Journal of Circuits, Systems, and Computers


# A Single-ended 28 nm CMOS 6T SRAM Design with Read-assist Path and PDP Reduction Circuitry

Chua-Chin Wang[†*], Zong-You Hou[†♭], Deng-Shian Wang[†♮], and Chia-Lung Hsieh[†©]
[†]National Sun Yat-Sen University, Department of Electrical Engineering,
70 Lian-Hai Rd., Kaohsiung, Taiwan
[*]ccwang@ee.nsysu.edu.tw
[♭]blazesky@vlsi.ee.nsysu.edu.tw
[♮]tonyhenry2@vlsi.ee.nsysu.edu.tw
[©]Chialung1218@vlsi.ee.nsysu.edu.tw

Received (Day Month Year)
Revised (Day Month Year)
Accepted (Day Month Year)

A Single-ended 6T SRAM cell composed of a 5T cell and a read-assist low $V_{th}$ PMOS as foot switch to prevent leakage damaging the data state is proposed in this work. Besides, a PDP (power-delay product) reduction circuitry design for nano-scale SRAMs is also proposed. The proposed PDP reduction circuitry design is composed of a AVD (adaptive voltage detection) circuit generating a boost enable signal if the process variation is over a pre-defined range and a HWB (half-period word-line boosting) circuit responding to the enable signal. The proposed SRAM is implemented using TSMC 28 nm CMOS logic technology. PDP reduction is verified to be 41.73% according to the measurement results. The energy per access is 0.0206 pJ given 800 mV power supply and 40 MHz system clock rate.

Keywords: single-ended SRAM cell; disturb-free; SRAM; power-delay product; word-line boosting

## 1. Introduction

SRAM has been recognized as an important role in many products, e.g., the cache of CPU. To extend the operation time and reduce power dissipation, SRAM is usually fabricated using advanced processes as long as they are available. However, as the technology evolving toward the nanometer scale, the leakage becomes a threat. The operation frequency and power consumption of SRAMs will be deteriorated because the leakage current increases exponentially with the drop of threshold voltage ($V_{th}$) and gate oxide thickness [1, 2]. Many prior works were proposed to resolve this low power dissipation issue. Certainly, the most straight forward approach is to carry out voltage scaling to reduce power dissipation. Another typical

*Corresponding author

example was to take advantage of Schmitt trigger positive feedback in the 10T SRAM, where the read stability is enhanced thanks to the extra 4 NMOS transistors [3]. Similarly, many researchers reduce the power dissipation and enhance the stability by paying the price of area overhead [4]. By contrast, a load-less 4T SRAM was proposed, where high $V_{th}$ devices are used as the storage latch and low $V_{th}$ devices are the driving control transistors [5]. This 4T SRAM demonstrated impressive 46.7% of leakage power reduction and 30% area reduction compared with traditional 6T SRAM at the same time. Nevertheless, the 4T SRAM cell stability was found to be threatened by the internal node with a floating voltage. That is, when "0" is the current data state, it could be flipped by accidental leakages or glitches.

To enhance the SRAM R/W stability, many circuities were proposed besides the change of SRAM cell structures, e.g., [6]. However, the endurance of voltage stress applied to the nano-scale transistors becomes an issue if nano-scale CMOS is used to fabricate the SRAM. Another approach was to take advantage of boosting wordline voltage [7]. The problems of this design are large die area demanded by the cell sitting in its own well and the stress endurance of the wordline transistors. Owing to lack of re-booting voltage design in this design, the wordline voltage might be dropped due to the leakage of the charging capacitor such that the stability is questionable.

To resolve all of the mentioned difficulties, the SRAM cell in this investigation is composed of ultra-high $V_{th}$ PMOS and ultra-low $V_{th}$ NMOS devices. Thus, not only the leakage is reduced, the driving current is also increased from the cell viewpoint. Most important of all, a low-$V_{th}$ PMOS device is added between the wordline access transistor and the 5T load-less SRAM cell to neutralize the severe leakage caused by the thin gate of the advanced CMOS process such that the data state of the cell will not be compromised and the access stability is ensured. Besides the addition of the low-$V_{th}$ PMOS device, a PDP reduction circuit consisting of AVD and HWB circuits is proposed to raise the logic "1" of the wordline signal from VDD to a higher VDDD for half of a cycle when the compensation is needed.

