

2026 VLSI Review

고려대학교 전기전자공학과 박사과정 한창우

Session 29 Advanced Memory Design

이번 2026 VLSI의 Session 29에서는 Advanced Memory Design을 주제로 총 5편의 논문이 발표되었다. 본 세션에서는 SRAM, DRAM, Mask ROM, Anti-Fuse OTP 및 STT-MRAM 등 다양한 메모리를 대상으로, 선단 공정에서의 고집적화와 저전력·저전압 동작, 동작 안정성 및 신뢰성 향상을 위한 회로 및 아키텍처 기술이 소개되었다. 특히 공정 미세화와 공급 전압 감소에 따라 발생하는 동작 마진 저하와 전력 소모 문제를 해결하면서도 높은 성능과 집적도를 확보하기 위한 다양한 설계 기법이 제안되었다.

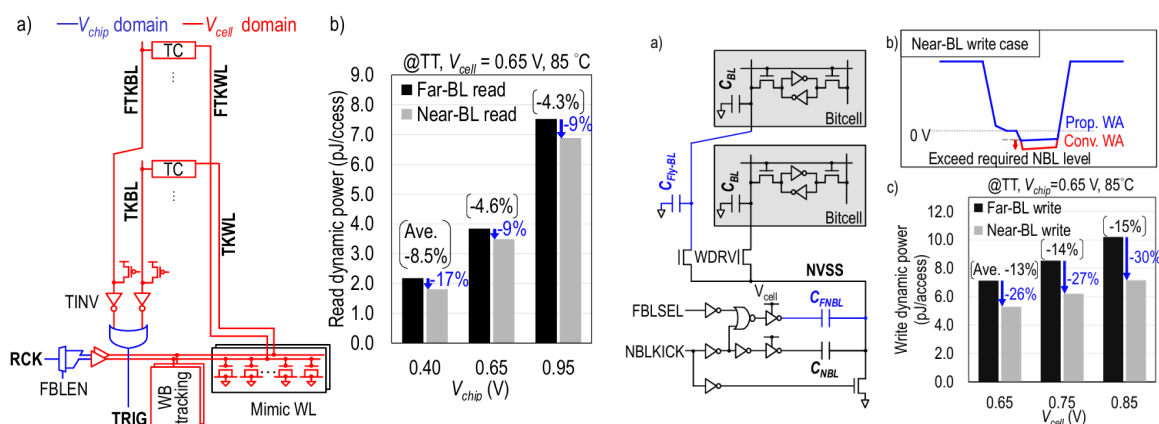
본 리뷰에서는 2 nm Nanosheet SRAM의 에너지 효율 향상 기술, 0.75 V 저전압 DRAM의 sensing 신뢰성 개선 기술, 그리고 PVT variation에 강인한 STT-MRAM의 dual-reference sensing 기술을 다룬 3편의 논문을 중심으로 살펴보고자 한다.

#29-1 본 논문은 TSMC에서 발표한 논문으로, 2 nm Nanosheet 공정 기반의 고밀도 6T SRAM에서 read/write 에너지 효율을 향상시키기 위한 Row-Access Aware Read Tracking 및 Write Assist(WA) 회로를 제안한다. 기존 dual-rail SRAM에서는 array와 peripheral의 공급 전압(V_{cell} , V_{chip}) 차이로 인해 read tracking과 실제 bit-line(BL) 방전 시간 사이에 오차가 발생하여 불필요한 동적 전력이 소모될 수 있다. 또한 기존 Negative Bit-Line(NBL) Write Assist는 worst-case bitcell을 기준으로 일정한 boost capacitance를 사용하기 때문에, 부하가 작은 near-array에서는 과도한 NBL boost로 인해 write power가 증가하는 문제가 존재하였다.

이를 해결하기 위해 near/far-array의 접근 위치를 고려한 read tracking 회로를 적용하고, tracking BL과 inverter를 memory array와 동일한 V_{cell} domain에서 동작시켜 실제 BL에 맞춰 Sense Amplifier Enable(SAE)과 WL 비활성화 시점을 조절하였다. 그 결과 worst-case에서 불필요한 BL differential swing을 45% 감소시키고, read dynamic power를 최대 8.7%, 전체 row access 평균 8.5% 절감하였다.

Write 동작에서는 접근 위치에 따라 NBL boost capacitance를 조절하는 Row-Access Aware Write Assist를 적용하여 near-array의 NBL over-boost를 억제하였다. 이를 통해 near-array의 write power를 26~30%, 전체 row access 평균 최대 15% 절감하였다. 최종적으로 2 nm 공정에서 제작된 539-kbit dual-rail SRAM macro는 37.42 Mbit/mm²의 bit

density와 125 °C에서 0.35~1.10 V의 동작 범위, 2.28 pJ/access의 read/write 에너지 효율을 달성하였다.

