Buried power SRAM DTCO and system-level benchmarking in N3

S. Salahuddin, M. Perumkunnil, E. Dentoni Litta, A. Gupta, P. Weckx, J. Ryckaert, M.-H. Na, A. Spessot
Imec, Kapeldreef 75, 3000 Leuven, Belgium
salahu43@imec.be

Abstract

The increased metal resistance degrades both the performance and write margin of SRAM circuits in sub-10nm nodes. This paper utilizes buried power distribution as SRAM performance and write ability booster in 3nm node. BPR-SRAM offers up to 34.5% read speed and 498.6mV write margin improvement over conventional SRAM. Gem5 system simulator predicts up to 28.2% performance gain with server-processor having BPR-SRAM in L2 and L3 cache as compared to the baseline.

Introduction

Scaling of the static random-access memory (SRAM) bit-cell area in each new technology generation improves the on-die memory density while lowers the bitline (BL) and wordline (WL) metal capacitance (Fig. 1a). Thus, the performance of embedded cache is expected to improve with scaling. SRAM read speed is however degraded (Fig. 1b) due to the increased metal resistance (Fig. 1a) in scaled nodes. Increased BL resistance also reduces the write margin (Fig. 1b). This work reports design technology co-optimization (DTCO) to improve SRAM performance and write margin in 3nm node (N3).

SRAM DTCO and System Performance

The conventional SRAM bit-cell (Fig. 3a) in scaled nodes consists of highly resistive 1CD (critical dimension) WL and BL. The buried power distribution is proposed in [1] as a performance booster. The buried rails in SRAM open-up space in both the BL and WL metal layers. Wide and thus low resistive metal tracks can be used for both the WLs and BLs without increasing bit-cell area (Fig. 3b). The buried power rail (BPR) integration flow is summarized in Fig. 2a [1], [2], [5]. SRAM design rules are listed in Table I. The schematic illustration of buried power SRAM (BPR-SRAM) and TEM of BPR with FEOL/BEOL are shown in Fig. 2 (b) and (c), respectively. High-density (111) BPR-SRAM offers 68.9% and 69.7% reduction in WL and BL resistance (Fig. 4a), respectively, due to wider (3CD) metal tracks as compared to the conventional SRAM. Wider metal tracks in BPR-SRAM however increase the WL and BL capacitance by 11.6% and 12.1%, respectively (Fig. 4b). Despite the increase in parasitic capacitance, 111 BPR-SRAM offers up to 34.5% performance improvement (Fig. 5) as compared to the SRAM without BPR.

SRAM performance is expected to improve with cell-sizing. Read speed however degrades with cell-sizing in N3 (see Fig. 6) in both the conventional and BPR-SRAM. BL in high-aspect-ratio cells can be widened to lower the resistance. Wider (ex. 5CD) BLs however cannot guarantee performance improvement (Fig. 7) due to the increased metal capacitance. Fig. 7 also indicates that BL is no longer the critical path in scaled technology nodes. In quest for enabling the performance improvement with cell-sizing, the dual wordline (M1+M3 for WL) scheme [3] is also evaluated in 3nm node. Dual-WL offers a 50% reduction in resistance at a cost of 2.1× increase in capacitance. Dual-WL thus cannot ensure SRAM performance enhancement with cell sizing (Fig. 8) in N3. For high-speed memory subsystems, high-density (111) BPR-SRAM with 3CD BL and WL is the optimal design choice.

The read current flows through the buried VSS (as illustrated in Fig. 9), thus its resistance affects both the read speed and read margin (RSNM). 50Ω/μm buried rail resistance, which is recommended for logic application [1], [4] may cause ~5% speed degradation (Fig. 10) due to lower read current as compared to fine-grain VSS distribution. The read margin however improves by ~3% (Fig. 10) under this constraint.

During a write operation, current flows from one of the data storage nodes to the write-driver. The increased bitline (BL) resistance limits the current and thus degrades the write margin in conventional high-density SRAM. The high-density BPR-SRAM provides 69.5% larger write margin (Fig. 11) due to wider (lower resistance) BL as compared to its conventional counterpart. High aspect-ratio bit-cells with and without BPR provide the same write margin (Fig. 11) since nonminimal size BL can be used in both cases.

The contention current which degrades the write margin flows from VDD towards the data storage node (illustrated in Fig. 9). Proper optimization of VDD resistance lowers the contention current and hence improves the write margin. The write margin of high-density BPR-SRAM with VDD distribution resistance is shown in Fig. 12. Mint based (column-wise) power distribution in BPR-SRAM offers an additional 193.9mV enhancement in write margin (Fig. 12) as compared to the fine-grain VDD distribution. High-density (111) BPR-SRAM with Mint based VDD distribution, thus provides 498.6mV, 127.6mV, and 106.8mV larger write margin (Fig. 12) as compared to the conventional 111, 122, and 133 SRAM, respectively. This improvement in write margin comes at the cost of hold margin (HSNM) degradation in the worst-case half-select cell (illustrated in Fig. 12). The HSNM in the worst-case cell is momentarily lowered by 16.6% (Fig. 12). However, the worst-case HSNM is 67.9% larger than RSNM. Data flip in half-select cells is therefore unlikely. Mint based VDD distribution is thus recommended for BPR-SRAM.

