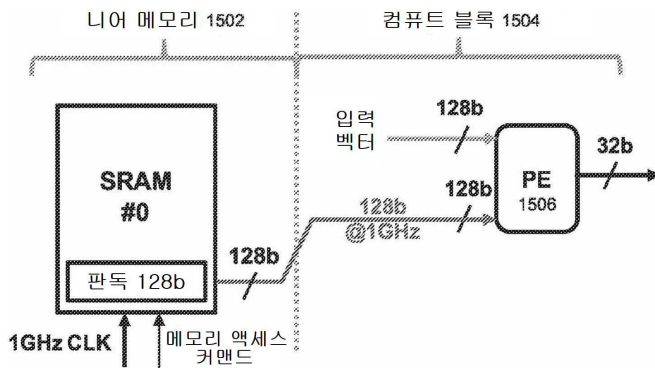
이러한 구성을 갖는



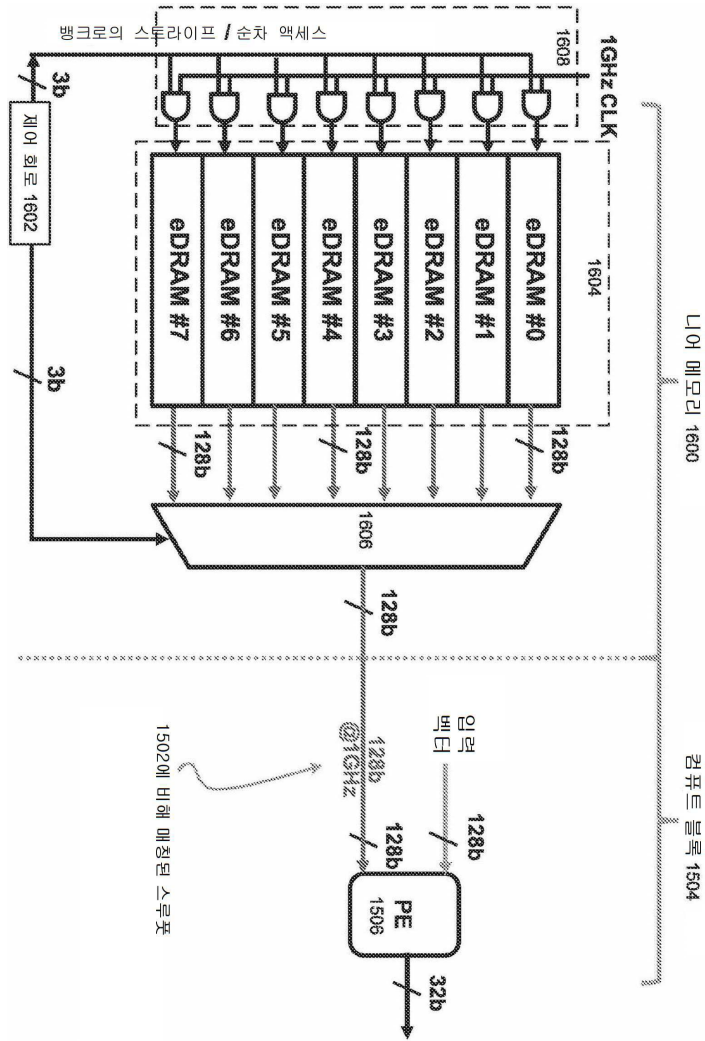
도면14



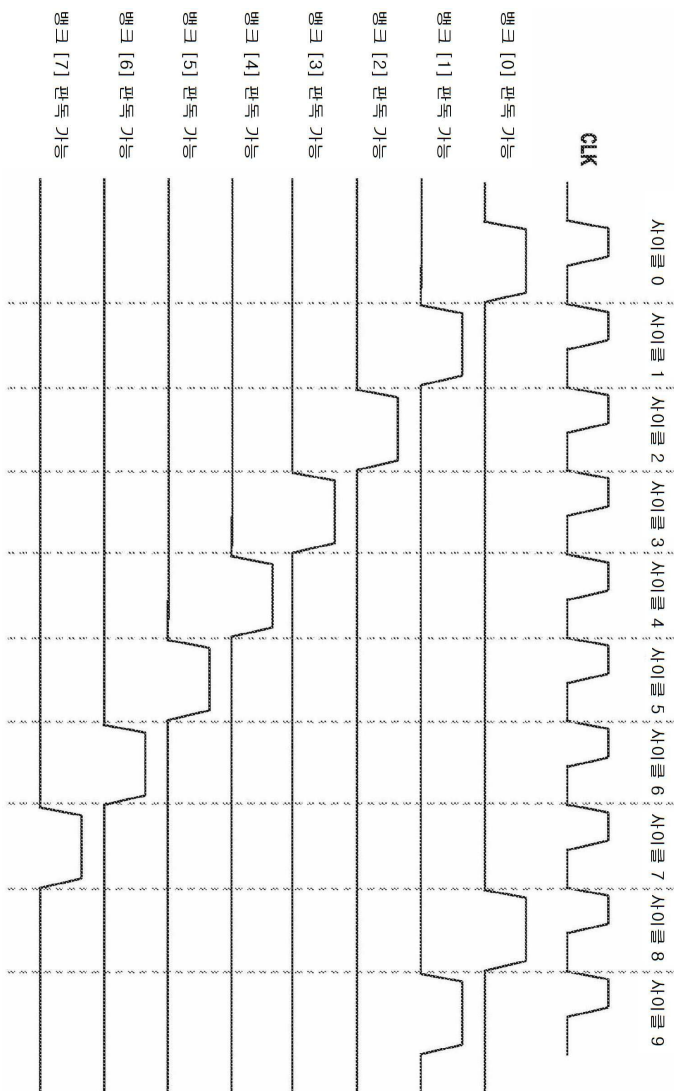
도면15



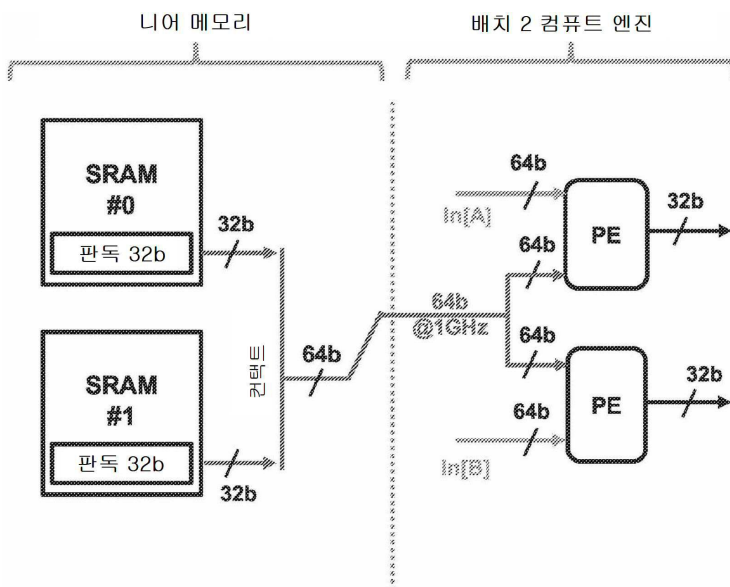
도면16



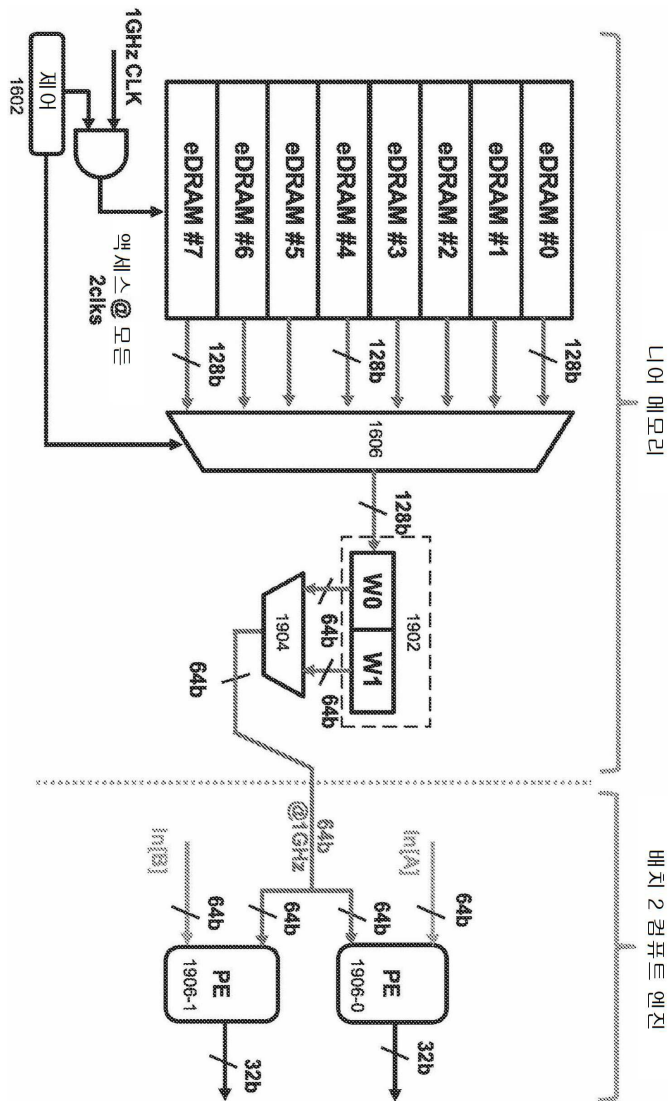
도면17



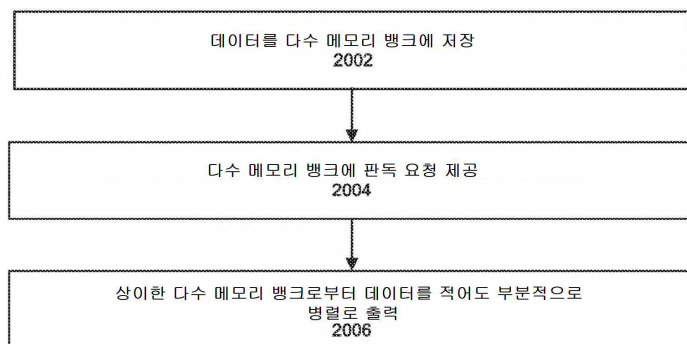
도면18



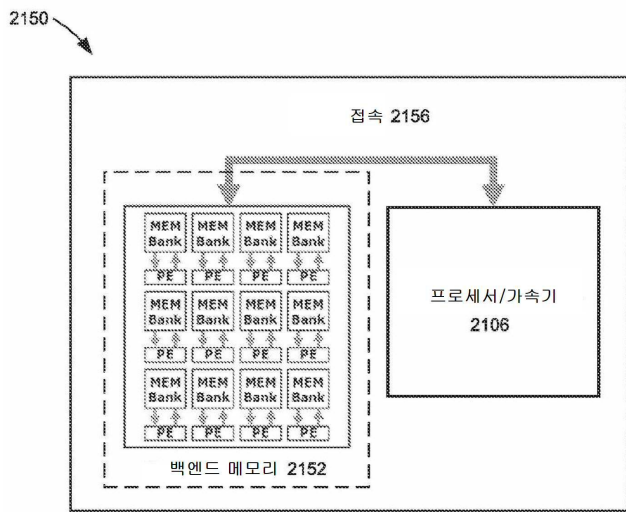
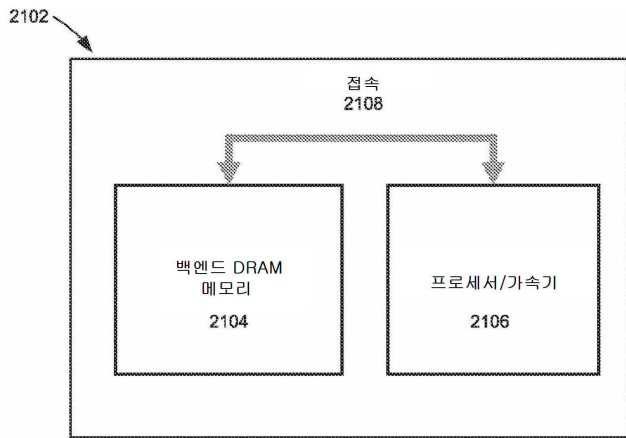
도면19



도면20

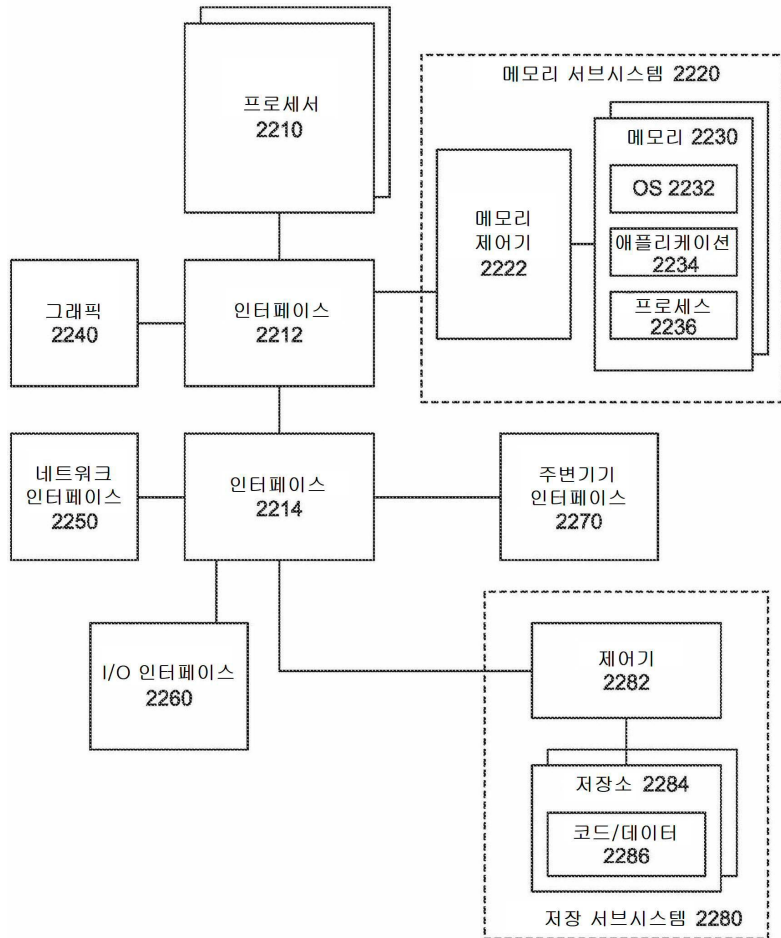


도면21



도면22

2200



도면23

