Dynamic Flash Memory with Dual Gate Surrounding Gate Transistor (SGT)

Koji Sakui and Nozomu Harada
Unisantis Electronics Singapore Pte Ltd., 03-17A, The Alpha Singapore Science Park II, Singapore, sakui@unisantis.com

***Abstract*—This paper proposes an ultra-scaled memory device, called 'Dynamic Flash Memory (DFM)'. With a dual-gate Surrounding Gate Transistor (SGT), a capacitorless $4F^2$ cell can be achieved. Similar to DRAM [1], refresh is needed, but high-speed block refresh can improve the duty ratio. Analogous to Flash [2], three fundamental operations of "0" Erase, "1" Program, and Read are needed, but fast random read and page program are available. Through 2D TCAD simulation, the three basic operations for Dynamic Flash Memory have been qualitatively validated.**

Keywords—Surrounding Gate Transistor (SGT), GAA, Z-RAM, Capacitorless DRAM, Floating Body

I. Introduction

During its development the DRAM has only made modest gains in scaling, as shown in Fig. 1 [3]-[9]. Reduction of the cell size from $6F^2$ to $4F^2$ [10] has been impacted by numerous difficulties, such as three dimensional capacitive coupling of WL-WL, BL-BL, and WL-transistor body, as well as high aspect ratio of the cell capacitor to store certain charges [11].

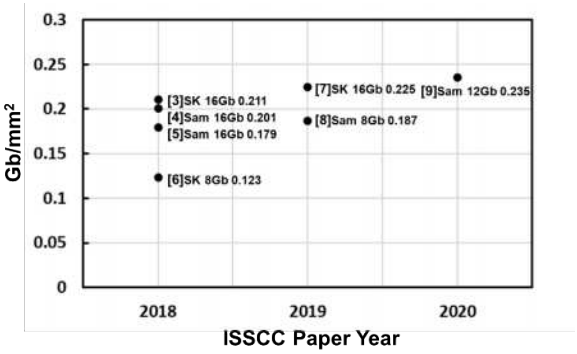


Fig. 1 DRAM density trend announced in recent ISSCC's.

The capacitor less 1-transistor DRAM (1T-DRAM) has been studied extensively [12]-[19] over the last 30 years. But the conventional 1T-DRAM had a serious problem of the Word Line (WL)-Floating Body (FB) capacitive coupling, which prevented 1T-DRAM from commercialization.

II. Conventional 1T-DRAM

Fig. 2 shows the data write and read mechanism. As for "1" write, electron-hole pairs are generated by the impact ionization near the drain. The generated excess holes boost the P-type FB of the NMOS on the SOI, while the generated electrons flow to the drain. With respect to "0" write, the drain bias goes negative, so that the holes are absorbed in the drain by the forward biased pn-junction between the drain N^+ and P-FB, as shown in Fig. 2. Whether the generated holes exist in the FB or not, the threshold voltages of the NMOS are changed because of the back-bias effect. Thereby, the stored data "1" and "0" are read. But the conventional 1T-DRAM had a serious problem of the WL-FB capacitive coupling, as shown in Fig. 3. Irrespective of whenever the WL goes up and down during writing and reading, the FB also goes up and down accordingly. Also, "1" and "0" are written by the positive and negative BL per page simultaneously, so that the reset and unselected WL voltages are forced to be negative so as not to destroy the stored data. As a result, the cell has a narrow margin between "1" and "0", and it always suffers from noise, which induces the leakage current, as illustrated in Fig. 3. Here, C_{WL}, C_{BL}, and C_{SL} are WL to FB, BL and SL junction capacitances, respectively.

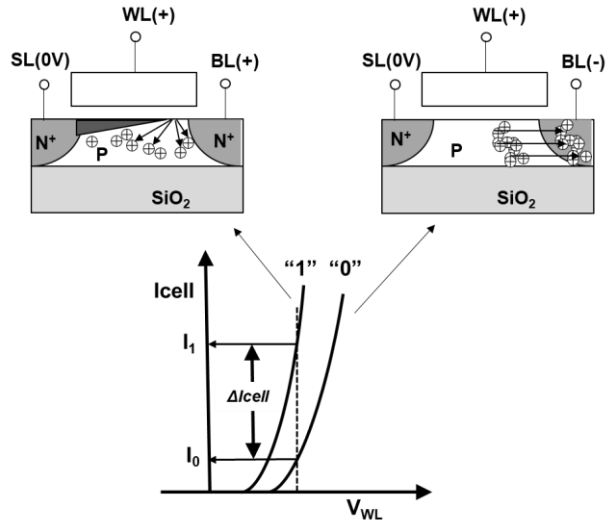


Fig. 2 Conventional 1T-DRAM cell operation mechanism.

