

2025 IEEE VLSI Review

고려대학교 전기전자공학과 박사과정 한창우

Session 4 SRAM and Mask ROM

이번 VLSI 2025의 Session 4에서는 고성능 및 저전력 요구를 동시에 만족시키는 차세대 SRAM 및 Mask ROM 기술을 중심으로 총 4편의 논문이 발표되었다. 공정·설계·구조적 차원에서의 최적화를 통해 SRAM의 속도, 면적, 전력, 안정성을 향상시키기 위한 다양한 접근이 제시되었으며, 특히 저전압 환경, 전력 효율성, 신뢰성에 이슈를 두고 소자와 회로의 경계를 넘나드는 기술들이 주목을 받았다.

Session 4의 발표 논문 중, SRAM을 핵심으로 다룬 2편의 논문에서는 각각 공정 단순화와 신뢰성 확보, 캐시 메모리 환경에서의 동작 최적화에 초점을 맞춘 최신 기술을 제안하였다. 본 리뷰에서는 이 두 편의 논문을 중심으로 기술적 내용을 살펴보고자 한다.

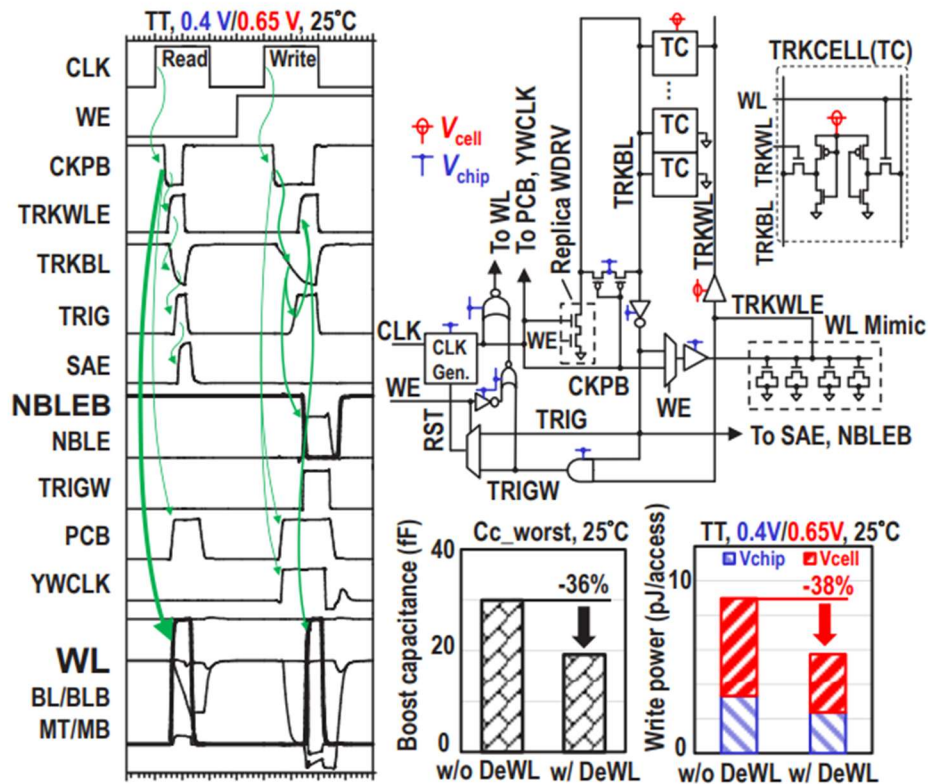
#4-1 본 논문은 TSMC에서 발표한 연구로, 3nm FinFET 공정 기반의 고밀도 단일 포트 SRAM 설계를 통해 에너지 효율과 스탠바이 누설 전류 저감을 동시에 달성한 내용을 담고 있다. 제안된 설계는 총 563kbit 용량에 35.5 Mbit/mm²의 높은 비트 밀도를 가지며, 3.89 pJ/access의 낮은 에너지 소모와 27.5 μ W/Mbit 수준의 저전력 모드(1-CLM)를 실현했다. 핵심 기법으로는 다음 두 가지가 있다.

첫째, eXtended Dual Rail(XDR) 아키텍처는 메모리 셀과 일부 회로만 고전압(V_{cell})으로 구동하고 나머지는 저전압으로 처리하여 레벨 쉬프터 없이도 동작하며 전력 소모와 면적을 동시에 줄였다.

