19.4 A Configurable SRAM with Constant-Negative-Level Write Buffer for Low-Voltage Operation with 0.149μm² Cell in 32nm High-κ Metal-Gate CMOS

Yuki Fujimura, Osamu Hirabayashi, Takahiko Sasaki, Azuma Suzuki, Atsushi Kawasumi, Yasuhisa Takeyama, Keiichi Kushida, Gou Fukano, Akira Katayama, Yusuke Niki, Tomoaki Yabe

Toshiba Semiconductor, Kawasaki, Japan

恒定负电平

This paper presents a configurable SRAM for low-voltage operation with constant-negative-level write buffer (CNL-WB) and level programmable wordline driver for single supply (LPWD-SS) operation. CNL-WB is suitable for compilable SRAMs and it improves write margin by featuring an automatic BL-level adjustment for configuration range of four to 512 cells/BL using a replica-BL technique. LPWD-SS optimizes the tradeoff between disturb and write margin of a memory cell, allowing a 60% shorter WL rise time than that of the conventional design [1] at 0.7V. A test-chip is fabricated in a 32nm high-κ metal-gate CMOS technology with a 0.149μm² 6T-SRAM cell. Measurement results demonstrate a cell-failure rate improvement of two orders of magnitude for an array-configuration range of 64 to 256 rows by 64 to 256 columns.

For a low power SoC, a configurable SRAM operating at low voltage is a key component. The characteristic variation of a scaled device has been increasing, thus compromising the low-voltage functionality of high-density SRAMs [2,3]. Functionality and high yield at low supply voltages such as 0.7V has become more difficult to achieve with the degradation of write margin and the optimization between read disturb and write margin. Among recently reported techniques, the scheme that implements BL negative biasing and WL level control [4] is an effective approach [1]. However, the conventional schemes are not suitable for configurable SRAM for two reasons. First, the circuit must be optimized for each different array configuration to create a suitable BL level. Second, the conventional scheme of WL-level control causes rise-time degradation of the WL at low voltage. This paper describes two circuit techniques for coping with the problems, CNL-WB to automatically adjust the BL bias to a suitable level in a wide range of array configurations and LPWD-SS for fast WL activation at low voltage.

Figure 19.4.1 shows the circuit schematic of the configurable SRAM. Each I/O block has the BL negative bootstrap circuit named CNL-WB. The BL capacitance monitor working as a replica of a BL generates a signal activating the bootstrap (boost_en). Every WL driver has a pull-down unit (PDU) to lower the WL voltage for optimizing the tradeoff between disturb and write margin.

Figure 19.4.2 shows the circuit schematic and simulated waveforms of CNL-WB. One issue of the conventional technique [1] is as follows: when using a negative-bootstrap circuit in a configurable SRAM, the bootstrap capacitance and the timing of activating the signal boost_en must be optimized depending on the memory cell count on a BL. The custom optimization makes this technique unsuitable for compilable SRAMs. Without such optimization, the negative BL level can become too high and results in insufficient write margin. On the other hand, if the level is too low it causes a reliability problem through data corruption on memory cells with unselected WLs on the accessed column. CNL-WB solves this problem with a bootstrap circuit by automatically adjusting the BL bias to an optimized constant negative level for different array configuration of four to 512 cells/BL. The key point of this feature is to automatically control the amount of charge stored in the bootstrap capacitor *C_boost* depending on the number of rows. To generate a constant BL level, the additional charge stored in *C_boost*, Δq, is proportional to the BL capacitance increase, ΔC_{BL}. The charge Δq is a product of the current *i_wb*/*n* and the additional time for pulling down the node *vn* by NB, Δt. Here, *n* is the current ratio of NA and NB. The time Δt is generated by the BL capacitance monitor that detects the level of the replica BL crossing $V_{DD}/2$. In the monitor, replica blt and blc are connected to each other so that the replica BL capacitance becomes $2C_{BL}$. As shown in the right portion of Fig. 19.4.2, CNL-WB adaptively sets Δq to be proportional to ΔC_{BL} by the BL-capacitance monitor. The simulated waveforms show the voltage of blt, blc, wl, boost_en, and vn. The solid and dashed curves show the results of the arrays implementing a longer BL with 512 cells and a shorter one with four cells, respectively. In either case, blt goes to approximately the same level, around -0.15V.