## 2. Single-ended 6T SRAM with Read-assist Path

The architecture of the proposed SRAM using single-ended load-less 6T cells is shown in Fig. 1, where two 1-Kb 6T load-less SRAM arrays are included. These two arrays share Column decoder, Row decoder, Control circuit, and BIST (built-in self test). The only difference is that the SRAM cells of the array on top consists of 6T cells, where a read-assist path is added in the cell, and the wordline control signals of the array, WL_c[31:0], are generated by the PDP reduction circuit to demonstrate the performance of the proposed compensation circuitry. Control circuit includes several MUXs which allow BIST_EN to select either the normal operation mode or the built-in-self-test (BIST) mode. WR_EN determines the read (= logic 0) or the write (= logic 1) is activated. Data state will be stored in the SRAM

array via Data_in. CLK represents the system clock. The output of the proposed design is Data_out_t, while the output of the traditional 6T load-less cell array design is Data_out [8]. Notably, PD[2:0] denotes the detected process corner and CS (Compensation Start) is the enable signals of the PDP reduction circuit. These signals could be generated by a PVT (process, voltage, temperature) detection circuit, e.g., [9], which is not shown in this work. Many of the blocks in Fig. 1 are realized by well-known circuits, e.g., Row and column decoders and Build-in Self-test (BIST), which will not be addressed in this investigation to avoid the waste of pages. By contrast, the details of Single-ended read-assist 6T Loadless SRAM Cell, PDP reduction circuit and Control circuit will be highly focused in the following text.

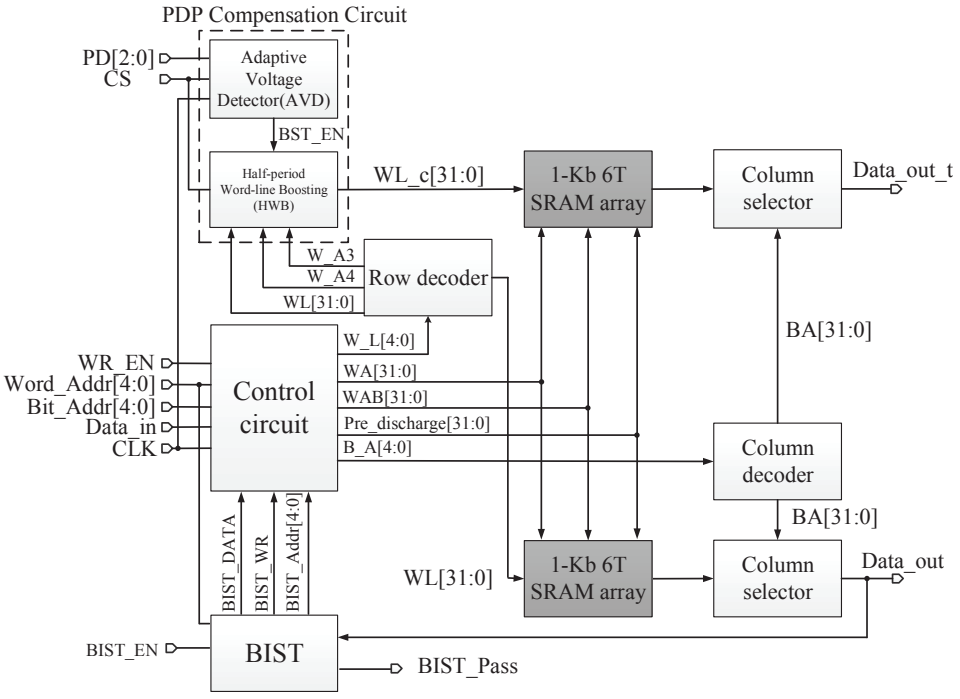


Fig. 1. Block diagram of the proposed SRAM

## 2.1. Single-ended read-assist 6T Loadless SRAM Cell

Fig. 2 shows a loadless 6T SRAM cell coupled with an inverter. To reduce the cutoff leakage current, high $V_{th}$ PMOS transistors, M101 and M102, consists of a latch-like storage. Low $V_{th}$ NMOS transistors, M103 and M104, are used as access switches. M105 is also a low $V_{th}$ NMOS transistor to supply large current

for driving the bitline (BLB). M105 is driven by WL (wordline) to access Qb such that the internal data state stored in Q won't be interfered by other cells coupled to BLB. The single-ended 6T structure is featured by M106 driven by the state of node Qb. Low $V_{th}$ PMOS transistor, M106, is added to provide an extra path to ground the leakage such that the retention fault is totally resolved.