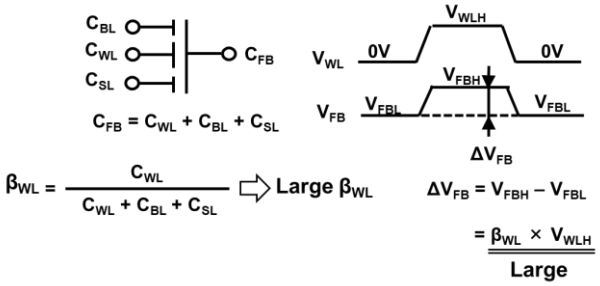
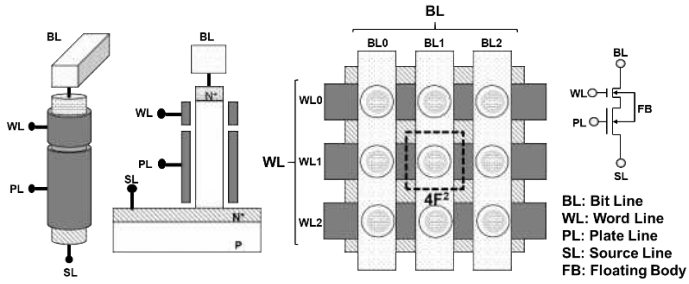


Fig. 3 WL-FB serious capacitive coupling.

III. Dynamic Flash Memory (DFM) Features

1. *Perfect $4F^2$ Cell Size*

Fig. 4 presents the DFM structure, which is composed of single SGT [20] with dual gates. Straight WL's run in the row direction, and straight BL's run in the column direction, so that ultra-scaled $4F^2$ cells exist in their cross-points.



(a) Bird's Eye (b) Cross Section (c) Top (d) Equivalent Circuit

Fig. 4 DFM structure.

2. PL Significantly Reduces WL Capacitive Coupling Ratio

On programming and reading, the PL gate capacitive coupling ratio to the FB of β_{PL} can significantly reduce the WL of β_{WL}, as illustrated in Fig. 5.

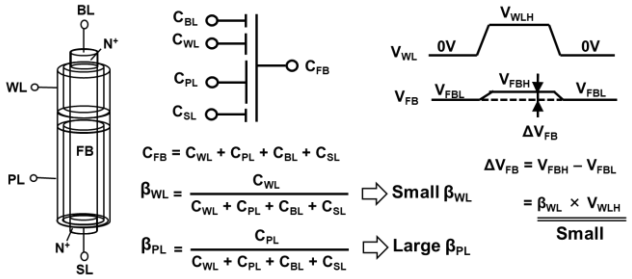


Fig. 5 The FB stabilization by the PL gate.

3. PL Creates Dramatical "1"-"0" Margin and Protection

As a certain fixed voltage is applied to the PL gate, the "0" erased threshold voltage of the PL gate can completely prevent the cell current from flowing. Basically, the PL and WL gates play role as a memory device, and a select gate with a variable threshold voltage, respectively. In the case of "1" program, when generated holes are in the FB, the PL gate can induce a current. Conversely, in the "0" erase case, as no holes are in the FB, the PL gate doesn't induce a current. As a result, the PL gate provides a significant "1" and "0" margin, as shown in Fig. 6. Furthermore, the programmed PL gate can create the inversion layer, so that it can shield the FB from external noises.

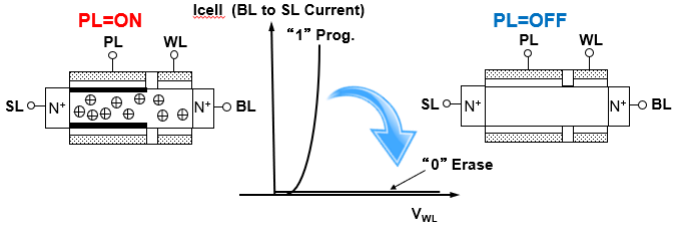


Fig. 6 Wide "1"-"0" margin and the inversion layer shield.