In order to predict system-level benefits, we simulated an 8-core 16-thread Intel Xeon 5 series processor in gem5. The processor consists of 64KB L1 cache per core, 256KB L2 cache per core, and 11MB shared L3 cache. Replacement of conventional SRAM bit-cells in L2 and L3 caches with BPR-SRAM offers up to 28.2% system-level performance boosting. A further performance gain is expected with the replacement of register files and L1 cache with BPR-SRAM.

Conclusions

Buried power SRAM DTCO is reported in this work for 3nm technology node. With a high resistive VDD and buried-VSS distribution networks, high-density (111) BPR-SRAM can outperform its high-aspect-ratio counterparts (122 and 133 bit-cells) in both the read speed and write margin. BPR-SRAM in L2 and L3 cache provides up to 28.2% speed gain in high-performance server processor as compared to the baseline.

References

[1] Salahuddin S. *IEEE EDL*, Vol. 40, No 8, pp. 1261-1264, Aug. 2019. [2] Gupta A. *IEEE IITC* 2018, pp. 4-6. [3] Chang J. *IEEE ISSCC* 2017, pp. 206-207. [4] Chava B. *Proc. SPIE*, vol. 10588, Mar. 2018. [5] Gupta A. submitted to *IEEE VLSI Symposium* 2020.

978-1-7281-6460-1/20/$31.00 ©2020 IEEE

Authorized licensed use limited to: Zhejiang University. Downloaded on December 24,2021 at 09:42:58 UTC from IEEE Xplore. Restrictions apply.

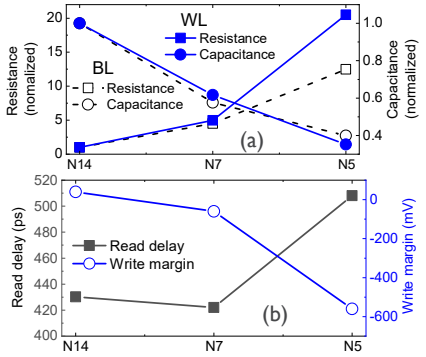

Fig. 1. (a) BEOL resistance and capacitance, (b) read delay and write margin with scaling.

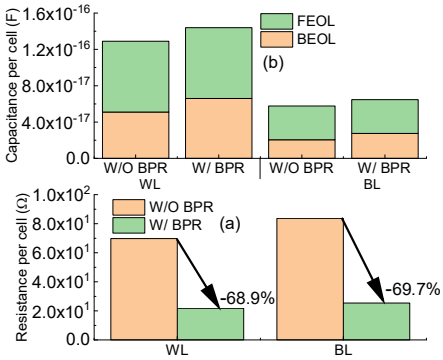

Fig. 4. SRAM WL/BL resistance and capacitance comparison.

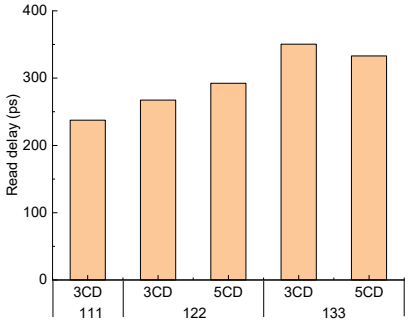

Fig. 7. BPR-SRAM ready delay with bitline width.

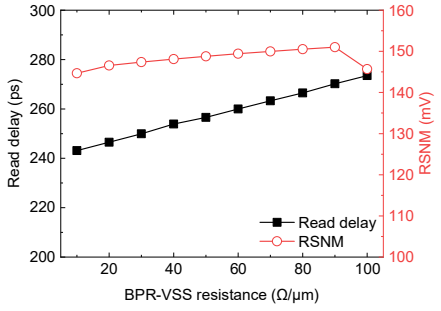

Fig. 10. Read delay and RSNM with BPR-VSS resistance.

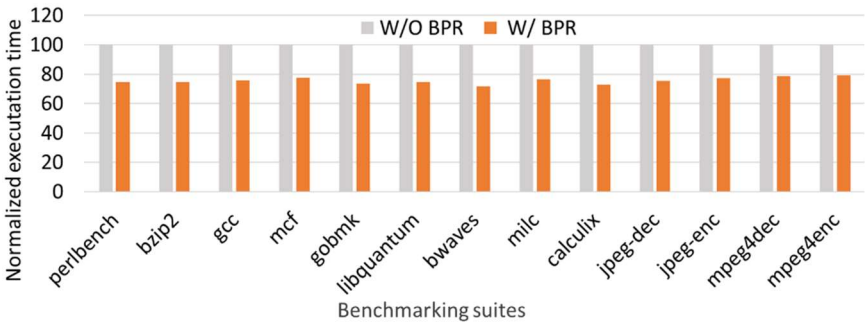

Fig. 13. System-level performance comparison.