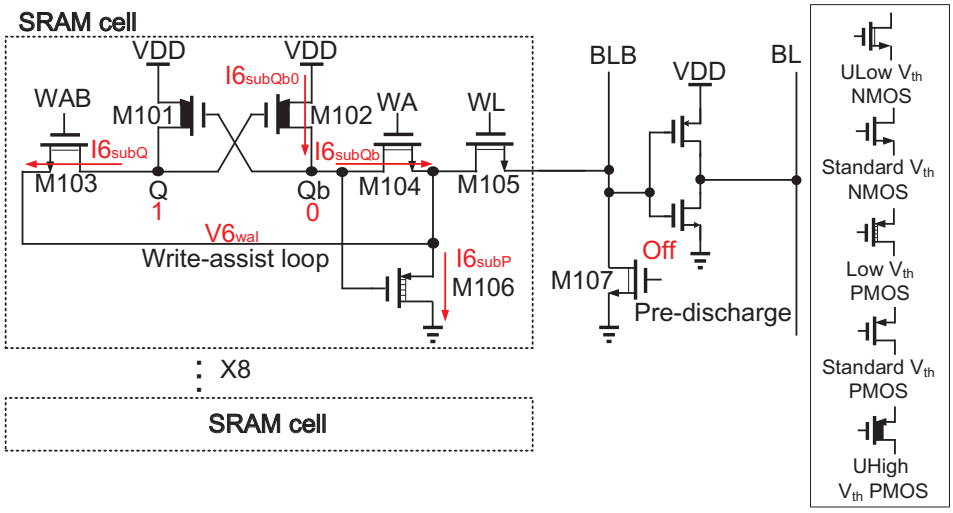


Fig. 2. Schematic of the proposed 6T SRAM cell

Referring to Fig. 2, all of the leakages at node Qb are as follows.
$I6_{subQ}$ = leakage current of M103
$I6_{subQb0}$ = leakage current of M102
$I6_{subQb}$ = leakage current of M104
$I6_{subP}$ = leakage current of M106

Apparently, the currents at the drain of M104 in the steady state are as follows.

$$I6_{subP} = I6_{subQb} + I6_{subQ} + I6_{leak,M105} \tag{1}$$

If M106 dose not exist, the entire cell will be the same the prior 5T cell in [10], which is suffered from the retention problem when "1" is managed to be written into the cell. Then, it will become a retention issue because now M104 is on due to WA = logic high. Qb = 0 will be contaminated without any protection mechanism. Thus, M106 is proposed to be added between the drain of M104 and ground. Thus, when Qb is stored with logic low, M106 is on to ground $V6_{wal}$ such that the stability is ensured and read-assist is achieved. Table 1 tabulates all the leakages of the PMOS with different $V_{th}$.

Notably, since subthreshold leakage is much larger than the gate leakage and

Table 1. Leakages of PMOS with different $V_{th}$

| Device | subthreshold leakage (pA) | gate leakage (pA) | reverse leakage (pA) |
|---|---|---|---|
| High $V_{th}$ | 19.48 | 0.48 | 6.70 |
| Low $V_{th}$ | 22790.00 | 0.27 | 22.50 |

reverse bias leakage, Fig. 3 shows only three subthreshold currents in M103, M014, and M106. $V6_{wal}$ of Fig. 2 is found graphically to be 0.038 V.

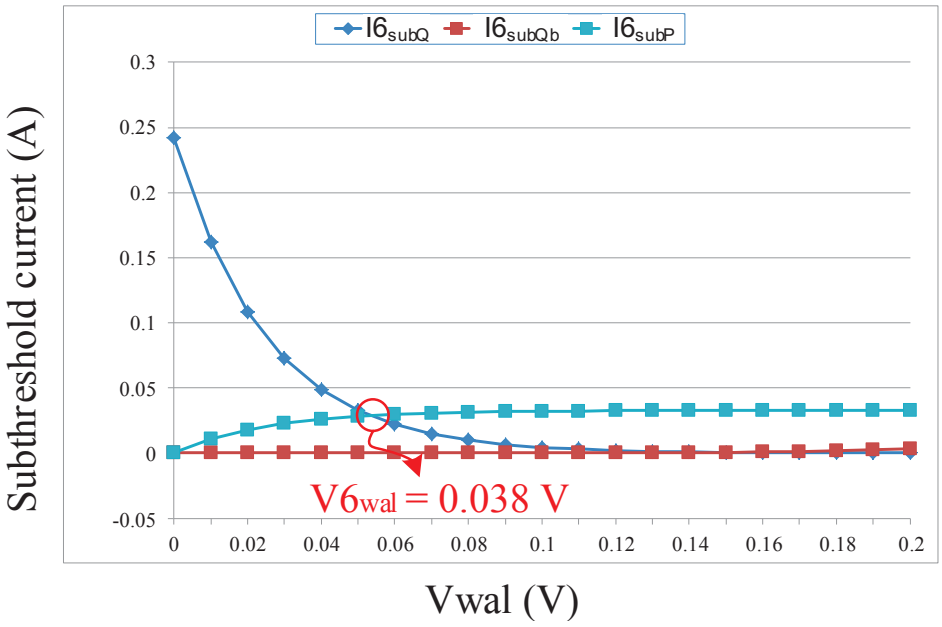


Fig. 3. Subthreshold leakage vs. $V6_{wal}$

Thus, Eqn. (1) is simplified as follows.

$$I6_{subP} = I6_{subQ} \tag{2}$$

The subthreshold current is known as the following equation.