IV. DFM Program and Erase Mechanism

1. Programmed by the Source-side Impact Ionization

As for "1" program, the source-side impact ionization has been utilized for the generation of hole-electron pairs, where the PL gate operates the saturation region, while the WL gate operates the linear region with the virtual drain, as shown in Fig. 7. When the WL and BL return to 0V after the generation of holes, they exert little influence on the FB voltage, because the PL gate stabilizes the FB by the large capacitive coupling ratio to the FB of β_{PL}.

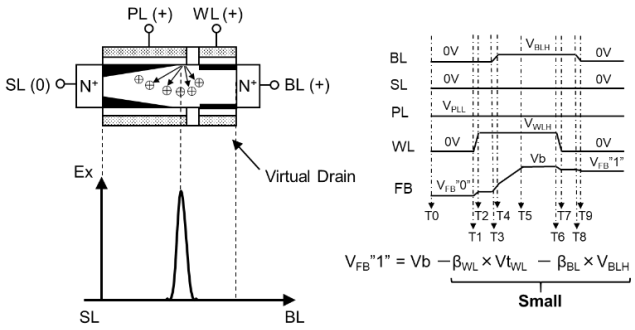


Fig. 7 The DFM "1" program mechanism.

2. Erase Scheme I: SL Negative Bias

Similar to the 1T-DRAM [12]-[19], holes can be ejected by the SL negative bias erasure with the floating BL. All the memory cells in a block, which is fabricated on the twin-well, are simultaneously erased by the negative biased SL. This design would require the negative charge pump, block decoder, and the twin-well process like the NAND Flash memory.

3. Erase Scheme II: Holes push-out by the PL coupling

In the design of the holes push-out erasure by the PL capacitive coupling, neither a negative charge pump nor twin-well process is needed, so that not only the power but also the layout area will be reduced. Furthermore, the simple design and fast erasure operation will be attractive and promising. At the start of the erase operation, prior to the PL and WL rise, both the BL and SL go to the high levels of V_{BLH} and V_{SLH} at T1~T2, respectively, so as not to generate the inversion layer, as shown in Fig. 8.

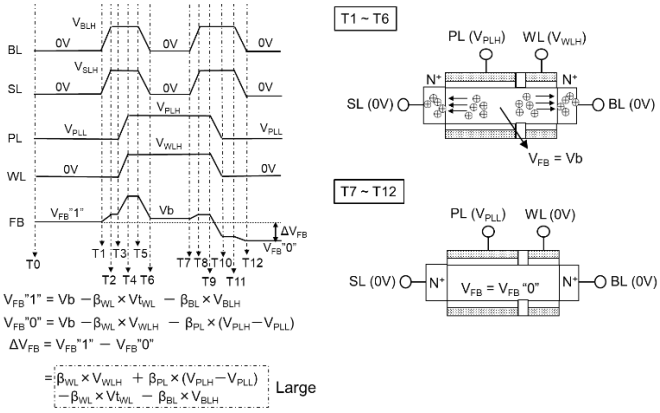


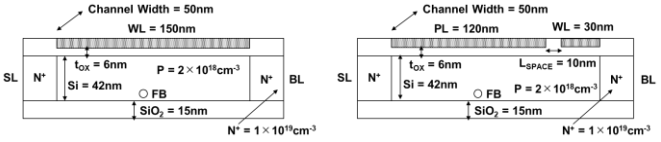
Fig. 8 Holes push-out by the PL coupling to the FB.

At T3~T4, the PL and WL go high, and then the FB is boosted due to their capacitive coupling to the FB. When the BL and SL return to 0V at T5~T6, both pn-junctions of the BL and SL should be forward-biased, so that the excess holes are pushed out. Consequently, the FB voltage of V_{FB} eventually becomes Vb. Again, prior to the fall of PL and WL, at T7~T8 both the

BL and SL go high to extinguish the inversion layer which has been generated at T5~T6. Then, when the PL and WL go low at T9~T10, the FB becomes deeply negative due to their capacitive coupling. Finally, the BL and SL fall at T11~T12, which can compensate their boosting the FB at T7~T8. The difference of ΔV_{FB} between the V_{FB}"1" and V_{FB}"0" is sufficiently large, because the sum of the β_{PL} and β_{WL} is close to 1. Therefore, the wide "0" and "1" margin can be attained without a negative bias circuit.