둘째, Delaying Write Word Line(DeWL) 기법은 NBL 방식에서 발생하는 셀-드라이버 충돌을 억제해 쓰기 전력 38% 절감을 달성했다.

또한 1-Cycle Latency Low-Leakage Mode(1-CLM)을 도입해 standby 중 비트라인을 차단하여 누설 전류를 약 10% 줄였으며, 순차적 pre-charge 회로를 통해 피크 전류도 504mA에서 93mA로 완화하였다.

종합적으로 본 논문은 초저전력 동작, 고집적, 고성능을 모두 만족하는 차세대 모바일용 SRAM 설계의 방향성을 제시하며, 실리콘 기반의 정량적 데이터로 그 효과를 입증했다.



[그림 1] Delaying Write Word Line (DeWL) 기법과 효과

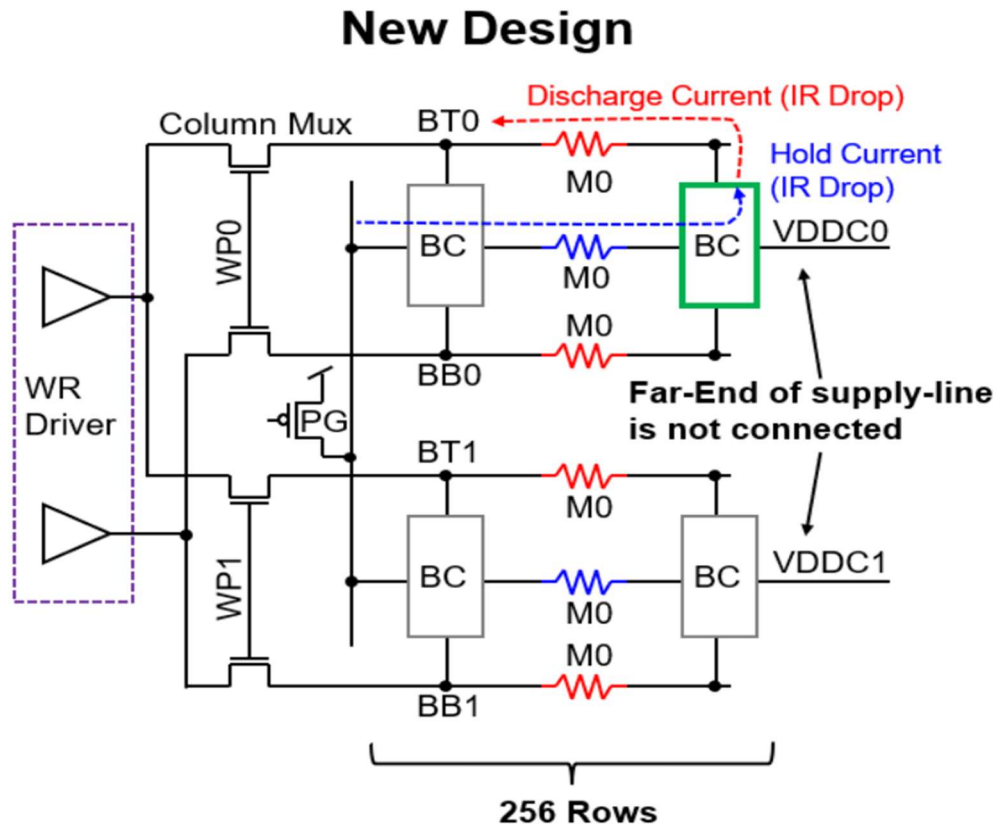
#4-4 본 논문은 Intel 18A 공정 기반으로 구현된 고밀도 SRAM 기반 Register File(RF) 구조를 제안하며, RibbonFET GAA 트랜지스터와 Back Side Power Delivery Network(BSPDN) 기술을 통해 최고 수준인 37.8 Mb/mm²의 밀도를 달성한 사례를 보고한다. 본 설계는 최대 262kb의 용량을 갖는 6T 셀 기반으로 구성되며, SoC에서 범용적으로 사용 가능한 SRAM 구조로 활용될 수 있다. 논문에서는 특히 두 가지 주요 기술이 도입되었다.