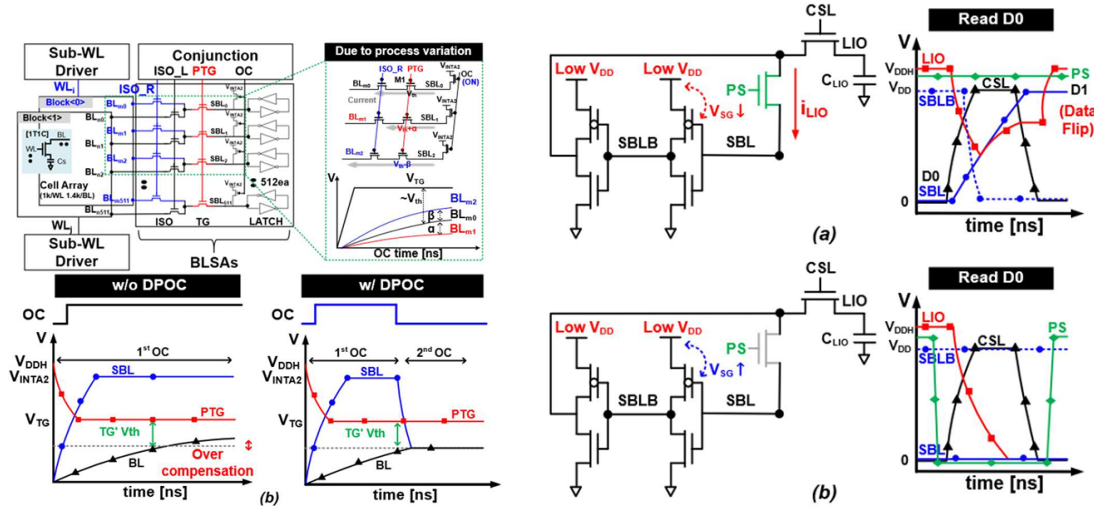


[그림 1] 제안된 Row-Access Aware Read Tracking 및 Write Assist 회로와 전력 절감 효과

#29-2 본 논문은 삼성전자에서 발표한 논문으로, 저전압 DRAM에서 sensing 신뢰성과 전력 효율을 향상시키기 위한 Fully Single-Ended Core Data-Line Architecture를 제안한다. 기존 DRAM은 공급 전압이 낮아질수록 bit-line의 charge-sharing voltage와 sense amplifier의 voltage headroom이 감소하여 안정적인 sensing이 어려워지며, single-ended charge-transfer sensing에서는 offset compensation 과정의 over-compensation과 CSL 활성화에 따른 latch disturb 문제가 발생할 수 있다.

이를 해결하기 위해 본 논문에서는 offset compensation을 두 단계로 분리하는 Dual-Phase Offset Compensation(DPOC) 기법을 적용하여 charge-transfer 이전의 over-compensation을 억제하였으며, 이를 통해 1 ns offset compensation window에서 Fail Bit Count(FBC)를 99.9% 감소시켰다. 또한 CSL pulse 동안 latch를 일시적으로 격리하는 Conditional Latch Gating(CLG) 기법을 적용하여 저전압에서 발생하는 data flip을 방지하였으며, $V_{DD}=0.75$ V에서 FBC를 99.4% 감소시켰다.

또한 BLSA부터 GIO까지 data path를 fully single-ended 구조로 구성하여 기존 differential 구조에서 요구되던 GIO metal layer 하나를 제거함으로써 약 2%의 backend cost를 절감하였다. 여기에 Data-State Inversion(DSI) 기법을 적용하여 동작 전류를 줄였으며, 기존 구조 대비 $IDD0$ 를 평균 27% 감소시켰다. 최종적으로 10-nm-class DRAM 공정에서 제작된 test chip을 통해 0.75 V의 저전압에서도 높은 sensing 신뢰성과 전력 효율을 동시에 확보할 수 있음을 실증하였다.



[그림 1] 제안된 DPOC 및 CLG 기법을 이용한 저전압 DRAM sensing 신뢰성 향상 구조

#29-5 본 논문은 성균관대학교와 삼성전자 공동 연구팀에서 발표한 논문으로, 28 nm FD-SOI 공정 기반 16 Mb STT-MRAM에서 PVT variation에 따른 read margin 저하를 개선하기 위한 Dual-Reference Sense Amplifier(DRSA)와 Array-Adaptive Dual-Reference(AADR) 기법을 제안한다. STT-MRAM은 높은 집적도와 비휘발성 특성을 갖지만, MTJ의 낮은 Tunnel Magnetoresistance(TMR)와 공정-전압-온도 변화로 인해 P/AP 상태의 저항 분포가 변하면서 read sensing margin과 access speed가 저하될 수 있다.

이를 해결하기 위해 본 논문에서는 하나의 reference를 사용하는 기존 방식 대신, P-state와 AP-state에 각각 대응하는 두 개의 reference를 사용하는 DRSA를 적용하였다. 또한 AADR을 통해 각 array의 저항 분포를 탐색하여 V_{REF_P} 는 P-state 분포의 상단에, V_{REF_AP} 는 AP-state 분포의 하단에 위치하도록 설정함으로써 local variation에 따라 reference를 보정하였다. 해당 calibration은 power-on 시 한 번 수행되고, 이후에는 저장된 digital code를 이용하여 동작하므로 runtime overhead를 최소화하였다.

그 결과 제안된 DRSA와 AADR 구조는 기존 single-reference 방식 대비 zero-BER read window를 -40°C , 25°C , 125°C 에서 각각 2.00배, 2.34배, 2.62배 확대하였다. 또한 0.9 V에서 25°C 기준 2.0 ns, 125°C 기준 3.3 ns의 read access time을 달성하였으며, 25°C 에서 0.6 pJ/bit의 read energy를 보였다. 이를 통해 넓은 온도 범위와 PVT variation 조건에서도 안정적이고 빠른 STT-MRAM read 동작이 가능함을 실리콘 측정을 통해 입증하였다.