Figure 19.4.3 shows simulated negative BL voltage for the case that the BL length is four to 512 cells/BL with supply voltage of 0.7V at 25°C. The target bias level is -0.15 ± 0.05V. The upper-side limit of -0.10V is defined by the necessity for improving cell failure rate by three orders of magnitude compared to the case without this technique. The lower-side limit of -0.20V is defined to prevent reliability problems. As shown in Fig. 19.4.3, the BL level with CNL-WB keeps within the target range in the range of four to 512 cells/BL. However, the BL level with conventional circuits optimized for either four or 512 cells/BL does not keep inside the target range for the BL length variation.

Figure 19.4.4 shows the circuit schematic and the simulation results of LPWD-SS. LPWD-SS controls the WL voltage (V_{WL}) to balance the tradeoff between disturb and write margin. This circuit is superior to the conventional design [1] in terms of WL rise time at low voltage. The conventional circuit supplies a suppressed voltage to the source of pull-up PMOS (PU). It decreases the overdrive voltage of PU, causing a slow rise time of WL with low supply voltage. LPWD-SS avoids the problem by lowering WL voltage with a pull-down unit (PDU) connected to the drain of PU. For the pull-down unit a PMOS connected to a polysilicon resistance, R_{poly} is introduced. The upper right portion of Fig. 19.4.4 shows the simulated V_{DD} dependence of the WL rise time. It shows that LPWD-SS has a 60% shorter WL rise time than that of the conventional design at 0.7V. The lower-right portion shows the simulated WL voltage on various PT conditions, showing the effect of R_{poly}. The target WL voltage is 0.65V at a supply voltage of 0.7V. R_{poly} gives a negative feedback effect on the pull-down current and reduces the voltage variation at the process corner with fast PMOS condition. The WL variation is 40% smaller with R_{poly} compared to the case without this resistor, which corresponds to a 75% lower cell failure rate.

Figure 19.4.5 shows the simulated cell-failure rate with CNL-WB and LPWD-SS. CNL-WB improves the cell failure rate by three orders of magnitude. In addition, the cell failure rate is reduced by about one order of magnitude by LPWD-SS.

Figure 19.4.7 shows the micrograph and features of the test-chip fabricated in 32nm high-κ metal-gate CMOS technology. A 0.149μm² memory cell is used [3]. The 1.25Mb macro includes subarray densities of 4Kb to 64Kb, which consist of 64 to 256 rows and 64 to 256 columns. The area overhead of CNL-WB and LPWD-SS is 2% and 3% of a 64Kb subarray, respectively. Figure 19.4.6 shows the measurement results. The upper-right portion of Fig. 19.4.6 shows the measured BL waveform with supply voltage of 0.7V at room temperature. The lower right portion shows the measured dependence of the BL bias level on the BL-length. The variation of the bias level is about 20mV for configurations of 64 to 256 cells/BL. The left portion shows the measured cell-failure rate. At 0.5V with 1.25Mb macro, CNL-WB and LPWD-SS improve the cell failure rate by two orders of magnitude for all configurations.

Acknowledgements:
The test-chips have been fabricated by 32nm shuttles of International Semiconductor Development Alliance (ISDA).

References:
[1] K. Nii et al., "A 45-nm Single-port and Dual-port SRAM family with Robust Read/Write Stabilizing Circuitry under DVFS Environment," *Symp. VLSI Circuits*, pp. 212-213, Jun. 2008.
[2] Y. Wang et al., "A 4.0 GHz 291Mb Voltage-Scalable SRAM Design in 32nm High-κ Metal-Gate CMOS with Integrated Power Management," *ISSCC Dig. Tech. Papers*, pp. 456-457, Feb. 2009.
[3] H.S. Yang et al., "Scaling of 32nm Low Power SRAM with High-κ Metal Gate," *IEDM Dig. Tech. Papers*, pp. 233-236, Dec. 2008.
[4] O. Hirabayashi et al., "A Process-Variation-Tolerant Dual-Power-Supply SRAM with 0.179μm² Cell in 40nm CMOS Using Level-Programmable Wordline Driver," *ISSCC Dig. Tech. Papers*, pp. 458-459, Feb. 2009.

 978-1-4244-6034-2/10/$26.00

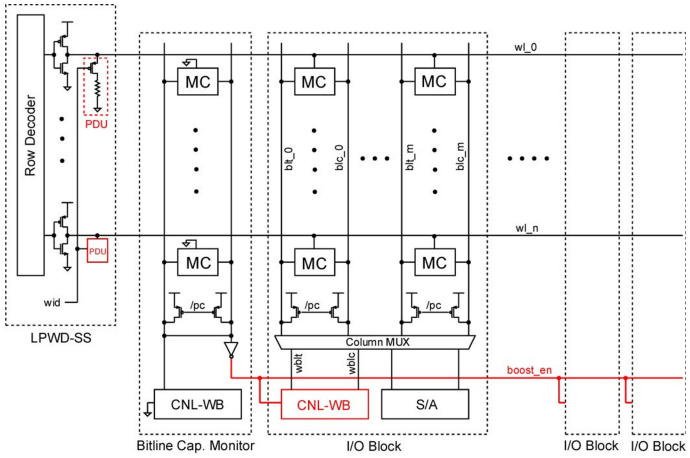


Figure 19.4.1: CNL-WL and LPWD-SS configurable SRAM.

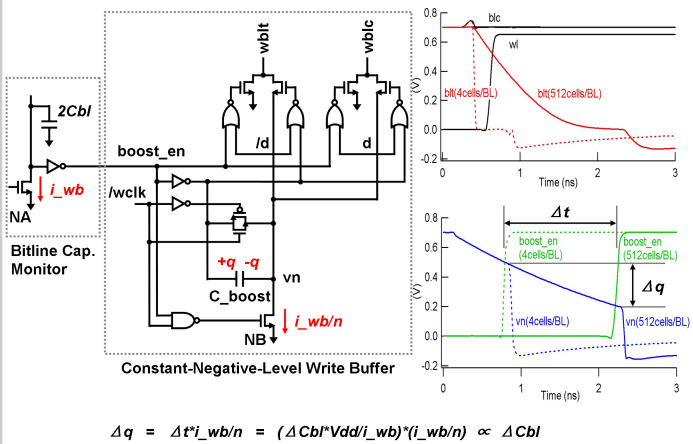


Figure 19.4.2: Constant-negative-level write buffer.

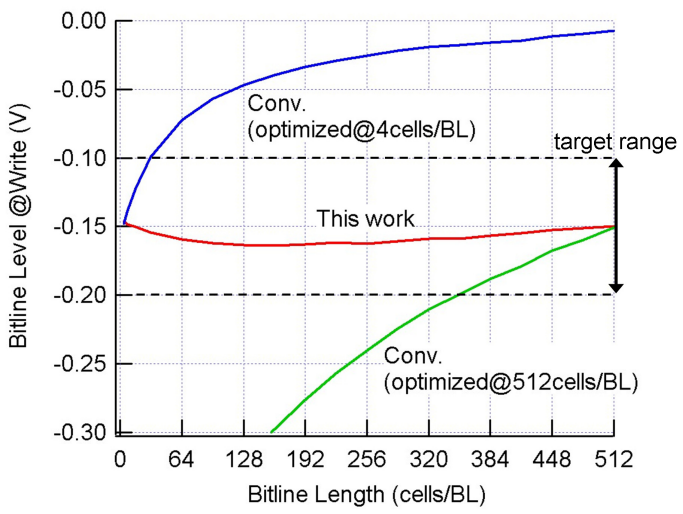


Figure 19.4.3: Simulated negative bitline level.

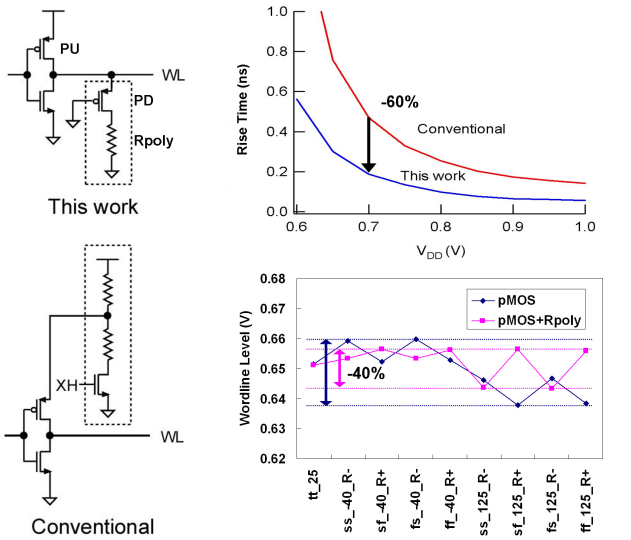


Figure 19.4.4: Level programmable wordline driver for single supply.

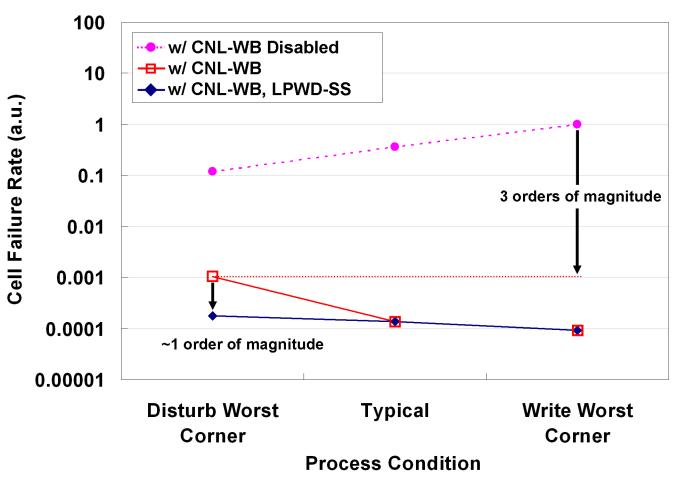


Figure 19.4.5: Simulated cell failure rate.

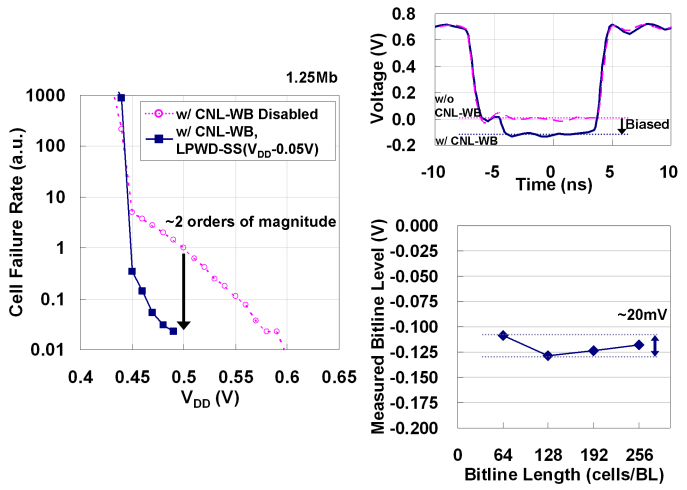
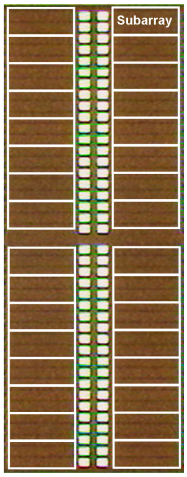


Figure 19.4.6: Measurement results at room temperature.

19



Technology	32nm High-k/Metal Gate CMOS, 8Metal Layers
Cell size	0.149μm^2
Chip Organization	4K – 64Kb Subarray X 32 (total 1.25Mb)
Subarray Configuration	Row 64 / 128 / 192 / 256 Column 64 / 128 / 192 / 256
Chip size	5.27mm^2

Figure 19.4.7: Test chip micrograph and features.

978-1-4244-6034-2/10/$26.00