



ORIGINAL RESEARCH

A 1.0 fJ energy/bit single-ended 1 kb 6T SRAM implemented using 40 nm CMOS process

Chua-Chin Wang^{1,2}  | Ralph Gerard B. Sangalang¹  | I-Ting Tseng¹ | Yi-Jen Chiu³ | Yu-Cheng Lin⁴ | Oliver Lexter July A. Jose^{1,5}

¹Department of Electrical Engineering, National Sun Yat-Sen University, Kaohsiung, Taiwan

²Institute of Undersea Technology, National Sun Yat-Sen University, Kaohsiung, Taiwan

³Department of Photonics, National Sun Yat-Sen University, Kaohsiung, Taiwan

⁴Department of Engineering Science, National Cheng Kung University, Tainan, Taiwan

⁵Department of Electronics Engineering, Batangas State University- The National Engineering University, Batangas City, Philippines

Correspondence

Chua-Chin Wang, Department of Electrical Engineering, National Sun Yat-Sen University, No. 70, Lian-Hai Rd., Kaohsiung City 80424, Taiwan.
Email: ccwang@ee.nsysu.edu.tw

Yu-Cheng Lin, Department of Engineering Science, National Cheng Kung University, No. 1, University Rd., Tainan City 70101, Taiwan.
Email: yuclin@mail.ncku.edu.tw

Funding information

The National Science and Technology Council of Taiwan, Grant/Award Numbers: MOST109-2218-E-110-007, 108-2218-E-110-011, 108-2218-E-110-002, 107-2218-E-110-002, 110-2221-E-110-063-MY2; National Applied Research Laboratories

Abstract

An ultra-low-energy SRAM composed of single-ended cells is demonstrated on silicon in this investigation. More specifically, the supply voltages of cells are gated by wordline (WL) enable, and the voltage mode select (VMS) signals select one of the corresponding supply voltages. A lower voltage is selected to maintain stored bit state when cells are not accessed, lowering the standby power. And when selecting a cell (i.e. WL is enabled) to perform the read or write (R/W) operations, the normal supply voltage is used. A 1-kb SRAM prototype based on the single-ended cells with built-in self-test (BIST) and power-delay production (PDP) reduction circuits was realised on silicon using 40-nm CMOS technology. Theoretical derivations and simulations of all-PVT-corner variations are also disclosed to justify low energy performance. Physical measurements of six prototypes on silicon shows that the energy per bit is 1.0 fJ at the 10 MHz system clock.

KEYWORDS

digital integrated circuits, logic design, low-power electronics, memory architecture, VLSI

1 | INTRODUCTION

Memory devices are known to be second only to CPU/MPU in terms of overall timing parameter performance of digital subsystems in electronic products. It will soon occupy 90% of the entire area in SOC (system on chip) according to the ITRS report [1]. Low-power memory devices will cut the total power consumption of these items, especially those that are battery powered and portable. Unlike DRAM, which is widely used as the main memory mechanism, SRAM has been used in most of CPU/MPUs as cache devices to speed up the access of recently used data. The performance of SRAM, regardless of its usage, has a significant influence on the power dissipation, which affects the overall efficiency of the system. Thus, it is

critical that SRAMs used in CPU/MPU have an energy saving feature to minimise their effect on its performance. Several variety of SRAM designs have been reported in past decades. Three key design methods were presented, specifically for the energy-saving and power demands of SRAMs:

- 1). Current-mode sense amplification [2, 3]: Since CMOS technology has been scaled down very quickly, the bitline capacitance has become too large for an SRAM cell to drive. During read operations, a sense amplifier pre-determines the output result by sensing the differential current on two bitlines, enabling high-speed and low-power operation. In this case, the bitline capacitance has less of an impact on the output delay.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2023 The Authors. *IET Circuits, Devices & Systems* published by John Wiley & Sons Ltd on behalf of The Institution of Engineering and Technology.

- 2). Secondary supply [4]: Adding an extra supply voltage, higher than the nominal VDD, can improve the access speed of the SRAM at the expense of a higher energy cost. The energy or standby power of the cells that are not accessed are often ignored.
- 3). Current compensation [5, 6]: When the SRAM is turned on, leakage current is detected by a current compensation circuit in each bitline to inject an additional current into the associated bitline. Thus, the SRAM's access speed is enhanced, even if the leakage current is not reduced. This method has little energy-saving benefits, especially in standby cells.

A 4T loadless SRAM has been reported for lower power usage, where high-threshold voltage transistors are used in data latches and low-threshold voltage transistors are used in bitline drivers [7], also called a P-latch N-drive 4T SRAM cell. Despite having self-refreshing paths to keep the stored bit state, instability and read/write disturbance become a threat to the access operations. These threats are particularly strong when loadless designs lacks any bitline isolation mechanism. What's even worse is that the weakening of the static noise margin (SNM), as mentioned in the works by Wang *et al.* [8], proved to be a hazard for such a cell structure. These SRAMs were found to be even more vulnerable when the supply voltage is lower. To resolve this issue, readout assist circuits were proposed for single-ended SRAMs [9]. This breakthrough accelerated the research of non-symmetrical R/W auxiliary circuit designs that are meant to provide disturbance isolation from bit-lines, for example, reports of Chen *et al.* [10] The disturbance isolation design becomes even more critical if the SRAM is meant to be fabricated using advanced CMOS technology nodes (e.g. <100 nm), or the SRAM cell is operating near the subthreshold region. The SRAMs that have write-assist loops were typical examples to demonstrate the disturbance-free feature [10, 11]. The two examples, however, used the design methodology for symmetrical R/W, and hence cannot be applied to single-ended SRAM cells. Other methods, such as the usage of asymmetric

Schmitt-trigger inverters, are also possible to improve cell performance [12].

In order to meet the low power dissipation demands for SRAMs implemented in advanced CMOS process, a supply voltage gate-control mechanism for every column of SRAMs is proposed in this work, wherein two supplies with different voltages are used and selected by WL (wordline) and associated signals to decrease the power dissipation while on standby. To compensate for the loss of R/W speed caused by reduced supply voltage, a voltage boost is given to the driving gate of the selected SRAM cells to provide speed and slew rate improvement. Detailed post-layout simulations and physical on-silicon measurements are demonstrated to justify the low power/energy feature. The proposed SRAM is fabricated using a typical 40-nm CMOS process, where 1.0 fJ energy/bit is measured at a 10 MHz system clock with an access delay of 52 ns.

2 | LOW ENERGY SRAM WITH SINGLE-ENDED CELLS

Referring to Figure 1, the proposed SRAM design consists of a memory array, a control circuit, a row and column decoder, a column select circuit, a build-in self-test (BIST) circuit, a power-delay product (PDP) reduction circuit, and a V_{DD} select circuit. The supply voltage of the SRAM cells is selected by the V_{DD} select circuit. A pass-transistor gate voltage boosting (PVB) and adaptive voltage detector (AVD) circuits make up the PDP reduction circuit. The functions of major signals in the proposed SRAM are summarised as follows:

1. Bit_Addr[4:0]: bitline addresses
2. Word_Addr[4:0]: wordline addresses
3. WR_EN: write/read enable (1/0)
4. Data_out, Data_in: data output and data input, respectively
5. CLK: system clock
6. BS: boost select