첫째, Per-Column Resistive Supply-Line(SL) 구조는 각 컬럼의 전원선(V_{DDC})을 독립적으로 구성하여 Bitline의 높은 저항이 쓰기 동작에 미치는 영향을 완화한다. 이 구조는 전통적인 파워 그리드 방식 대신 컬럼 단위로 IR drop을 유도함으로써, 쓰기 대상 셀의 Pull-Up 트랜지스터 게이트 전압을 감소시키고, 추가적인 write-assist 없이도 쓰기 동작이 성공하도록 한다.

둘째, Parallel Boost Injection 기반 Write-Assist(WA) 회로는 기존의 3단 스택 구조에서 발생하던 비효율을 제거하고, 2스택 구조를 기반으로 Boost Cap의 충전 경로를 간소화함으로써 write driver 면적을 35% 감소시킨다. 이를 통해 전체 WA 회로의 면적 오버헤드는 단 1.5% 수준으로 최소화되었다.

실리콘 측정 결과, 해당 구조는 WA 미사용 시에도 0.675 V까지 동작 가능하며, WA를 사용하면 최저 V_{DD} 0.5 V에서도 동작 가능하도록 약 175mV의 V_{MIN} 개선이 확인되었다. 이는 -40°C 환경에서도 유효했으며, 고온(125°C) 조건에서도 안정적인 동작이 입증되었다.

총평하자면, 본 논문은 고집적, 저전압, 고신뢰성을 동시에 만족하는 차세대 RF용 SRAM 설계 기술을 제시한 것으로, 특히 Flying-BL을 사용하지 않고도 최고 수준의 RF 셀 밀도를 달성했다는 점에서 기존 구조와 차별화된 가치를 지닌다



[그림 2] 본 논문에서 제안하는 Per-Column Resistive Supply-Line(SL) 구조

Session 22 OTP and Nonvolatile Memory

이번 VLSI 2025의 Session 22에서는 OTP(One-Time Programmable) 및 차세대 비휘발성 메모리(NVM) 기술을 중심으로 총 4편의 논문이 발표되었다. 각 논문은 소자의 미세화와 고집적화를 위한 3D 적층 구조, 고온 환경에서의 데이터 유지, 그리고 BEOL 공정 호환성을 확보하면서도 높은 신뢰성을 유지하는 구조 설계 등의 기술적 이슈를 다루고 있다.

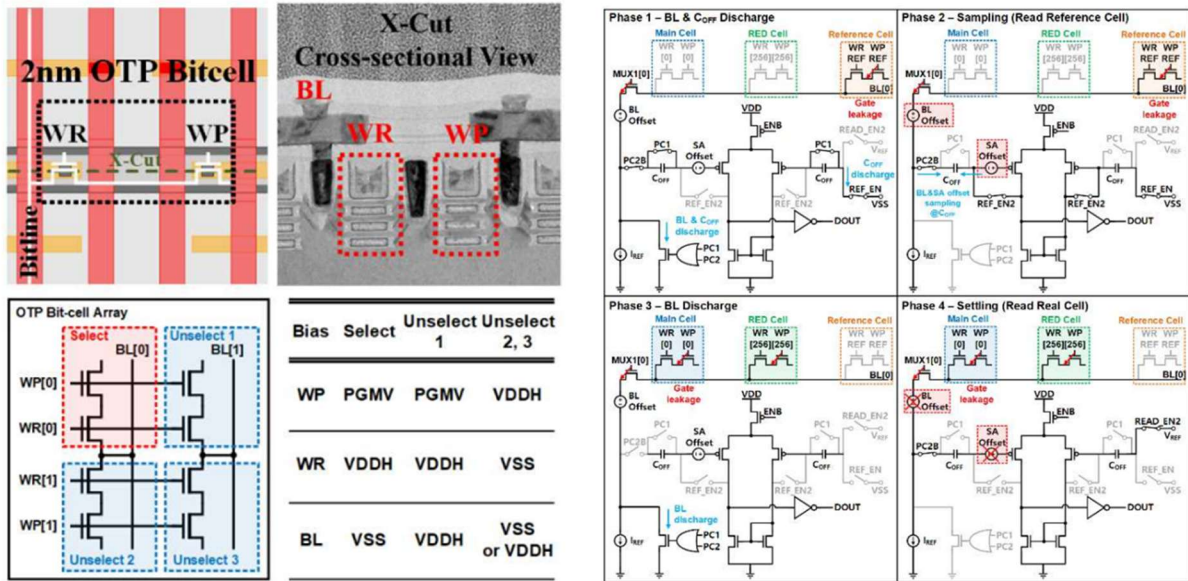
특히 저전압 동작, 열 안정성 확보, 공정 복잡도 감소와 같은 실질적 기술 장벽을 돌파하기 위한 소재 및 구조적 접근이 주목을 받았으며, OTP를 비롯한 다양한 NVM 기술의 scaling 한계 극복 및 logic-on-memory 구현 가능성을 강조하는 방향성이 돋보였다.