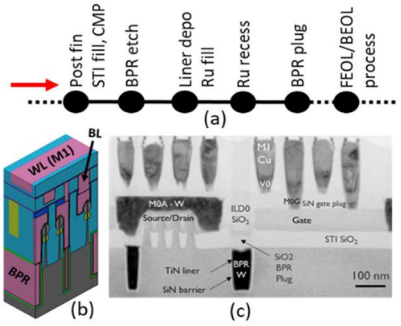

Fig. 2. (a) BPR integration flow. (b) Illustration of BPR in SRAM. (c) TEM cross-section of BPR.

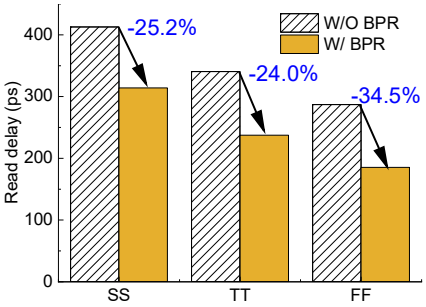

Fig. 5. SRAM performance comparison under different process corners.

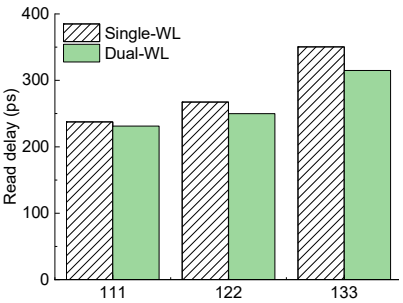

Fig. 8. BPR-SRAM ready delay with Single-WL and dual-WL.

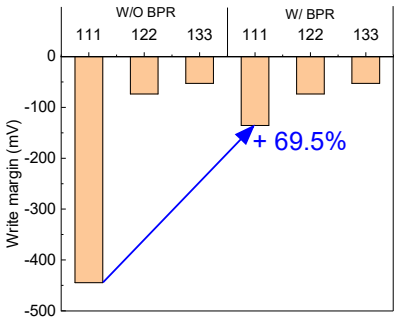

Fig. 11. Write margin with cell sizing.

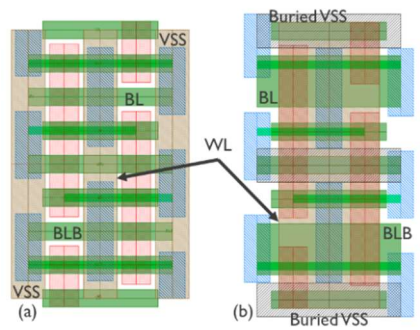

Fig. 3. SRAM bit-cell. (a) without BPR. (b) with BPR.

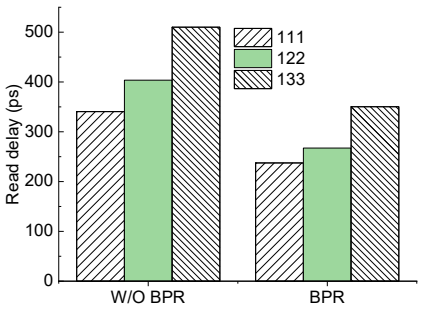

Fig. 6. SRAM read delay with cell sizing.

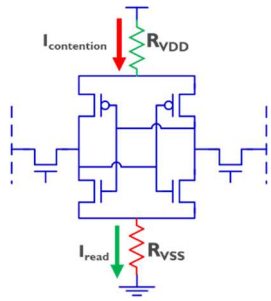

Fig. 9. Illustration of read current path and contention current path during write operation.

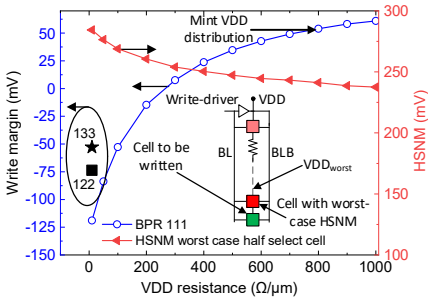

Fig. 12. Write margin and worst-case hold static noise margin (HSNM) with VDD rail resistance.

TABLE I
LAYOUT DESIGN RULES ASSUMED

Parameters	Value (nm)
Fin pitch	21
Connected gate pitch	45
Metal pitch	21
Fin to BPR separation	8
Buried rail width	21
Buried rail pitch	84
Buried rail depth	147

Authorized licensed use limited to: Zhejiang University. Downloaded on December 24,2021 at 09:42:58 UTC from IEEE Xplore. Restrictions apply.