V. TCAD Simulation Results

DFM operation has been qualitatively validated by the 2D T-CAD simulation of Advance/TCAD/Device by AdvanceSoft Corporation in comparison with the 1T-DRAM, as illustrated in Fig. 9. The model parameters of 1T-DRAM are referred to T. Ohsawa's paper [19]. Fig. 10 shows the simulation result of 1T-DRAM. Due to the large WL coupling ratio, the FB goes up and down, according as the WL movement. Also, "1" and "0" are written by the positive and negative BL per page simultaneously, so that the reset and unselected WL voltages are forced to be negative. Thereby, a very small margin remains. Conversely, with a high capacitive coupling ratio of the PL to the FB, and the block erase operation, DFM has successfully attained a huge wide margin of "1" program and "0" erase, as shown in Figs. 11 and 12. With respect to "0" read, the WL voltage has been intentionally increased up to 3V, so that the PL gate ON/OFF can be verified. The V_{WL}-Icell's for both erase schemes were similar.



(a) 1T-DRAM (b) DFM

Fig. 9 Simulation models for (a) 1T-DRAM and (b) DFM.

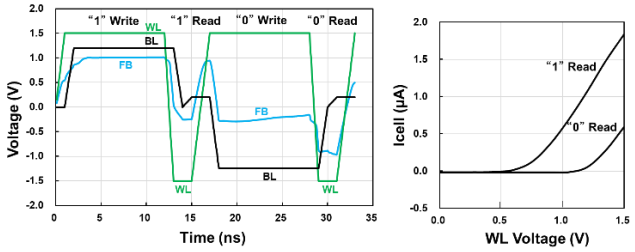


Fig. 10 1T-DRAM simulation result.

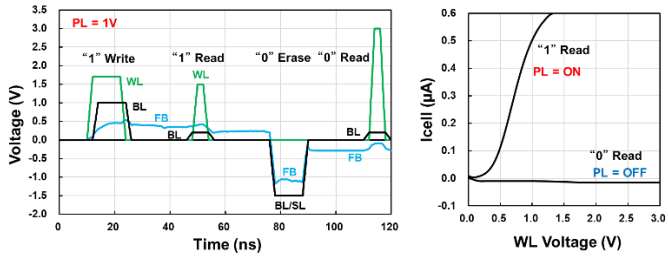


Fig. 11 Erase Scheme I: SL negative bias.

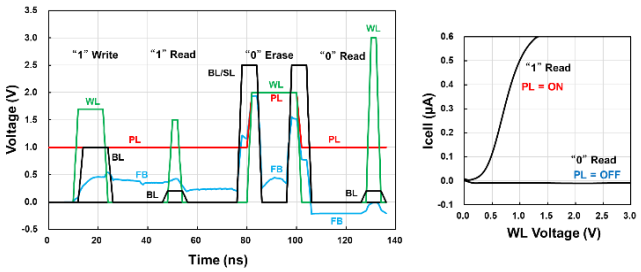


Fig. 12 Erase Scheme II: Holes push-out.

VI. Block Erase Opetaion

Fig. 13 explains the block erase operation. Like Flash, a controller can manage the Logical-to-Physical Look-up Table (LP-Table). New data is copied and reprogrammed into one of erased blocks.

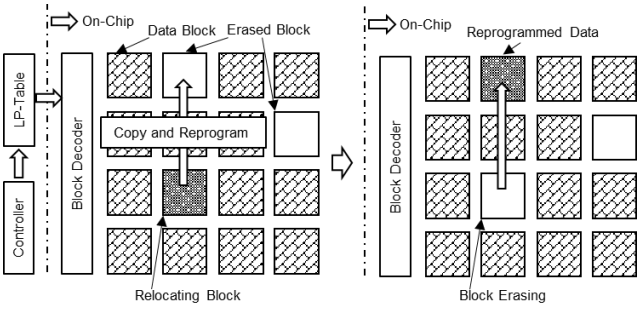
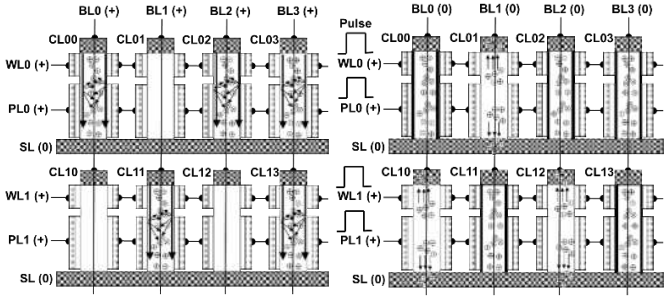


Fig. 13 Block Erase Operation.