Session 22의 발표 논문 중, BEOL 공정과의 통합성, 3D 적층 기반 고속 OTP 구현 등 상이한 기술적 접근을 통해 미래형 OTP/NVM 구현 방향을 제시한 2편의 논문을 중심으로 기술적 내용을 살펴보고자 한다.

#22-1 본 논문은 삼성 파운드리에서 발표한 연구로, 2nm Gate-All-Around(GAA) 공정 기반의 128Kb Anti-Fuse One-Time Programmable(OTP) 메모리를 세계 최초로 구현하고, 초미세 공정에서 문제가 되는 읽기 마진 저하를 해결하기 위한 새로운 기법을 제안하였다.

OTP 메모리는 SoC에서 메모리 리페어, 보안, 아날로그 캘리브레이션 등에 반드시 요구되지만, 공정이 미세화되면서 게이트 산화막이 얇아져 게이트 누설 전류가 증가하고, 그 결과 비프로그램 셀 전류(ICELL-0)가 증가하며 프로그램 셀 전류(ICELL-1)는 감소하여 두 전류 간의 차이가 줄어드는 문제가 발생한다. 이는 Sense Amplifier(SA) 오프셋과 Bit-Line(BL) 오프셋의 영향을 더욱 크게 만들며, 결과적으로 안정적인 데이터 판별이 어려워진다.

이를 해결하기 위해 본 연구는 동적 BL 및 SA 오프셋 캔슬레이션(Dynamic BL & SA Offset Cancellation) 기법을 제안하였다. 추가적인 Reference Cell Array를 두어 BL 누설 패턴을 보상하고, Sense Amplifier 단계에서 오프셋을 함께 제거함으로써, 기존 방식 대비 낮은 IREF 환경에서도 안정적인 판별이 가능해졌다. 실험 결과, 제안된 기법은 Fail bit count를 크게 줄이며 데이터 무결성을 강화했고, Shmoo plot을 통해 BL 누설 전류 허용 범위가 넓어지고 전원 동작 마진이 확대됨을 확인하였다. 특히, VDD 0.42V, VDDH 0.8V까지 동작 가능하며, 접근 시간은 0.9V에서 33ns, 0.5V에서 66ns를 달성하였다. 또한 CMOS 로직 호환 2nm GAA 기반 OTP 메모리 매크로는 실리콘 검증에서 100% 수율을 보이며, 초미세 공정 환경에서도 안정적인 OTP 구현 가능성을 입증하였다. 본 연구는 향후 2nm 급 SoC에서 필수적인 OTP 메모리의 신뢰성과 저전압 동작을 동시에 보장할 수 있는 핵심 기술로 평가된다