$$I_{sub} = \mu_{eff} \cdot C_{ox} \cdot (m\text{-}1) \cdot \frac{1}{V_T^2} \cdot \frac{W}{L} \cdot e^{\frac{-V_{th}}{nV_T}} \left(1 - e^{\frac{-V_{wal}}{V_T}}\right) \tag{3}$$

$$\frac{\mathrm{W_{M103}}}{\mathrm{W_{M106}}} = \frac{\mu_{\mathrm{eff,P}} \cdot \mathrm{C_{ox,P}} \cdot e^{\frac{-V_{thP}}{nV_T}} (1 - e^{\frac{-\mathrm{V_{wal}}}{V_T}})}{\mu_{\mathrm{eff,N}} \cdot \mathrm{C_{ox,N}} \cdot e^{\frac{-V_{thN}}{nV_T}} (1 - e^{\frac{-\mathrm{V_{wal}}}{V_T}})} \tag{4}$$

where $\mathrm{C_{ox}}$ is the unit capacitance, $\mu_{\mathrm{eff}}$ is the device mobility, $V_T$ is the thermal voltage, $V_{th}$ is the threshold voltage of the device, $m$ is the process parameter, $W$ and $L$ denote the width and the length, respectively. In short, as long as $\mathrm{V6_{wal}}$ is known and all the other parameters in Eqn. (4) are given, the size ratio of M103 (and M104) vs. M106 can be derived to ensure the stability of Qb as well as Q.

### 2.2. PDP reduction circuit

PDP (power-delay product, i.e., energy) reduction circuit in Fig. 1 is shown in Fig. 4, where Adaptive Voltage Detector (AVD) and Half-period Wordline Boosting (HWB) circuits are included.

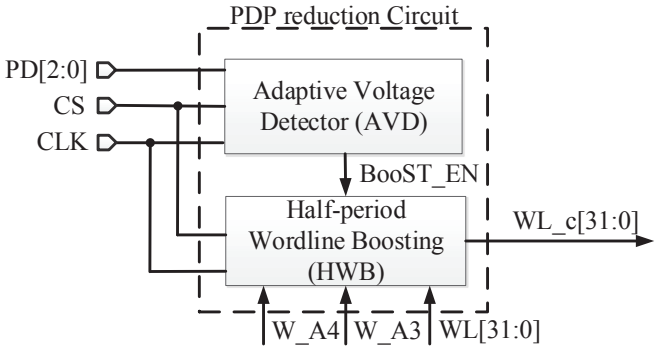


Fig. 4. PDP reduction circuit

#### 2.2.1. Adaptive Voltage Detector (AVD)

Referring to Fig. 5, the voltage at VP0 is determined by the total currents of paths, R201, R202, R203, which are controlled by 3 external signals PD[0]-PD[2], respectively, and the resistance of pulldown path composed of G-D-shorted M213 and M214, where M214 is driven by another external start signal, CS. VP0 is coupled to the gates of M216 and M217, which is an inverter clocked by CLK and CLKb with a state-flipping voltage at $\mathrm{V_{trip}}$. Thus, the clocked inverter compares VP0 with $\mathrm{V_{trip}}$ and then the resulted state, VP1 (and VP2), is kept by the clocked latch comprising inv201, inv202, and a transmission gate driven by CLK and CLKb. Finally, Boost_EN is generated at the NAND gate driven by CS and VP2.

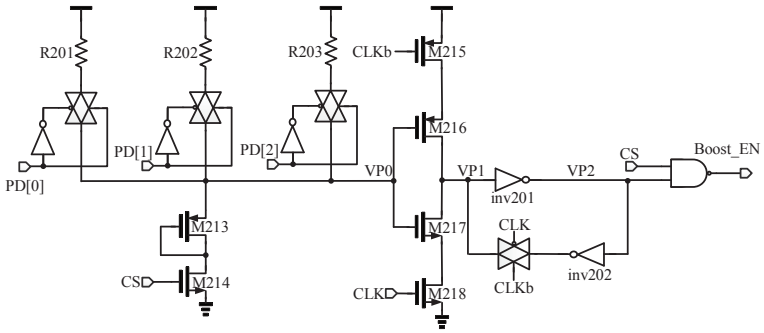


Fig. 5. Schematic of AVD

When VP0 is low, M213 is off. When VP0 rises to a certain voltage, M213 is driven into the saturation region. However, VP0 voltage will be clamped finally by the voltage drop of R201, R202, and R203, with the increasing saturation current of M213, as shown in Fig. 6 (a). The next thing to be resolved is the voltage of $V_{trip}$. Apparently, the power dissipation of transistors will exponentially increase as the supply voltage increases, as shown in Fig. 6 (b). It is found that 0.68 V is the voltage that the power consumption starts to soar exponentially such that it is used as point for the $V_{trip}$ and VP0 to cross each other in Fig. 6 (a). In other words, 0.68 V is used as the transition voltage for the size tuning target for the clocked inverter consisting of M215-M218.