VII. Block Refresh Opetaion



(a) "1" Refresh (b) "0" Refresh

Fig. 14 Block Refresh Operation.

Fig. 14 presents the block refresh operation. With respect to the "1" refresh (a), both the threshold voltages of the WL and PL gates are low for "1" programmed cells of CL00, CL02, CL03, CL11, and CL13. Therefore, certain positive biases are applied to the all BL's, WL's, and PL's, the currents can flow on the "1" programmed cells, so that those currents generate electron-hole pairs for the refresh operation. As for "0" refresh (b), the pulses are input to the PL and WL with BL=0V and SL=0V. The threshold voltages of the "0" erased cells of CL01, CL10, and CL12 are high, so that no inversion layers are generated. As a result, excess holes are pushed out from the SL and BL pn-junctions, because the FB voltages are boosted by the PL and WL capacitive couplings to the FB. Conversely, the

threshold voltages of "1" programmed cells of CL00, CL02, CL03, CL11, and CL13 are low, so that the inversion layers are generated to shield the excess holes from the PL and WL capacitive couplings. Unlike DRAM, as the refresh operation can be done by block, the duty cycle has been greatly improved.

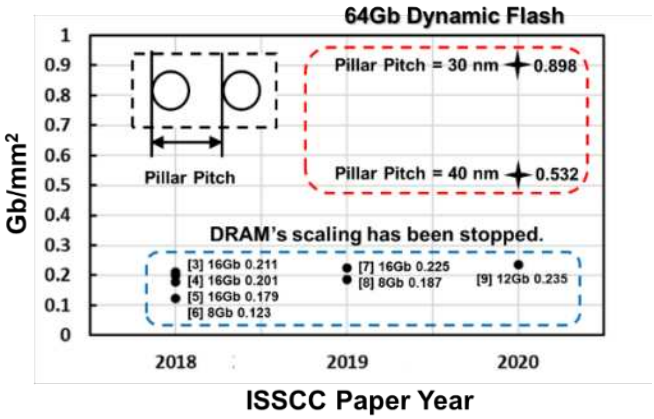


Fig. 15 Density comparison between DFM and DRAM.

VIII. Competitive DFM Strategy

Fig. 15 compares the density between the DFM and DRAM presented at recent three year ISSCC's. The density of the DFM should be potentially four times higher than DRAM. Table 1 summarizes the main features of the DFM. In the memory devices, there are three key metrics; 1) Speed, 2) Power, and 3) Bit Cost. The most significant element among three is 3) Bit Cost. Therefore, the low-cost DFM has a high potential to replace the existing DRAMs, the market for which is shown, in Fig. 16.

Table 1 Main features of DFM.

Simple Cross-Point Structure	Easy to Fabricate
$4F^2$ Cell Size	Ultra Small
PL Stabilization and Block Erase	Wide Margin between "1" and "0"
Capacitor Less	High Speed
Block Refresh	Low Refresh Duty Ratio
Low Leak Current	Long Refresh Cycle

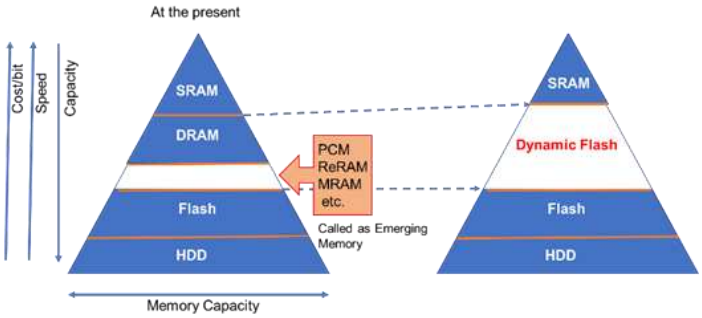


Fig. 16 Potential DFM positioning.

IX. Conclusions

A novel memory device of the Dynamic Flash Memory (DFM) has been introduced. The DFM can exist at a tight cross-point of the WL and BL with the memory cell size of $4F^2$. A low cost DFM is a very promising candidate next to DRAM and NAND in the memory hierarchy.

Acknowledgements

The authors wish to thank Mr. A. Hagiwara of AdvanceSoft Corporation for the TCAD simulation to validate the DFM operation.