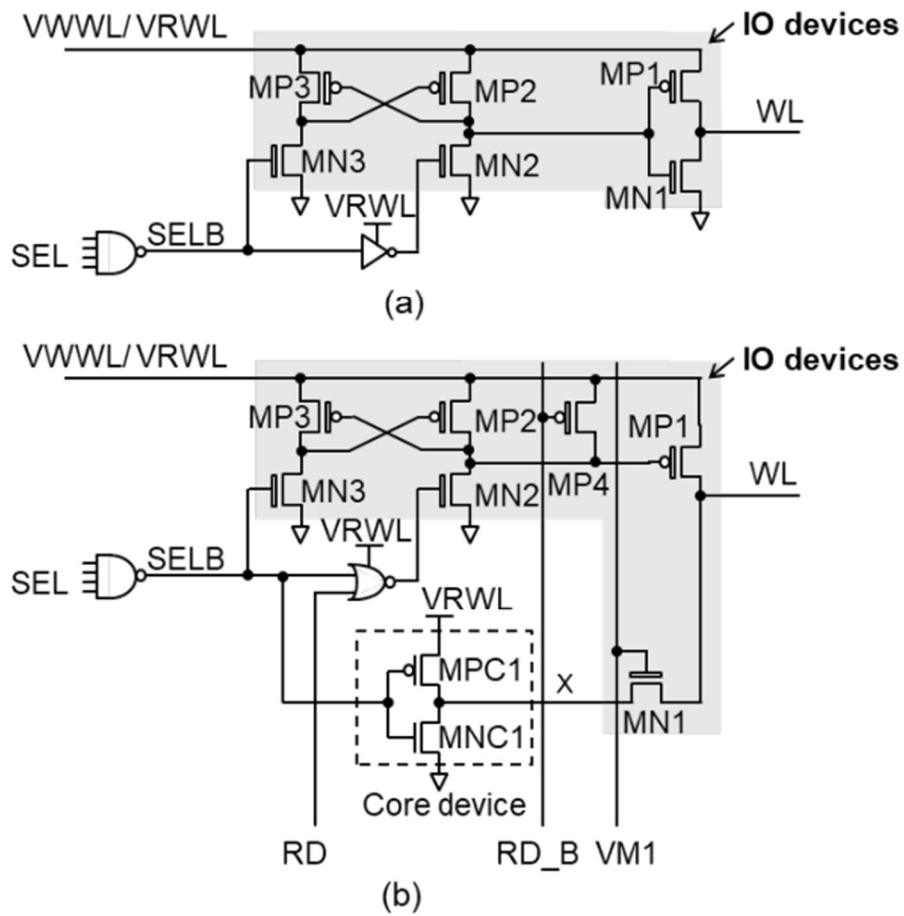
[그림 3] 2nm Anti-Fuse OTP 비트셀 구조 및 제안된 읽기 센싱 동작 원리

#22-4 본 논문은 TSMC에서 발표한 연구로, 55nm BCD 공정 기반의 16Mbit RRAM 매크로를 구현하여 자동차용 고온 환경(150°C)에서의 안정적 동작을 실현한 내용을 다룬다. RRAM은 단순 구조와 높은 저항 스위칭 비율, BEOL 호환성을 바탕으로 차세대 비휘발성 메모리로 주목받고 있으나, 자동차 응용에서는 150°C 이상의 고온에서 retention과 endurance 확보가 큰 과제이다. 이를 해결하기 위해 본 연구는 세 가지 주요 기술을 제안하였다.

첫째, 5V IO MOS와 1.2V Core MOS를 결합한 새로운 Wordline Driver(WLDRV) 구조를 도입하여, 기존 IO 기반 WLDRV 대비 읽기 속도 저하와 고전압 스트레스 문제를 해결하면서도 면적을 56.2% 절감하고, 매크로 전체 면적을 10.5% 줄였다.

둘째, Master/Local Current Limiter(LLVCL) 구조를 적용해 칩 전역의 V_{SS} 전위 기울기에 따른 쓰기 전류 불균일 문제를 억제함으로써 균일한 쓰기 특성을 확보하였다. 셋째, 온도 보상형 Local Write Driver를 도입하여, 고온에서 발생하는 Bit-Line 전압 강하와 쓰기 전류 감소 문제를 해결하였다. Bandgap Reference 기반의 V_{ptat} 전압을 활용해 온도가 높아질수록 보상 전압을 제공함으로써 150°C에서도 안정적인 쓰기 동작을 가능하게 했다.

실험 결과, 본 매크로는 2.56 Mbit/mm²의 집적도를 보였으며, 150°C 환경에서 1K endurance test 이후에도 충분한 읽기 윈도우를 유지하였고, 55MHz의 읽기 속도를 달성하였다. 결론적으로, 본 연구는 면적 효율성과 고온 신뢰성을 동시에 만족하는 RRAM 기반 매크로의 가능성을 입증하며, 자동차용 메모리 응용에 RRAM의 상용화를 앞당길 수 있는 중요한 성과로 평가된다.



[그림 4] 본 논문에서 제안하는 Wordline Driver(WLDRV) 아키텍처

저자정보



한창우 박사과정

- 소속 : 고려대학교 전기전자공학과
- 연구분야 : 차세대 반도체 소자 및 회로
- 이메일 : cwoo0105@naver.com
- 홈페이지 : <https://sites.google.com/view/kudclab>