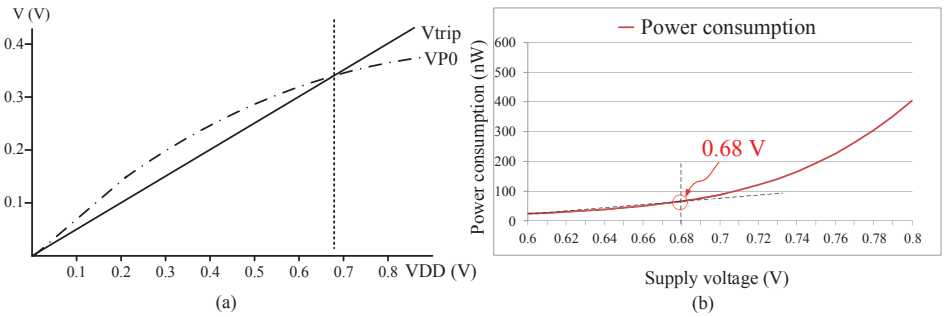


Fig. 6. (a) VP0 vs. $V_{trip}$ (b) Power dissipation vs. supply voltage

### 2.2.2. Half-period Wordline Boosting (HWB)

HWB in Fig. 5 is disclosed in Fig. 7. To reduce the loading of the output, all the 32 wordlines are equally divided into 4 groups. W_A3 and W_A4 are directly

coupled with Word_Addr[3]-[4].

1). When CS is low, M219 is on such that WL_bst = VDD. Meanwhile, C201 is not charged. C202 is not charged, either, regardless what Boost_EN is high or low.
2). Notably, regardless the state of Boost_EN, WL_bst will be raised to VDD+△V thanks to C201. However, when Boost_EN is high, it means the entire circuit is driven into a high-power mode, e.g., VDD ≥ 0.68 V, which results in the leakage of C201. This problem will not ensure that WL_bst is kept at VDD+△V. Thus, Boost_EN driving inv206 to turn on the transmission gate for C202 such that C202 is added as a parallel device to C201 to compensate the drop of the voltage thereof.

In short, the gate voltage of the M105, namely WL, will be raised to be higher than supply voltage. The boosting voltage not only speeds up the SRAM into the stable state, but also increases the subthreshold leakage current flowing to ground to reduce the impact caused by PMOS pair, namely M103 and M104. The timing diagram is illustrated in Fig. 8.

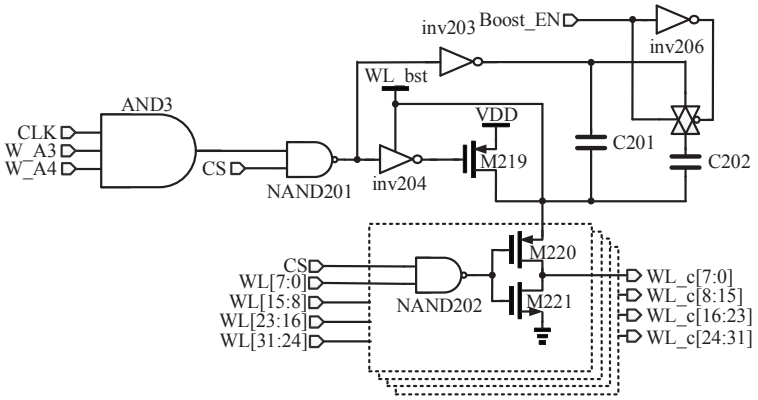


Fig. 7. Schematic of HWB

## 2.3. Control circuit

Referring to Fig. 1 again, Control circuit is apparently a traditional logic circuit with the function summarized as follows.

1). The memory address to be accessed is decoded by Word_Addr[4:0] and Bit_Addr[4:0], respectively, using row decoder and column decoder at the rising edge of the system clock.

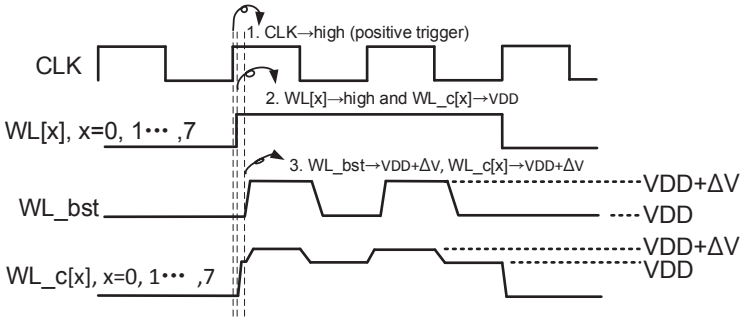


Fig. 8. Timing of HWB

2). The R/W operation of the selected memory bit is determined by the decoding of WR_EN and Data_in (data input), as tabulated in Table 2. Besides, the WA, WAB, and Precharge of unselected memory bits all are logic 0 at the same WL bit.
3). When BIST_EN is valid, Control circuit enters BIST mode, where several self-testing signals will be generated, i.e., BIST_Data, BIST_WR, and BIST_Addr[4:0].

Table 2. Control signals

| operation | Write0 | Write1 | Read | Standby |
|---|---|---|---|---|
| WL | 1 | 1 | 1 | 1 |
| WA | 0 | 1 | 1 | 0 |
| WAB | 1 | 0 | 0 | 0 |
| Precharge | 1 | 1 | 0 | 1 |

## 3. Implementation and Measurement

The proposed SRAM is carried out by using TSMC 28 nm CMOS process. The prototype of the proposed single-ended 6T SRAM is shown in Fig. 9, where the chip area is 600×534 $\mu m^2$. Fig. 10 shows the measurement setup, where the power supply (Agilent E3631Ax2) deliver 3 VDDs (VDDa, VDDb, VDDc) to Control circuit, SRAM array without PDP reduction, and SRAM array with PDP reduction, respectively. Function Generator (Agilent 81250) generates CLK, WE_EN, and Data_in. OSC (Agilent 54855A) displays the outcome waveforms, including Data_out and Data_out_t.

Referring to Fig. 11, which is an measurement snapshot of R/W functions, an

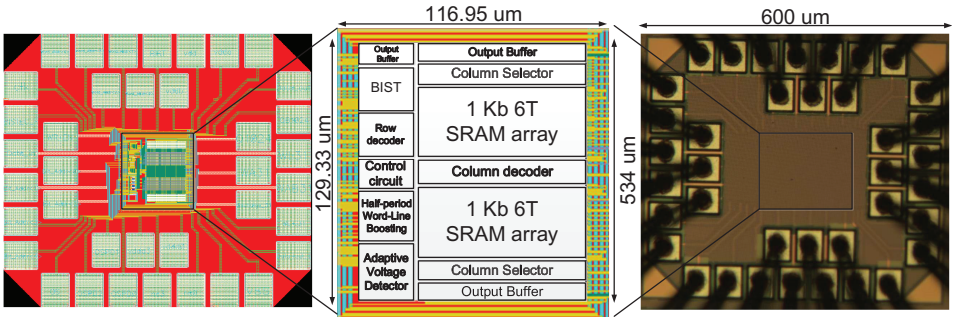


Fig. 9. Layout and diephoto of the proposed SRAM

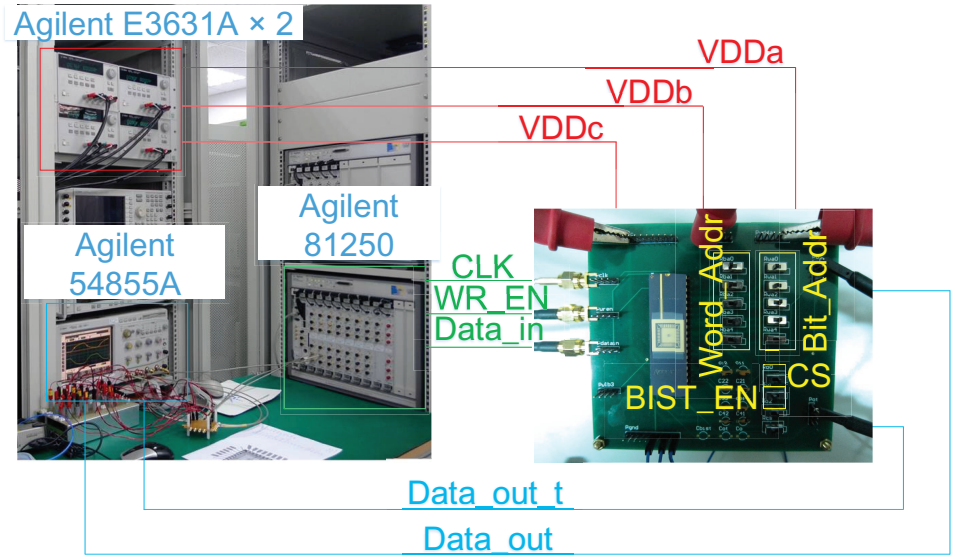


Fig. 10. Measurement setup

access sequence, W0→R0→W1→R1, is generated by the combination of WR_EN and Data_in.

During the first cycle, W0, outputs remains unchanged. In the second cycle, R0, both outputs are pulled down low. Then, they are pulled high in the third cycle, W1, due to the discharge. Lastly, the outputs are still kept high in the fourth cycle, R1. The above R/W operation was tested over 1000 times to justify the correct functionality. The PDP performance given different CLK rate, from 1 MHz to 40 MHz, are tabulated in Table 3. The reduction ratio is found to be about 40% no matter what clock rate is. A brief comparison with several recent SRAM works is

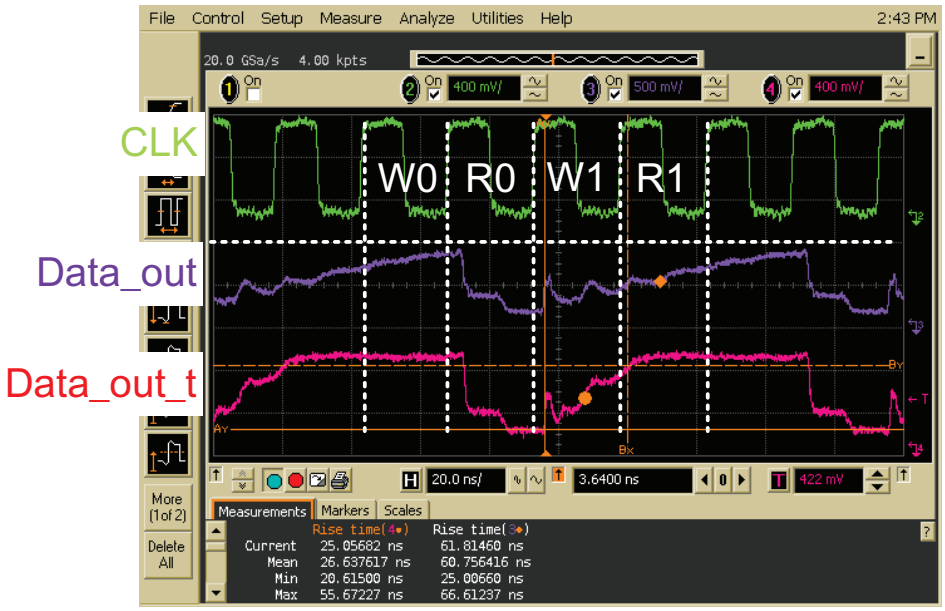


Fig. 11. 40 MHz waveforms

Table 3. PDP reduction performance

| CLK(MHz) | without PDP reduction | with PDP reduction | reduction ♭ |
|---|---|---|---|
| 1 | 28.7 fJ | 17.3 fJ | 39.67% |
| 4 | 91.6 fJ | 52.9 fJ | 42.22% |
| 15 | 320.2 fJ | 192.4 fJ | 39.93% |
| 30 | 694.8 fJ | 377.6 fJ | 45.65% |
| 40 | 755.6 fJ | 444.5 fJ | 41.17% |
| Average | | | 41.73% |

Note: ♭ reduction $= \frac{\text{without PDP reduction}-\text{with PDP reduction}}{\text{without PDP reduction}}$

summarized in Table 4. Our design attains the best FOM, and the second best operating frequency.

## 4. Conclusion

A novel single-ended 6T SRAM cell design with PDP reduction circuitry is presented in this paper, where an extra grounding PMOS transistor is added to

Table 4. Performance Comparison

| | TCAS-I [11] | TVLSI [10] | TVLSI [12] | ICECS [13] | ICETCCT [14] | TCAS-I [15] | This work |
|---|---|---|---|---|---|---|---|
| Year | 2014 | 2015 | 2015 | 2017 | 2017 | 2018 | 2018 |
| Technology | 40 nm CMOS | 40 nm CMOS | 22 nm FinFET | 180 nm CMOS | 130 nm CMOS | 28 nm CMOS | 28 nm CMOS |
| Cell Architecture | 12T | 5T | 9T | N/A | 5T | 10T | 6T |
| Capability (kb) | 4 | 4+1 | 32 | 0.0625(64 bits) | 1/1024(1 bit) | 32 | 1+1 |
| Read PDP (fJ) | N/A | N/A | N/A | N/A | N/A | N/A | 444.5@0.8 V |
| Frequency (MHz) | 11.5 | 54 | N/A | 100 | 50 | 0.03 | 40 |
| Standby Power ($\mu$W) | 22 | N/A | 2.1 | N/A | N/A | 0.465 | 7.5@0.8 V |
| Core Area ($\mu m^2$) | 134×132 | 136.5×181.2 | N/A | 1744×2078.44 | 3×7.7 | 271×394 | 116.95×129.33 |
| $\frac{\text{Normailized Area}}{\text{Cap}}$ ($\mu m^2$/bit) $^\flat$ | 2.7 | 3.01 | 2.19 | 1.79 | 1.4 | 4.256 | 9.646 |
| Energy/access (pJ) | 1.91 | 0.94 | 2.7 | 13.7 | 3.53 | 2.92 | 0.0206 |
| FOM$^\ddagger$ | 0.193 | 0.353 | 0.169 | 0.0408 | 2.023 | 0.805 | 5.033 |

Note: $^\flat$ $\frac{\text{Normailized Area}}{\text{Cap}} = 1000 \times \frac{\text{Core Area}}{\text{Capability}\times\text{Process}^2}$ ,

$^\ddagger$ $\text{FOM} = \frac{1}{\frac{\text{Energy}}{\text{access}} \cdot \frac{\text{Normailized Area}}{\text{Cap}}}$

resolve the retention problem of the prior 5T SRAM cell. Besides, to demonstrate the effect of the proposed PDP reduction design, two SRAM arrays, 1-Kb with the PDP reduction circuit and 1-Kb without, are realized on the same chip to make a fair comparison. By the measurement results, over 40% of PDP reduction (energy) per access is attained with the assistance of the PDP reduction circuit. Last but not least, the proposed loadless single-ended 6T SRAM is physically proved to attain the best FOM to date compared with recent SRAM designs.

Acknowledgement

This investigation is partially supported by Ministry of Science and Technology under grant MOST 107-2218-E-110-004- and MOST 107-2218-E-110-016-. The authors would like to express their deepest gratefulness to Taiwan Semiconductor Research Institute for their thoughtful chip fabrication service and EDA tool support.

References

1. F. Moradi, D. T. Wisland, S. Aunet, H. Mahmoodi, and T.-V. Cao, "65nm subthreshold 11T-SRAM for ultra low voltage applications," IEEE Trans. on System-on-Chip Conf., pp. 113-118, Sep. 2008.
2. M. Goudarzi and T. Ishihara, "SRAM leakage reduction by row/column redundancy under random within-die delay variation," IEEE Trans. on Very Large Scale Integration Systems, vol. 18, no. 12, pp. 1660-1671, Sep. 2009.

3. J. Kulkarni, K. Kim, and K. Roy, "A 160 mV robust Schmitt trigger based subthreshold SRAM," IEEE Journal of Solid-State Circuits, vol. 42, no. 10, pp. 2303-2313, Oct. 2007.
4. A. Islam and M. Hasan, "Leakage characterization of 10T SRAM cell," IEEE Trans. Electron Devices, vol. 59, no. 3, pp. 631-638, Mar. 2012.
5. C.-C. Wang, C.-L. Lee, and W.-J. Lin, "A 4-Kb 500-MHz 4-T CMOS SRAM using low-VTHN bitline drivers and high-VTHP latches," IEEE Trans. on Very Large Scale Integration (VLSI) Systems, vol. 12, no. 9, pp. 901-909, Sept. 2004.
6. H. Tanaka, M. Aoki, T. Sakata, S. Kimura, N. Sakashita, H. Hidaka, T. Tachibana, and K. Kimura, "A precise on-chip voltage generator for a gigascale DRAM with a negative word-line scheme," IEEE J. Solid-State Circuits, vol. 34, no. 8, pp. 1084-1090, Aug. 1999.
7. M. Bhargava, Y. Chong, V. Schuppe, B. Maiti, M. Kinkade, H. Y. Chen, A. W. Chen, S. Mangal, J. Wiatrowski, G. Gouya, A. Baradia, S. r. Thyagarajan, and G. Yeung, "Low VMIN 20nm embedded SRAM with multi-voltage wordline control based read and write assist techniques," in 2014 Symposium on VLSI Circuits Digest of Technical Papers, pp. 1-2, 2014.
8. Y. Singh, D. K. Pradhan, and S.-P. Mohanty, "A single ended 6T SRAM cell design for ultra-low-voltage applications," IEICE Electron. Exp., vol. 5, no. 18, pp. 750-755, Sep. 2008.
9. C.-C. Wang, W.-J. Lu, and T.-Y. Tsai, "Analysis of calibrated on-chip temperature sensor with process compensation for HV chips," IEEE Trans. on Circuits & Systems - II (TCAS-II), vol. 62, no. 3, pp. 217-221, Mar. 2015.
10. C.-C. Wang, D.-S. Wang, C.-H. Liao, and S.-Y. Chen, "A Leakage Compensation Design for Low Supply Voltage SRAM," IEEE Trans. on Very Large Scale Integration (VLSI) Systems, vol. 24, no. 5, pp. 1761-1769, Oct. 2015.
11. Y.-W. Chiu, Y.-H. Hu, M.-H. Tu, J.-K. Zhao, Y.-H. Chu, S.-J. Jou, and C.-T. Chuang, "40 nm bit-interleaving 12T subthreshold SRAM with data-aware write-assist," IEEE Trans. on Circuits and Systems I : Regular Papers, vol. 61, no. 9, pp. 2578-2585, Oct. 2014.
12. Y. Yang, J. Park, S. C. Song, J. Wang, G. Yeap, and S.-O. Jung, "Single-ended 9T SRAM cell for near-threshold voltage operation with enhanced read performance in 22-nm FinFET technology," IEEE Transactions on Very Large Scale Integration (VLSI) Systems, vol. 23, no. 11, pp. 2748-2752, Nov. 2015.
13. L. M. G. Rocha, G. Paim, G. M. Santana, B. A. Abreu, R. Ferreira, E. A. C. d. Costa, and S. Bampi, "Physical implementation of an ASIC-oriented SRAM-based viterbi decoder," IEEE International Conference on Electronics, Circuits and Systems (ICECS), pp. 526-529, Dec. 2017.
14. P. C. Shrivastava, P. Kumar, M. Tiwari, and A. Dhawan, "A novel approach for low voltage, low power deep sub-threshold 5-T SRAM cell," International Conference on Emerging Trends in Computing and Communication Technologies (ICETCCT), pp. 1-5, Dec. 2017.
15. Y.-C. Chien and J.-S. Wang, "A 0.2 V 32-Kb 10T SRAM with 41 nW standby power for IoT applications," IEEE Trans. on Circuits and Systems I : Regular Papers, no. 99, pp. 1-12, Jan. 2018.