References

[1] R. H. Dennard, "Field-Effect Transistor Memory," *U.S. Patent* 3,387,286, Jun. 4, 1986.

[2] F. Masuoka, "Semiconductor Memory Device," *U.S. Patent* 4,437,174, Mar. 13, 1984.

[3] S. Shim et al., "A 16Gb 1.2V 3.2Gb/s/pin DDR4 SDRAM with Improved Power Distribution and Repair Strategy," in *ISSCC*, pp. 212-213, Feb. 2018.

[4] K. C. Chun et al., "A 16Gb LPDDR4X SDRAM with an NBTI-Tolerant Circuit Solution, an SWD PMOS GIDL Reduction Technique, an Adaptive Gear-Down Scheme and a Metastable-Free DQS Aligner in a 10nm Class DRAM Process," in *ISSCC*, pp. 206-207, Feb. 2018.

[5] Y.-J. Kim, "A 16Gb 18Gb/s/pin GDDR6 DRAM with Per-Bit Trainable Single-Ended DFE and PLL-Less Clocking," in *ISSCC*, pp. 204-205, Feb. 2018.

[6] K.-D. Hwang et al., "A 16Gb/s/pin 8Gb GDDR6 DRAM with Bandwidth Extension Techniques for High-Speed Applications," in *ISSCC*, pp. 210-211, Feb. 2018.

[7] D. Kim et al., "A 1.1V 1ynm 6.4Gb/s/pin 16Gb DDR5 SDRAM with a Phase-Rotator-Based DLL, High-Speed SerDes and RX/TX Equalization Scheme," in *ISSCC*, pp. 380-381, Feb. 2019.

[8] K.-S. Ha et al., "A 7.5Gb/s/pin LPDDR5 SDRAM with WCK Clocking and Non-Target ODT for High Speed and with DVFS, Internal Data Copy, and Deep-Sleep Mode for Low Power," in *ISSCC*, pp. 380-381, Feb. 2019.

[9] H.-J. Chi et al., "An 8.5Gb/s/pin 12Gb-LPDDR5 SDRAM with a Hybrid-Bank Architecture using Skew-Tolerant, Low-Power and Speed-Boosting Techniques in a 2nd generation 10nm DRAM Process," in *ISSCC*, pp. 332-333, Feb. 2020.

[10] H. Chung et al., "Novel $4F^2$ DRAM Cell with Vertical Pillar Transistor(VPT)," in *Proc. ESSDRC*, pp. 211-214, 2011.

[11] D. S. Woo, "DRAM: Its Challenging History and Future," in the *Proc. 2018 IEDM Short Course 2*, pp. 3-57, 2018.

[12] K. Sakui, "Semiconductor Storage Device," *Japanese Patent Application*, Number: 01-311386, Date of Filling: 30.11.1989.

[13] S. Okhonin et al., "A SOI capacitorless 1T-DRAM concept," in *Proc. IEEE Int. SOI Conf.*, pp. 153–154, Oct. 2001.

[14] P. C. Fazan et al., "A simple 1-transistor capacitorless memory cell for high performance embedded DRAMs," in *Proc. IEEE Custom Integrated Circuits Conf.*, pp. 99–102, May 2002.

[15] P. C. Fazan et al., "Capacitorless 1-transistor DRAM," in *Proc. IEEE Int. SOI Conf.*, pp. 10–13, Oct. 2002.

[16] T. Ohsawa et al., "Memory design using a one-transistor gain cell on SOI," *IEEE JSSC*, vol.37, No.11, pp1510-1522, Nov. 2002.

[17] M. Blagojevic et al., "Capacitorless 1T DRAM Sensing Scheme With Automatic Reference Generation," *IEEE JSSC*, Vol. 41, No. 6, pp. 1463-1470, Jun. 2006

[18] T. Shino et al., "Floating Body RAM Technology and its Scalability to 32nm Node and Beyond," in the technical digest of *2006 IEDM*, Dec. 2006.

[19] T. Ohsawa, "SOI DRAM with One-Transistor FET Cell," *Oyo Buturi*, vol.75, No.9, pp1131-1135, 2006.

[20] H. Takato et al., "High Performance CMOS Surrounding Gate Transistor (SGT) for Ultra High Density LSIs," in the technical digest of *1988 IEDM*, pp.222-225, Dec. 1988.