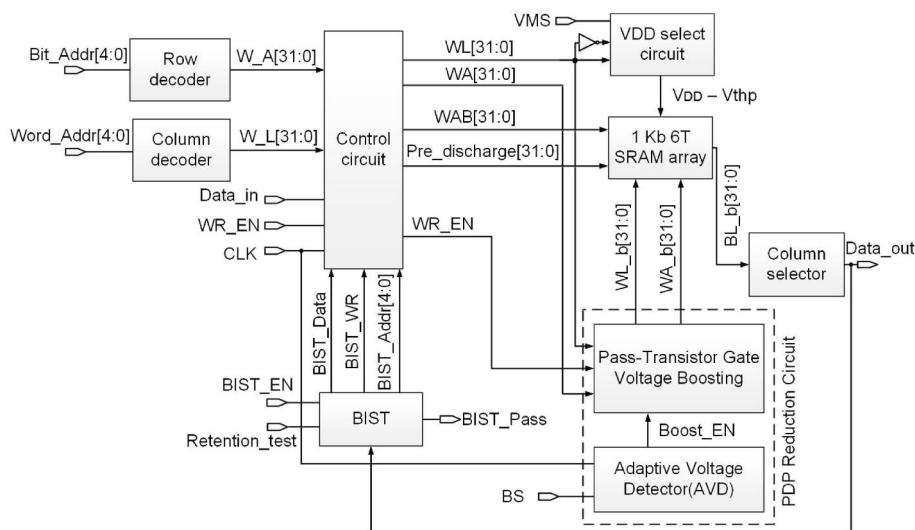


FIGURE 1 1-kb SRAM with single-ended cells system block diagram

7. VMS: voltage mode select
8. BIST_Pass: BIST pass (or not)
9. BIST_EN: BIST enable

2.1 | Single-ended SRAM cell circuit analysis

Works related to the reduction of leakage has been introduced in 8T SRAM cells through the use of low- V_{th} auto-gating transistors to minimise leakage at the expense of lower speeds [13]. Meanwhile in the works of Chen *et al.* [10], a 5T single-ended cell was reported, which introduced a cell isolation mechanism to prevent noise interference. However, leakage current was a main issue, causing retention fault, particularly for advanced CMOS processes. This issue was not present in the works of Terada *et al.* [14], because they introduced a transistor in their design to act as a leakage bypass. This was at the expense of a larger area overhead.

All the SRAM cells discussed before still exhibit high standby power, since all of the idle cells are directly coupled to regular power, thus consuming a significant amount of standby power. A new cell-column structure was presented in our previous report [15], as presented in Figure 2. This was to reduce the standby power, and in turn the overall power dissipation, when most cells are not accessed. Though the cell with the associated power-gated mechanism was described in the mentioned article, operation details, theoretical analysis, and the on-silicon verification were never disclosed. The proposed design operates as follows:

- 1). In the event that any cell is being accessed, WLB goes low and WL goes high. Transistor M_{301} is then turned on, providing the regular V_{DD} to the cells in the same column.

- 2). If all the cells in the column were not being accessed, WLB goes high and WL goes low. A lower supply voltage, $V_{DD} - V_{thp}(M_{306})$, is coupled to all cells' supply nodes in the same column. That is, the voltage supplied to the cells are decreased by the threshold voltage of transistor M_{306} to save power and still maintain the stored bit states.

Let us consider a typical CMOS process for our SRAM design vehicle. The supply voltage V_{DD} is 0.8 V for the typical 40-nm CMOS process such that the reduced supply voltage for those unaccessed cells becomes $V_{DD} - V_{thp} = 0.68$ V. Referring to Figure 2, the current supplied through the low- V_{th} PMOS devices is limited by the width thereof. Thus, auxiliary circuits driven by a VMS signal are needed to prevent possible R/W errors caused by the insufficient supply current.

- Access operation: $WL = 1$ and $WLB = 0$. $VMS1 = VMS (= 0)$ or $WLB = 0$ such that M_{305} is turned on to supply extra current.
- Hold operation: $WL = 0$ and $WLB = 1$. $VMS2 = VMS (= 0)$ or $WL = 0$ such that M_{310} and M_{315} are on to supply extra current.
- No auxiliary: $VMS = 1$. All of the auxiliary circuits are off. This is for the purpose of testing to validate the proposed power-gated mechanism.

2.2 | SRAM cell transistor sizing

The transistor sizes of the SRAM cells are determined by the current that will pass through the transistors for every operation. To attain reliable Q and Q_b with a symmetric feature, currents through M_{201} and M_{202} must be the same. Hence, $(W/L)_{201} = (W/L)_{202}$. Transistor M_{206} will drain the current when $Q = 0$ such that $(W/L)_{206}$ is chosen to be the minimum size to have the lowest current passing to it. The write-assist transistor M_{203} should be able to draw the same current passing to M_{201} when writing logic '1'. Transistors M_{203} and M_{204} are equally sized to have the ratio of M_{201} versus M_{203} equal to that of M_{202} versus M_{204} . Access transistor M_{205} is chosen to be equal to M_{203} versus M_{204} , since it should have current passing to them while doing write operation, which is equal.

To further decrease power consumption, transistors inside the cells are chosen to have high- V_{th} , while the write-assist and access transistors are chosen to have low- V_{th} to have faster access and write operations.

2.3 | Analysis of power dissipation and area overhead

Every column of memory cells in the proposed design has a power-gated mechanism. The biggest challenge is to optimise the size of the gating transistors to retain the correct operating region and reduction of power dissipation at the same time. Awkward scenarios at the FS corner (fast NMOS, slow PMOS) might be a problem of this power-gated mechanism, since slow

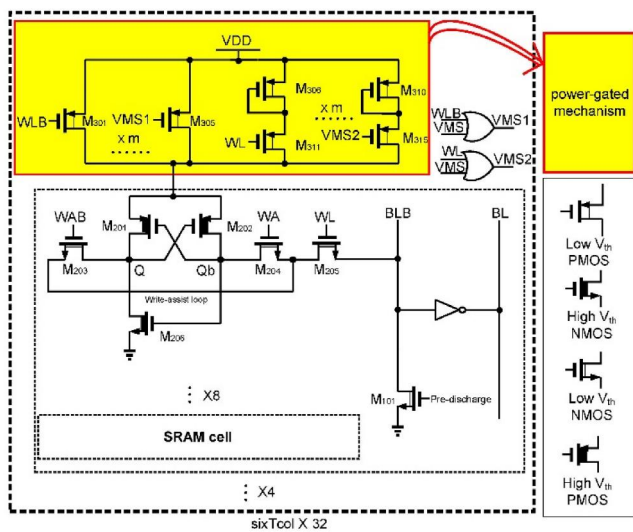


FIGURE 2 Proposed ultra-low power 6T SRAM cell with power gating

PMOS devices provide weak current. The expected result is that it will be hard for Q_b to stay high (as $Q = 0$). This current shortage issue can be overcome through analytical solutions to know what the minimal PMOS size should be provided that the number of cells is given. Assume that every column has a total of n cells. Assume that I_{act} stands for the current needed by the cell to access, and I_{idl} denotes the required current by idle cells. The power PMOS device's drain current should satisfy the following equation:

$$I_D \geq I_{act} + (n - 1) \times I_{idl} \quad (1)$$

If $n = 32$ and the implementation is on typical 40-nm CMOS technology, the total required current for cells being accessed in a column is $I_D = 39.5 \mu\text{A} = 31.5 \mu\text{A}$ (1 cell accessed) + $31 \times 245 \text{ nA}$ (other 31 cells idle). According to the saturation current I_D equation for MOS transistors, the minimal width of the power PMOS to supply such a current is $3.75 \mu\text{m}$. Therefore, a total of $5 \times 750\text{-nm}$ PMOS transistors in parallel are used, which are $M_{301}\text{--}M_{305}$. Similarly, the number and the size of those power-gated transistors for idle cells in Figure 2, namely $M_{306}\text{--}M_{315}$ can be determined as well.

Based on the proposed power-gated mechanism to decrease the standby power for most of the cells, the layouts of the proposed ultra-low power 6T SRAM cell and the power-gated mechanism circuit are shown in Figures 3 and 4, respectively, where the area of the single cell is $1.8 \times 2.1 \mu\text{m}^2$, and that of the power-gated mechanism circuit is $2.8 \times 3.6 \mu\text{m}^2$. Namely, the area overhead cost is $8.33\% = \frac{2.8 \times 3.6}{32 \times 1.8 \times 2.1}$ for every 32 cells, if the area of wiring is not taken into account, to share 1 power-gated mechanism circuit. If the length of the cell array is increased to 1024, the area penalty becomes only 0.26%, which is considered relatively small and negligible.

Referring to Figure 5, the layout size comparison between the proposed cell to a conventional 6T cell, both cells have similar transistor sizes, and the proposed cell occupied 21% larger area compared to a conventional cell layout, since the proposed cell

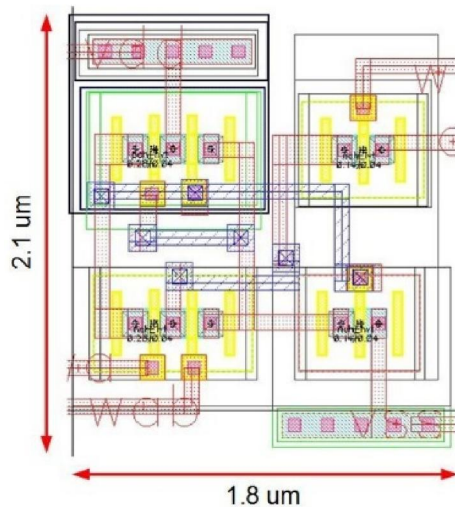


FIGURE 3 Layout of the proposed ultra-low-power 6T SRAM cell

used HVT and LVT devices such that the shared diffusion layout (as what is done for the conventional cell) cannot be utilised to minimise the area of the cell.

The proposed ultra-low-energy 6T single-ended SRAM cell design with the power-gated mechanism does not need a sense amplifier (SA) like prior single-ended SRAMs [10, 14]. In contrast, traditional SRAMs [13] need SAs to accelerate access operations, because they use the bitline and bitline together to access the data nodes. This is another reason why the proposed SRAM has to be more energy efficient.

2.4 | Read/write cycles

Read/write cycles of the proposed SRAM are described as follows. The read cycles are shown in Figure 6. The read operation in the cell is shown in Figure 7.

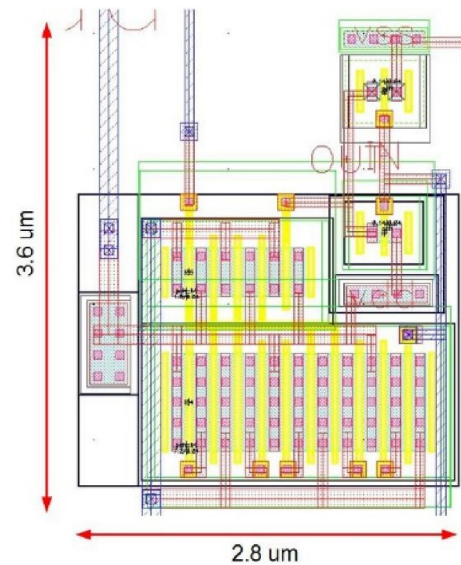


FIGURE 4 Layout of the proposed power-gated mechanism

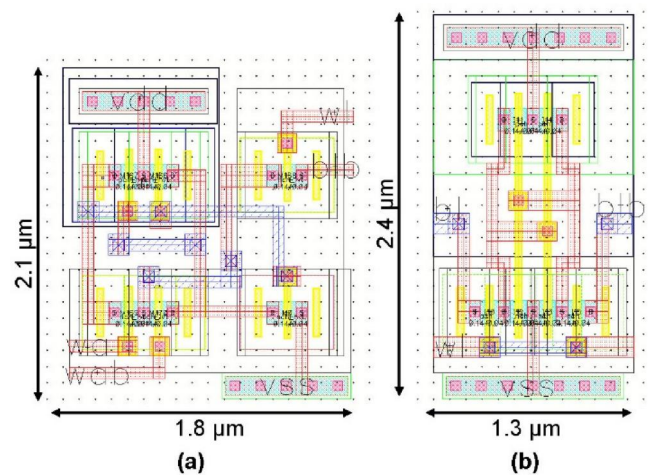


FIGURE 5 Layout size comparison (a) proposed 6T cell; (b) conventional 6T cell

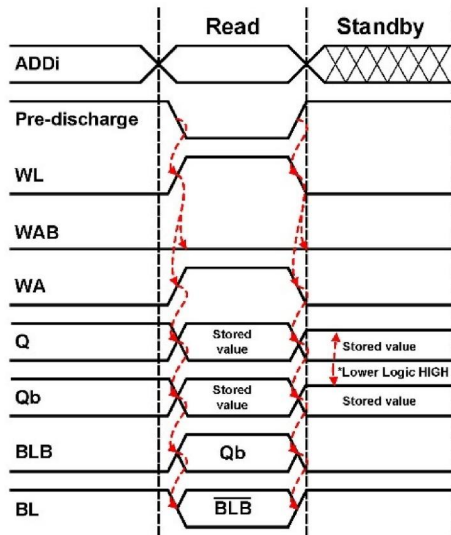


FIGURE 6 Read cycle timing diagram

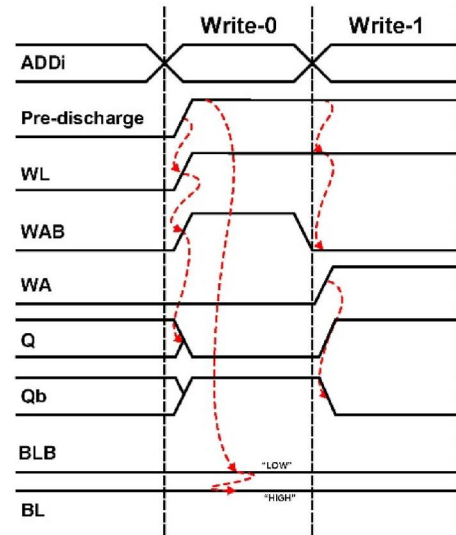


FIGURE 8 Write cycle timing diagram

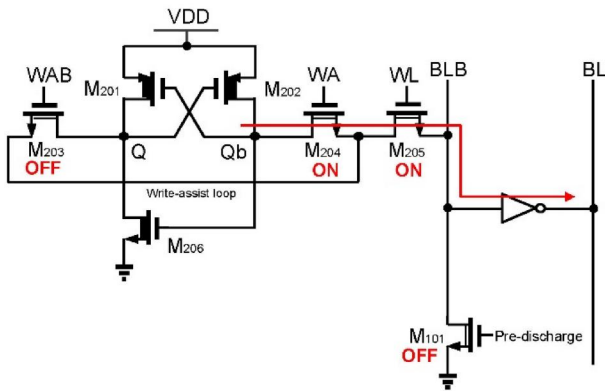


FIGURE 7 Read operation

- Before any operation, pre-discharge will cause BLB to be grounded to prevent state “0” from being disrupted by leakage and noise.
- The matching decoders choose the cell once if the row and column addresses are available.
- WA and WL are then pulled high, turning M_{204} (M_{203}) and M_{205} on. Regardless of Read1 or Read0, WAB is pulled low to turn M_{203} off. Qb will then be coupled to BLB through M_{205} and M_{204} .

The write operation timing diagram is shown in Figure 8. The write-0 and write-1 operations are shown in Figures 9 and 10, respectively.

- Write-0: WA is set to low and WAB is then pulled high to turn transistors M_{203} and M_{204} on and off, respectively. Q is then pulled down to the ground using the pre-discharge signal.

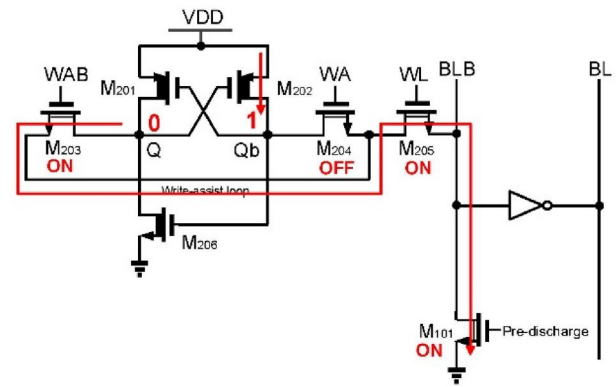


FIGURE 9 Write-0 operation

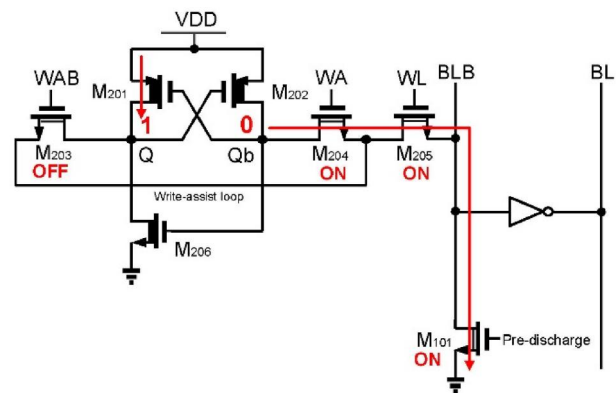


FIGURE 10 Write-1 operation

- Write-1: WA is pulled high to turn transistor M_{204} on. WAB is low to turn transistor M_{203} off. Pre-discharge then pulls Q high by pulling Qb down.

Table 1 tabulates the overall R/W operation and related control signals of the proposed SRAM.

2.5 | Hold/standby operation

The hold/standby operation is also shown as part of the read cycle in Figure 6. All the access transistors (M_{203} , M_{204} , and M_{205}) are disconnected to isolate the memory cell. During this operation, the power gating circuit is also enabled to reduce the supply voltage of the inactive cells. This is shown in Figure 11. A reduced voltage of around 0.68 V will now be used to supply the cells. This ensures the low power standby operation of the cells.

Since the low- V_{th} access transistors are used in the design, it is important that these transistors do not leak in the worst process corner. A long transient simulation during hold operation is presented in Section 3 to show that the design will not droop during the worst-case corner.

2.6 | PDP reduction circuit

Aside from the power-gating method used in the preceding sections for each column of cells, a power-delay product (PDP,

TABLE 1 Read and write operation

	Write 1	Write 0	Read (1/0)	Standby
Predischarge	1	1	0	1
WL	1	1	1	0
WA	1	0	1	0
WAB	0	1	0	0
BL	1	1	1/0	1
BLB	0	0	0/1	0

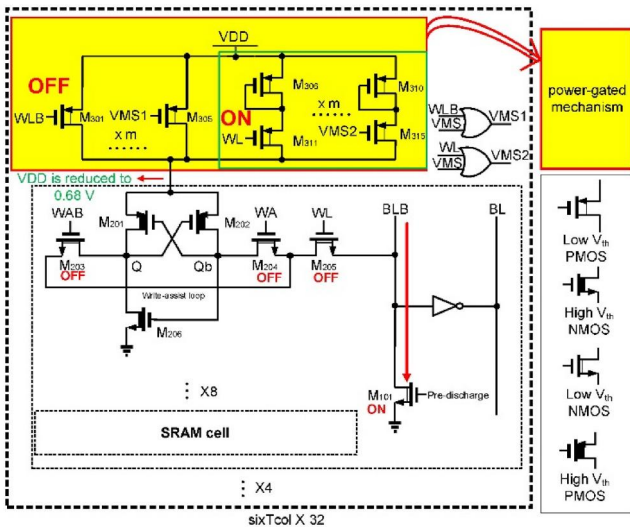


FIGURE 11 Standby/hold operation

namely “energy”) reduction circuit is used to further minimise energy consumption in each R/W operation [11, 15, 16]. Referring to Figure 12, the adaptive voltage detector and pass-transistor gate voltage boosting are the two sub-circuits of the PDP reduction circuit [11]. For high-speed access operations, when boost select (BS) is set to high, the AVD circuit gives a boost enable (Boost_EN). This changes the voltage supply of the cells to be accessed from V_{DD} to V'_{DD} (a voltage greater than V_{DD}).

2.6.1 | Adaptive voltage detector (AVD)

Referring to Figure 13, the adaptive voltage detector circuit used for generating the boost enable signal for the pass transistor voltage boosting circuit is presented. A common source amplifier (composed of M_{1301} , M_{1302} , and R_{1301}) generates the VP0 signal, once the BS signal is enabled, which is then fed into the current-starved inverter composed of M_{1304} and M_{1305} , and has a precise switching voltage to adjust to slight variations in the BS signal. The inverter's output is then latched to keep track of whether the pass transistor gate boosting circuit has to be enabled or disabled. The Boost_EN signal will be high if the output of inv_{1303} and the latched voltage VP2 are both low.

2.6.2 | Pass-transistor gate voltage boosting (PVB):

The PDP reduction circuit initially operates in the waiting mode. It will stay in the waiting mode, since the AVD circuit has not yet finished the system voltage detection (the inverter

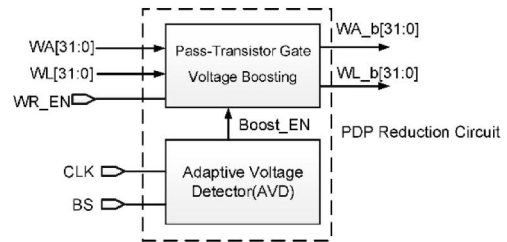


FIGURE 12 PDP reduction circuit

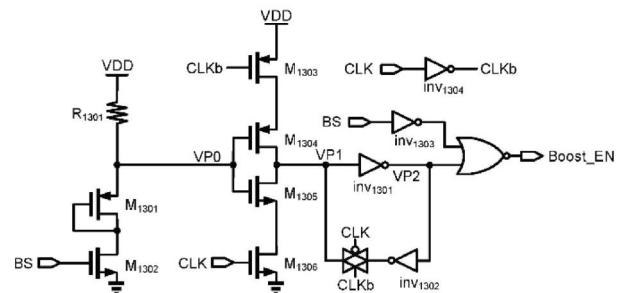


FIGURE 13 AVD circuit

switching voltage vs. VP0). The PDP reduction circuit enters the PDP reduction mode after exiting the waiting mode. The operation is as follows, as seen in Figure 14; C₁₄₀₁'s top plate will be pulled down to the ground by inv₁₄₀₃, while the bottom plate will be pulled up to VDD by M₁₄₀₁. On the other hand, as soon as Boost_EN is pulled up high, the PVB circuit starts to work. If WR_EN is high, meaning one of the SRAM cells is being accessed (read or write), the PVB circuit enters the voltage boosting mode, turning M₁₄₀₁ off. Then the top plate of C₁₄₀₁ is pulled to a higher voltage through the pull-up circuit of inv₁₄₀₃. Now, the supply voltage V_{DD} is at a higher level than when the PDP is on the waiting mode, V_{DD} = V_{DD} - V_{DS1401}. This now has a value V_{DD} = V_{DD} + ΔV, which in this 40-nm CMOS process is 1.0 V. The illustrative timing diagram of the PVB circuit is presented in Figure 15.

2.7 | Built-in self-test (BIST)

In every memory system, as shown in Figure 1, a BIST circuit is essential for high reliability. The BIST circuit block diagram is shown in Figure 16. It is composed of a pattern generator, a controller, and an output response analyser. The BIST circuit implements the control and output response pattern based on the March C-algorithm [17], which has moderate complexity and fault coverage. It was tested for transition faults, address-decoder faults, stuck-at faults, coupling faults, data retention faults etc. The March C-algorithm, which is the most reliable

of BIST circuits, has a complexity of 10×N, where N specifies the memory size, which in this study is 1024. The March C-algorithm, as shown in Eqn. (2), is outlined as follows:

$$\{\updownarrow(w0); \uparrow(r0, w1); \uparrow(r1, w0); \downarrow(r0, w1); \downarrow(r1, w0); \updownarrow(r0)\} \tag{2}$$

where *w* represents write access operation, *r* means read operation, \updownarrow represents either up or down count, \downarrow represents down counts, and \uparrow represents counting up. The timing diagram for the BIST is shown in Figure 17; it has two testing modes, a normal testing mode and a retention testing mode. A linear feedback shift register (LFSR) pseudo-random number generator is used to generate the test patterns with a characteristic equation as follows:

$$f(x) = x^5 + x^4 + x^3 + x + 1 \tag{3}$$

3 | CHIP IMPLEMENTATION AND MEASUREMENT

The prototype SRAM was implemented using a TSMC 40-nm CMOS process. Figure 18 shows the layout of the entire system, with the floor-plan at the right-hand side, where the overall chip area is 0.276 mm² and the core area is

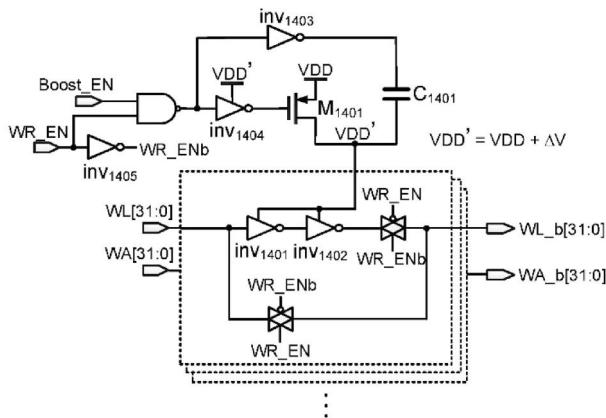


FIGURE 14 Pass-Transistor Voltage Boost (PVB) circuit

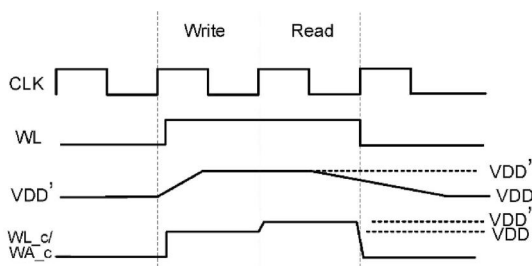


FIGURE 15 Gate drive boosting timing diagram

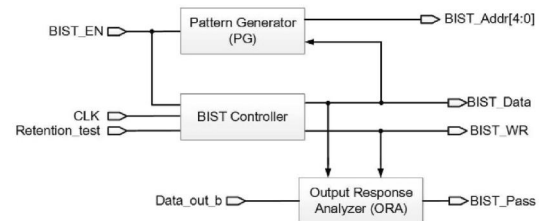


FIGURE 16 BIST block diagram

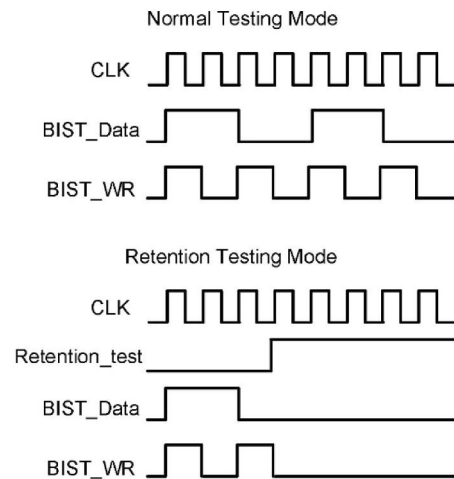


FIGURE 17 BIST timing diagram

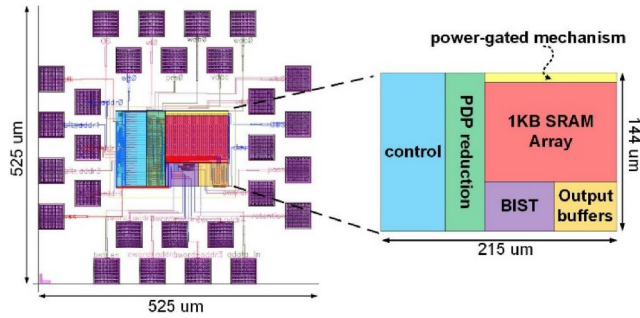


FIGURE 18 Layout of the proposed 1-kb SRAM

0.031 mm². The validation of the design is carried out in the following sections: post-layout simulations and on-silicon measurements.

3.1 | All-PVT-corner post-layout simulations

A post-layout all process, voltage, and temperature (PVT) variations simulation is required before the system is fabricated on silicon. The corners used in the post-layout simulations are as follows: 5 process variations (TT, SS, SF, FS, and FF), 3 supply voltage variations ($0.9 \times V_{DD}$, V_{DD} , and $1.1 \times V_{DD}$), and 3 temperature variations (0, 25, and 75) °C. Monte-Carlo simulations were executed a total of 100 times. All tests performed showed the correct functionality of the system. It was also ensured that there is high speed access when the AVD circuit is activated.

Two of the most important measures for the quality of SRAM operations are the static and dynamic noise margins. The static noise margin of the proposed SRAM cell is assessed by turning off the write-assist loop and turning on M_{604} , M_{605} in Figure 2. Then, the voltage on BLB is changed from logic “0” (GND) to logic “1” (V_{DD}). The curves of the induced voltages on Q versus Qb are plotted, where the biggest square in the generated transfer functions diagram is the SNM of this cell. Figure 19 shows the static noise margin of the cells showing the worst case of 412.3 mV. The static noise margin has an asymmetrical shape, different from traditional differential SRAM cells, since the design employed single-ended architecture.

For the dynamic noise margin (DNM) assessment, it can be found by using pulses with varying amplitude and pulse width applied at node WL to find out if the state of node Q is compromised. The dynamic noise margin of the designed SRAM cell is shown in Figure 20, which shows that V_{DD} for the proposed SRAM cell to operate correctly can be as low as 0.3 V.

Figures 21 and 22 show all the PVT (process, voltage, and temperature) corners simulations when the AVD is disabled and enabled. It can be seen that when the AVD circuit is enabled, the access speed of the SRAM has been improved.

Post-layout simulation shows that the standby power goes down from 17.734 to 3.432 μ W, which is by 80.65% reduction.

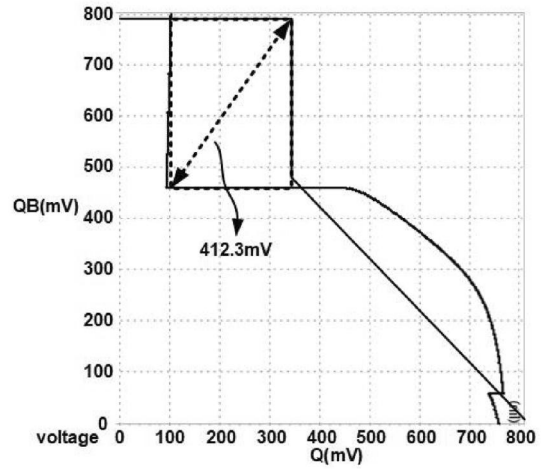


FIGURE 19 Static noise margin

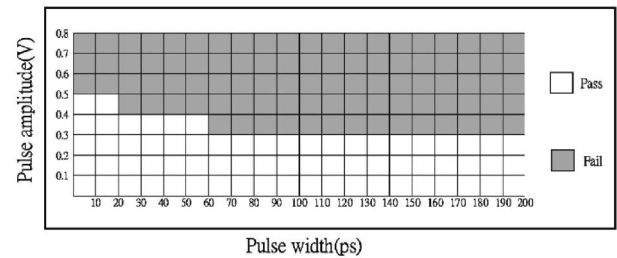


FIGURE 20 Dynamic noise margin

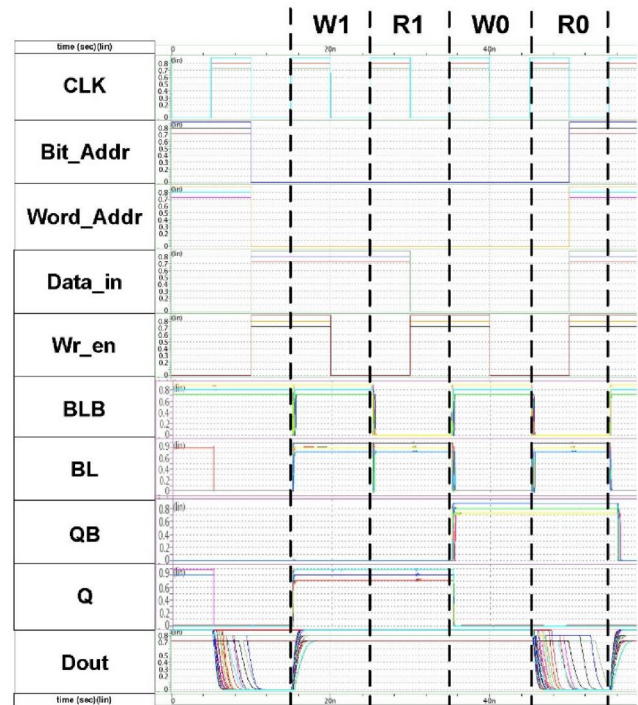


FIGURE 21 Simulations with AVD circuit disabled

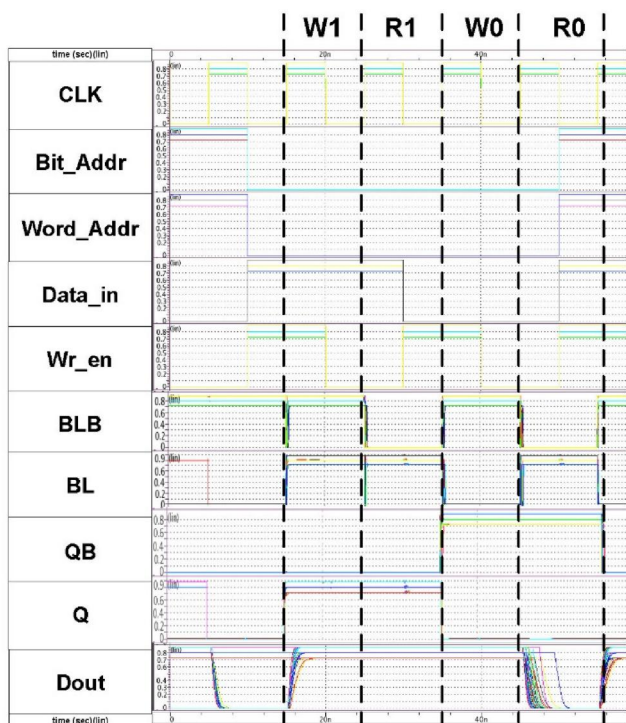


FIGURE 22 Simulations with AVD circuit enabled

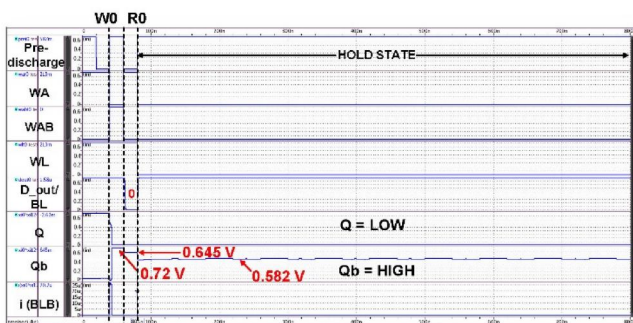


FIGURE 23 Hold state transient simulation in the worst corner (SS corner, $V_{DD} = 0.72$ V, $T = 0$ °C)

Referring to Figure 23, the post-layout simulation of the proposed design in the worst corner (SS corner, $V_{DD} = 0.72$ V, and $T = 0$ °C) is presented. A data bit is written in the cell first, then the data is read, and afterwards, the cell is placed on its hold state. The simulation run shown is 800 ns with the cell in the hold state for over 710 ns. It can be seen that the data bit written in Qb remains at the same value after a long standby state; hence there is no droop in its content. It also shows the lowered voltage level for the cell during standby operation.

3.2 | On-chip measurements

The die photo of the proposed SRAM array (1 kb) is shown in Figure 24. The details are hard to be observed because the prototype is covered by metal layers due to the minimum

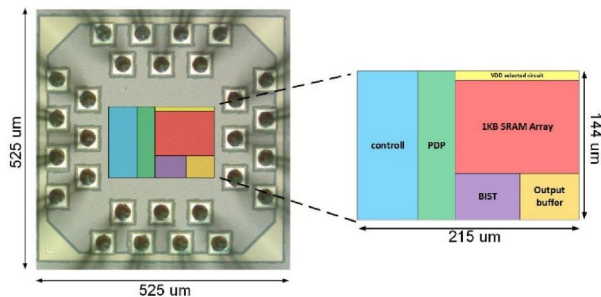


FIGURE 24 Die photo of our SRAM prototype

metal density rules required for the 40-nm CMOS process. A total of 6 chips, measured 50 times each, were used for testing in the measurement site as shown in Figure 25, where this site is in the Tainan branch of TSRI (Taiwan Semiconductor Research Institute). The power supply is Agilent N6761A, the Agilent 81250 pattern generator is used as test vector generator, and the voltage measurements were taken using the Agilent 54855A oscilloscope. The measurements show an improvement in the read delay of the system when boost select is enabled.

For a system clock of 2 MHz, the read delay improvement is from 229 to 64 ns. Figure 26a,b show the read/write timing waveforms when BS = 0 and BS = 1, respectively, at the system clock = 2 MHz.

Upon increasing the system clock to 10 MHz, the read delay is reduced from 148 to 61 ns. Figure 27a,b show the read/write timing waveforms when BS = 0 and BS = 1, respectively, at the system clock = 10 MHz. The proposed AVD produced a huge reduction in the access delay as expected. It was measured that the write-0 delay is 0.250 ns, and the write-1 delay is 0.233 ns. The overall power consumption is found to be 0.8 V \times 30 μ A = 24.0 μ W.

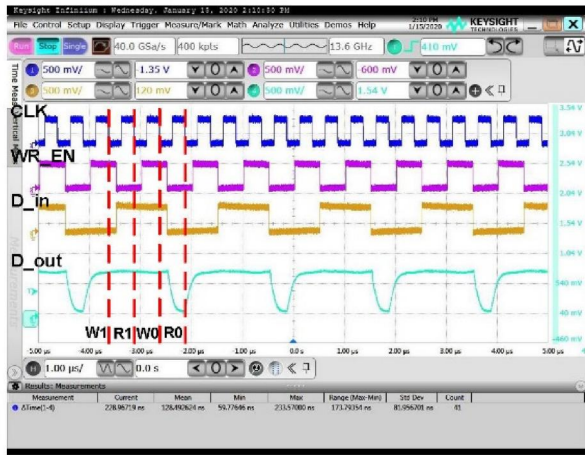
Table 2 tabulates several previous low power/energy SRAM designs using 28, 40, and 65-nm CMOS technology nodes in the past years. The energy/access is defined as the average energy dissipated while executing write-0-read-0 and write-1-read-1 operations divided by the system clock rate. The energy/bit, on the other hand, is defined as the energy/access divided by the number of bits per word. The proposed SRAM attained the second lowest energy per access with a value of 32 fJ, second only to 20.6 fJ [18].

It does, however, demonstrate the lowest energy per bit among all SRAMs designed and implemented using the 40-nm CMOS technology and measured on silicon. It even has a lower energy per bit compared to one work implemented in more advanced CMOS technology, that is, 28-nm CMOS technology [18]. Referring to Figure 28, the SNMs of the SRAMs in the previous decade are compared. The graph is normalised to the respective voltage supply. As observed, the proposed SRAM's SNM is the closest to 50% of the supply voltage, resulting in good noise immunity.

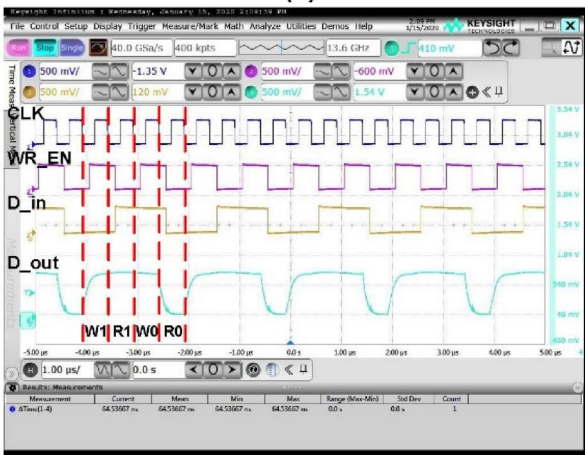
These observations validate the fact that the proposed power-gating PMOS devices indeed account for the reduction of the standby power for those cells that are not activated.



FIGURE 25 Measurement setup for the SRAM prototype



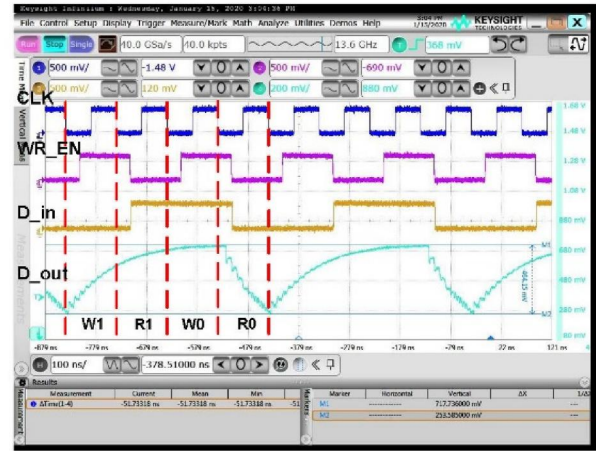
(a)



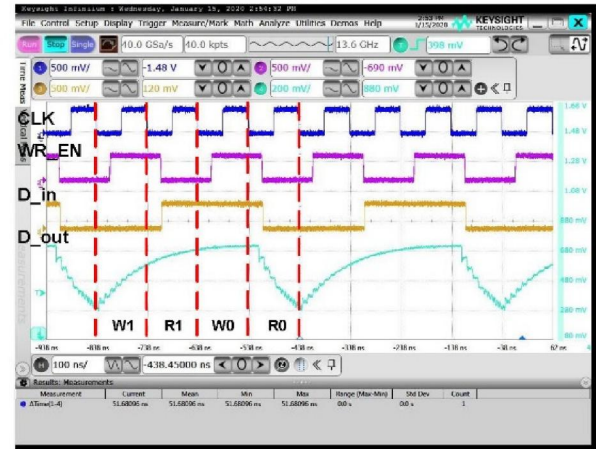
(b)

FIGURE 26 Access operation at (a) 2 MHz (BS = 0) and (b) 2 MHz (BS = 1)

Referring to the technology roadmap shown in Figure 29, the proposed SRAM achieved the historical second lowest energy per bit in the last decade compared with previous works. If the CMOS processes normalise the energy per bit, namely 40-nm (ours) versus 28-nm [18], as well as the PDP reduction, the proposed SRAM is in fact the historical lowest one. The major reason is the addition of the power gating circuit in the single-ended cell wherein the standby power is significantly reduced. Meanwhile, the proposed AVD circuit manages to compensate for the access speed loss by generating a higher supply voltage



(a)



(b)

FIGURE 27 Access operation at (a) 10 MHz (BS = 0) and (b) 10 MHz (BS = 1)

when the access operations are asserted. All of the measurement results substantially prove the energy efficiency of the proposed design.

4 | CONCLUDING REMARKS

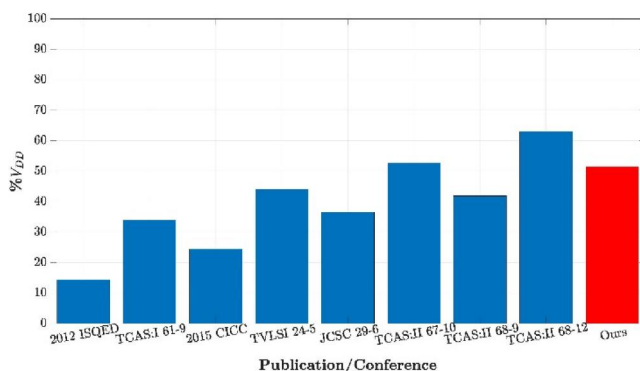
This investigation demonstrates a very low-power SRAM architecture on silicon that has power supply gating that responds to the cell operations. The supply voltage is kept at a lower level for SRAM cells that are not being accessed, which in turn creates a substantial decrease in standby power. Aside from the supply voltage gating, a power delay product reduction circuit is added to the design to further reduce the power dissipation by decreasing the transient time of states. On the other hand, this extra circuit elevates the supply of the read/write circuit to a higher level when the read/write circuit is accessed, in which the delay is significantly reduced. Post-layout simulations verify the ultra-low-power performance, and the physical measurement also showed the expected low power/energy performance. The same design

TABLE 2 Performance comparison of current state-of-the-art low-energy/power SRAM designs

Year	ISQED 2012		TCAS-I 2017				
	VLSIC 2011 [19]	[14]	TCAS-I 2014 [11]	CICC 2015 [20]	TVLSI 2016 [16]	TVLSI 2017 [21]	[22]
CMOS Tech. (nm)	40	40	40	28	40	65	65
Cell	8T	8T	12T	8T	5T	6T	9T
Supply volt. (V)	0.5	0.6	0.35	0.7	0.8	1.2	0.35
Verification ^a	Meas.	Meas.	Meas.	Meas.	Meas.	Simu.	Meas.
SNM (mV)	N/A	86	119	171	353	N/A	N/A
Read PDP (fj)	88	N/A	N/A	650	N/A	17.5	N/A
Capacity (kb)	512	4 + 1	4	64	4 + 1	1	4
Word length	16	16	16	16	5	32	64
Frequency (MHz)	6.25	10	11.5	50	54	100	0.741
Energy/access (pJ)	8.8	2.24	1.91	0.65	0.941	2.2	0.229
Energy/bit (fj)	550	140	119.4	40.625	188.22	68.75	3.58
Core area (mm ²)	0.73	1.278	0.018	0.73	0.024	0.013	0.011

Year	TCAS-II 2018	JCSC 2019 [18]	TCAS-II 2020	TCAS-I 2021	TCAS-II 2021	TCAS-II 2021	This work 2022
	[23]						
CMOS Tech. (nm)	65	28	65	55	40	16	40
Cell	6T	6T	8T	6T	6T	6T	6T
Supply volt. (V)	1.2	0.8	0.36	1.2	0.9	0.8	0.8
Verification ^a	Simu.	Meas.	Meas.	Meas.	Meas.	Meas.	Meas.
SNM (mV)	N/A	292	190	N/A	377	504.76	412.3
Read PDP (fj)	N/A	444.5	4454.4	233.38	47.382	2.69	2.0592
Capacity (kb)	8	1 + 1	32	4	1	1	1
Word length	32	32	128	32	32	32	32
Frequency (MHz)	20	40	0.25	935	200	500	10 (typ.) 15 (max.)
Energy/access (pJ)	0.592	0.026	0.3	1.04	0.2313	0.219	0.032
Energy/bit (fj)	18.5	0.6	2.34	32.5	7.23	6.8	1.0 @10 MHz 4.3 @2 MHz
Core area (mm ²)	0.019	0.025	0.015	0.018	0.01	0.02	0.02

^aSimu. - Simulations or Meas. - Measurements on-chip.

**FIGURE 28** Comparison of SNMs for SRAMs (normalised to V_{DD})

methodology is expected to be applied in more advanced SRAM technology nodes, for example, 22, or even 16-nm FinFET nodes.

AUTHOR CONTRIBUTION

Chua-Chin Wang: Funding acquisition, Visualisation, Formal analysis, Investigation, Methodology, Writing – review & editing. Ralph Gerard B. Sangalang: Formal analysis, Investigation, Methodology, Writing – review & editing. I-Ting Tseng: Conceptualisation, Methodology, Software, Validation. Yi-Jen Chiu: Funding acquisition, Visualisation. Yu-Cheng Lin: Funding acquisition, Visualisation, co-corresponding. Oliver Lexter July A. Jose: Resources, Writing – review & editing.

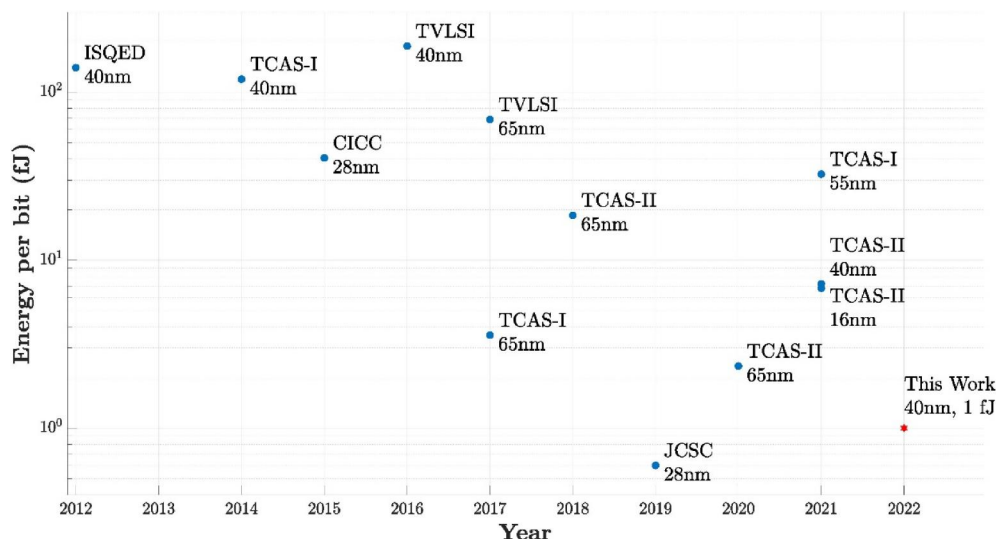


FIGURE 29 Technology roadmap of energy per bit for recent SRAMs

ACKNOWLEDGEMENT

The National Science and Technology Council of Taiwan funded this study in part via grants MOST109-2218-E-110-007, 108-2218-E-110-011, 108-2218-E-110-002, 107-2218-E-110-002, and 110-2221-E-110-063-MY2. The authors would like to express their profound appreciation to TSRI (Taiwan Semiconductor Research Institute) in NARL (National Applied Research Laboratories), Taiwan, for providing EDA tool support, fabrication service, and measurement setup.

CONFLICT OF INTEREST STATEMENT

None.

DATA AVAILABILITY STATEMENT


The data that support the findings of this study are available from the corresponding author upon reasonable request

PERMISSION TO REPRODUCE MATERIALS FROM OTHER SOURCES

None.

ORCID

Chua-Chin Wang  <https://orcid.org/0000-0002-2426-2879>

Ralph Gerard B. Sangalang  <https://orcid.org/0000-0002-4120-382X>

REFERENCES

- Morifuji, E., et al.: Supply and threshold-voltage trends for scaled logic and SRAM MOSFETs. *IEEE Trans. Electron. Dev.* 53(6), 1427–1432 (2006). <https://doi.org/10.1109/TED.2006.874752>
- Xu, H., et al.: A current mode sense amplifier with self-compensation circuit for SRAM application. In: *Proceedings of the 2013 IEEE 10th International Conference on ASIC*; 2013 Oct 28–31; Shenzhen, China, pp. 1–4. IEEE, New York (2013). <https://doi.org/10.1109/ASICON.2013.6812020>
- Do, A.-T., et al.: Design and sensitivity analysis of a new current-mode sense amplifier for low-power SRAM. *IEEE Trans. VLSI Syst.* 19(2), 196–204 (2011). <https://doi.org/10.1109/TVLSI.2009.2033110>
- Kim, D., et al.: A 1.85fW/bit ultra low leakage 10T SRAM with speed compensation scheme. In: *Proceedings of the 2011 IEEE International Symposium of Circuits and Systems (ISCAS)*; 2011 May 15–18; Rio de Janeiro, Brazil, pp. 69–72. IEEE, New York (2011). <https://doi.org/10.1109/ISCAS.2011.5937503>
- Ruixing, L., et al.: Bitline leakage current compensation circuit for high-performance SRAM design. In: *Proceedings of the 2012 IEEE Seventh International Conference on Networking, Architecture, and Storage*; 2012 Jun 28–30; Xiamen, China, pp. 109–113. IEEE, New York (2012). <https://doi.org/10.1109/NAS.2012.19>
- Agawa, K., et al.: A bitline leakage compensation scheme for low-voltage SRAMs. *IEEE J. Solid State Circ.* 36(5), 726–734 (2001). <https://doi.org/10.1109/4.918909>
- Wang, C.-C., et al.: 4-kB 500-MHz 4-T CMOS SRAM using low-V_{THN} bitline drivers and high-V_{THP} latches. *IEEE Trans. VLSI Syst.* 12(9), 901–909 (2004). <https://doi.org/10.1109/TVLSI.2004.833669>
- Wang, C.-C., Lee, C.-L., Lin, W.-J.: A 4-kB low-power SRAM design with negative word-line scheme. *IEEE Trans. Circuits-I* 54(5), 1069–1076 (2007). <https://doi.org/10.1109/TCSI.2006.888767>
- Wang, D.-S., Su, Y.-H., Wang, C.-C.: A readout circuit with cell output slew rate compensation for 5T single-ended 28 nm CMOS SRAM. *Microelectron J.* 70, 107–116 (2017). <https://doi.org/10.1016/j.mejo.2017.11.001>
- Chen, S.-Y., Wang, C.-C.: Single-ended disturb-free 5T loadless SRAM cell using 90 nm CMOS process. In: *Proceedings of the 2012 IEEE International Conference on IC Design and Technology*; 2012 May 30–Jun 1; Austin, TX, USA, pp. 1–4. IEEE, New York (2012). <https://doi.org/10.1109/ICICDT.2012.6232848>
- Chiu, Y.-W., et al.: 40nm bit-interleaving 12T subthreshold SRAM with data-aware write-assist. *IEEE Trans. Circuits-I* 61(9), 2578–2585 (2014). <https://doi.org/10.1109/TCSI.2014.2332267>
- Reddy, S., Sangalang, R.G.B., Wang, C.-C.: Sub-0.2 pJ/access Schmitt trigger-based 1-kb 8T SRAM implemented using 40-nm CMOS process. In: *Proceedings of the 2022 International Conference on IC Design and Technology (ICICDT)*; 2022 Sep 21–23; Hanoi, Vietnam, pp. 24–27. IEEE, New York (2022). <https://doi.org/10.1109/ICICDT56182.2022.9933116>
- Frustaci, F., et al.: Techniques for leakage energy reduction in deep submicrometer cache memories. *IEEE Trans. VLSI Syst.* 14(11), 1238–1249 (2006). <https://doi.org/10.1109/TVLSI.2006.886397>
- Terada, M., et al.: A 40-nm 256-kb 0.6-V operation half-select resilient 8T SRAM with sequential writing technique enabling 367-mV V_{DDmin} reduction. In: *Proceedings of the 13th International Symposium on Quality Electronic Design (ISQED)*; 2012 Mar 19–21; Santa Clara, CA,

- USA, pp. 489–492. IEEE, New York (2012). <https://doi.org/10.1109/ISQED.2012.6187538>
15. Wang, C.-C., Tseng, I.-T.: Ultra low power single-ended 6T SRAM using 40 nm CMOS technology. In: Proceedings of the 2019 International Conference on IC Design and Technology (ICICDT); 2019 Jun 17–19; Suzhou, China, pp. 1–4. IEEE, New York (2019). <https://doi.org/10.1109/ICICDT.2019.8790848>
 16. