

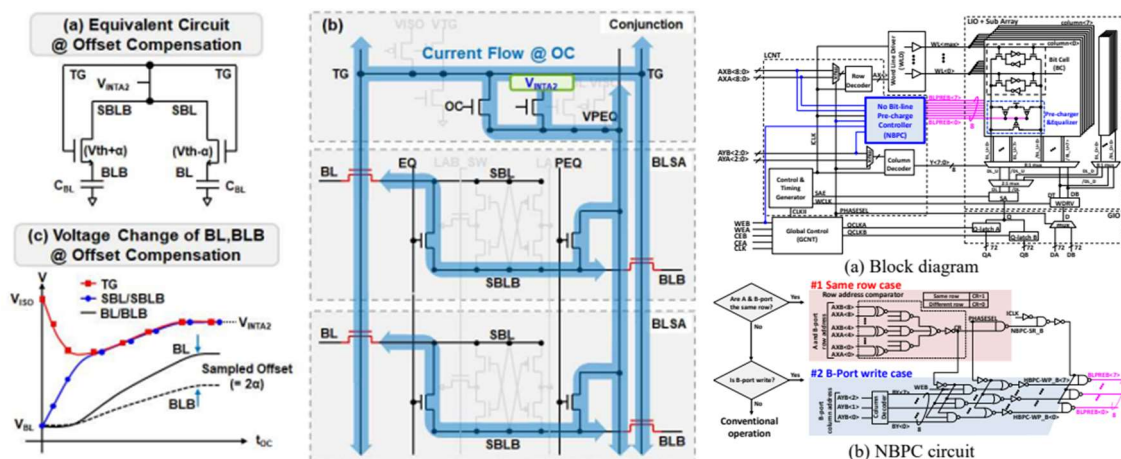
2024 IEEE VLSI Review

서울대학교 전기정보공학부 박사과정 박현준

Session 16 Memory Circuits

Session 16은 최신 반도체기술을 기반으로 한 다양한 고성능 메모리 설계 및 회로 최적화 기법들이 소개되었다. 이 세션에서는 SRAM, MRAM 등 다양한 메모리 기술이 고속 동작, 저전력 소모, 그리고 높은 신뢰성을 확보하기 위한 방법론을 제시하며, 특히 Automotive grade 1 및 HPC 애플리케이션을 목표로 한 논문들이 소개되었다.

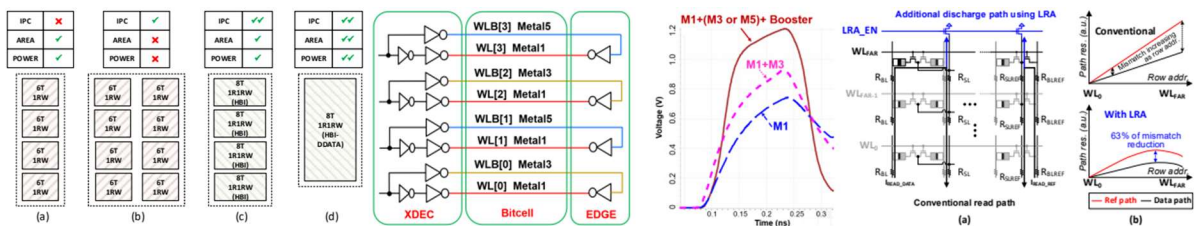
#16-1 이 논문은 저전압 DRAM을 위한 OC-CTPS BLSA를 제안한다. 제안된 비트 라인 센스 앰프(BLSA)는 14nm DRAM 공정으로 구현되었으며, 오프셋 보정을 통해 트랜지스터의 V_{th} 불일치를 보정함으로써 충전 전달(CT)의 정확성을 높였다. 기존 BLSA 방식에서 공정 변동으로 인한 V_{th} 오프셋은 충전 전달의 불안정성을 유발해 불량 비율(FBR)을 증가시키는 문제가 있었으나, OC-CTPS 방식은 이를 다이오드 연결을 통해 해결하였다. 이는 추가적인 회로 크기 증가 없이 BLSA의 성능을 개선할 수 있게 해준다. 제안된 OC-CTPS BLSA는 2.75ns의 오프셋 보정 시간(t_{OC})과 1.5ns의 충전 전달 시간(t_{CT})을 통해 안정적인 감지 기능을 제공한다. 실험 결과, -25°C 와 100°C 에서 각각 250ps와 500ps의 3시그마 창을 확보하였으며, 이는 다양한 온도 조건에서도 전압 임계값 변화를 안정적으로 추적할 수 있음을 보여준다. 또한, 0.75V의 낮은 동작 전압에서도 기존 BLSA보다 94% 향상된 불량 비율을 기록하며, 저전압 환경에서의 안정성을 증명하였다. 결론적으로, 이 논문에서 제안된 OC-CTPS BLSA는 저전력, 저전압 환경에서 DRAM의 신뢰성을 크게 향상시킬 수 있는 효과적인 솔루션을 제공한다.



[그림 2] (좌) 16-1 (우) 16-2

#16-2 이 논문은 3nm Fin-FET 기술을 기반으로 한 19.87 Mbit/mm²의 비트 밀도를 달성한 2RW(2-읽기/쓰기) 의사 듀얼 포트(PDP) 6T SRAM 설계를 제안한다. 이 설계는 성능을

개선하기 위해 고저항 와이어 추적을 도입하여 다양한 PVT(공정, 전압, 온도) 조건에서 최적의 속도를 유지하고, 순차적 접근을 고려한 동적 전력 감소 기법을 통해 불필요한 비트라인 프리차지를 제거함으로써 동적 전력을 최대 43%까지 절감한다. 이 SRAM은 288-kbit PDP 매크로로 제작되었으며, 2RW 듀얼 포트 SRAM과 비교하여 면적 효율성을 극대화하였다. 6비트 DAC 아키텍처를 사용하여 데이터를 처리하며, 비트라인 프리차지를 제어하는 새로운 회로 구조를 통해 전력 소모를 크게 줄일 수 있었다. 특히, 동일 행(row)을 순차적으로 접근할 때 1차 프리차지를 생략하여 전력 절감을 실현했다. 또한, 고저항 와이어가 포함된 회로에서 발생할 수 있는 타이밍 문제를 해결하기 위해 비트라인 프리차지 타이밍 최적화를 도입하였다. 이를 통해 다양한 PVT 조건에서도 안정적인 동작을 보장하며, SRAM의 성능과 전력 효율성을 향상시켰다. 제안된 SRAM은 43%의 비트라인 전력 절감을 달성했으며, 1.64 GHz에서 동작할 수 있다. 최종적으로, 제안된 3nm PDP SRAM은 경쟁 기술에 비해 더 높은 비트 밀도와 전력 효율성을 제공하여 차세대 SoC 및 FPGA 설계에 적합한 솔루션을 제시한다.



[그림 2] (좌) 16-3 (중) 16-4 (우) 16-5

#16-3 이 논문은 3nm 기술에서 7GHz 이상의 성능을 제공하는 1R-1RW(1 읽기-1 쓰기/쓰기) 고대역폭 SRAM을 Arm HPC(High Performance Computing) 프로세서에 통합한 설계를 소개한다. 기존 8T-1R1W 메모리 아키텍처를 개선하여 추가적인 읽기 포트를 제공함으로써 1R1RW 기능을 구현하였으며, 이를 통해 L1-데이터 캐시에서 읽기 대역폭을 두 배로 증가시키고 Instruction Per Cycle을 1% 이상 향상시켰다. 또한, 이 새로운 HBI(High Bandwidth Instance) 메모리 아키텍처는 CPU의 물리적 설계에서 라우팅 혼잡과 지연을 줄여 13%의 면적 감소와 10-15ps의 라우팅 지연 감소를 달성하였다. 제안된 HBI-1R1RW 메모리는 기존 6T-1RW 메모리와 비교했을 때 33% 더 작은 L1-데이터 캐시 면적을 구현했으며, 두 개의 읽기 포트를 통해 읽기 성능을 개선하였다. 이 아키텍처는 word line 길이를 절반으로 줄이고, 두 개의 메모리 뱅크를 동시에 접근할 수 있는 구조를 사용하여 메모리 접근 시간을 최적화했다. 또한, RC 불일치 문제를 해결하기 위해 이중 word line 기술을 사용하여 두 번째 포트에서의 성능을 개선하였으며, 고속 다이내믹 멀티플렉서와 센싱 회로를 적용하여 첫 번째 포트에서도 빠른 읽기 속도를 보장하였다. HBI-Ddata 메모리는 네 개의 HBI-1R1RW 뱅크를 통합하여 CPU 설계에서 핀 접근성을 개선하고 라우팅 혼잡을 완화하였다. 이를 통해 CPU 성능 향상과 함께 전력 효율성을

극대화하였다. 또한, 백엔드 금속을 사용하여 내부 메모리 라우팅의 저항과 결합 커패시턴스를 줄임으로써 메모리 성능을 최적화하였다. 3nm 공정으로 제작된 테스트 칩에서는 7GHz 이상의 주파수와 11.2 Mbit/mm²의 8T 기반 SRAM 비트 밀도를 기록하였으며, 0.945V에서 108ps의 빠른 메모리 접근 시간(0.75V, TT에서 85°C 기준)을 달성하였다. 최종적으로, HBI-1R1RW 메모리는 Arm HPC 프로세서의 L1-데이터 캐시에서 사용되며, 높은 성능과 효율성을 제공하여 CPU의 전반적인 주파수를 향상시킬 수 있음을 입증하였다.

#16-4 이 논문은 3nm FinFET 기술을 사용하여 1024x640 및 2048x640 크기의 단일 포트 고속 멀티뱅크 SRAM 매크로를 설계한 내용을 다룬다. 이 설계는 고성능 컴퓨팅(HPC) 애플리케이션을 대상으로 하며, single-port 6T 비트셀을 기반으로 하여 최고 3.3GHz에서 작동하는 SRAM 매크로를 구현하였다. 또한, 0.55V-1.35V의 넓은 전압 범위에서 동작이 가능하여 DVFS와 같은 전력 효율성 개선 기술을 지원한다. 제안된 SRAM 설계에서는 워드라인, 글로벌 클록, 글로벌 비트라인 부스팅 기술을 통해 속도를 개선하고, 디코더 신호용 스플릿 드라이버를 도입하여 속도를 37% 향상시켰다. 또한, 읽기 보조 회로를 통해 최대 전압(V_{max})을 1.35V로 높여 성능을 최적화했다. 이 모든 기술을 적용한 결과, 1V 100°C에서 3.3GHz의 성능을 달성하였다. 이 설계는 멀티뱅크 구조로 구현되었으며, 이를 통해 워드라인과 X-Dir 신호 전송의 효율성을 높였다. 부스팅 회로와 라우팅 최적화를 통해 리피터 추가 없이 신호 전송 속도를 개선하였다. 특히, Y방향에서는 워드라인과 비트라인 신호를 개선하기 위해 보조 인버터와 글로벌 신호 부스터를 사용했으며, X방향에서는 글로벌 클록(GCK) 지연과 X 프리디코드 신호를 최적화했다. 결과적으로, 이 논문에서 제시한 회로 기술은 전력 효율성을 개선하고, SRAM의 면적 효율성을 극대화하였다. 이를 통해 기존의 멀티뱅크 SRAM 설계보다 더 큰 매크로 크기를 구현하면서도 밀도×주파수/전력 지표에서 최고의 성능을 달성했다.

#16-5 이 논문은 14nm FinFET 공정에서 구현된 128Mb eMRAM 설계를 제안하며, 고온에서의 읽기 마진 저하 문제를 해결하기 위한 회로 기술을 소개한다. LRA를 통해 데이터와 참조 읽기 경로 사이의 저항 불일치를 최소화하고, 온도 추적 회로를 사용하여 참조 저항을 적응적으로 변경하여 고온에서도 신뢰할 수 있는 읽기 동작을 보장한다. 이 설계를 적용한 256-IO MRAM 매크로는 17.88Mb/mm²의 고밀도를 달성하였으며, 0.60V 코어 전압에서 -40°C에서 160°C까지의 온도 범위에서 11ns의 읽기 속도를 성공적으로 구현하였다. STT-MRAM(스핀 전이 토크 MRAM)**은 높은 내구성, 유지력, 빠르고 저전력 쓰기 성능 등의 장점이 있지만, 터널 자기 저항 비율(TMR)이 고온에서 저하되는 문제가 있다. 이는 데이터 '0'과 '1' 사이의 저항 차이가 감소하여 읽기 마진이 줄어드는 문제를 초래한다. 이 논문에서는 이러한 문제를 해결하기 위해 오프셀 누설 전류와 저항 불일치를 줄이고, 고온에서도 신뢰할 수 있는 읽기 성능을 유지하는 방법을 제안한다. LRA 데이터와 참조 경로 간의 저항 불일치를 완화하여 읽기 속도와 전류를 증가시키며, TTC는 온도

변화에 따라 참조 저항을 적응적으로 조정하여 TMR 저하와 저항 불일치를 보상한다. 또한, NWL 생성기를 도입하여 선택되지 않은 셀에서 발생하는 누설 전류를 제거하고 접근 트랜지스터의 신뢰성을 개선하였다. 128Mb MRAM 매크로는 7.16mm²의 면적을 차지하며, 이전 연구에 비해 40% 향상된 읽기 마진을 기록하였다.

저자정보



박현준 박사과정 대학원생

- 소속 : 서울대학교
- 연구분야 : HBM, Chord Signaling, Information Theory
- 이메일 : spp098@snu.ac.kr
- 홈페이지 : <https://sites.google.com/view/wschoi?pli=1>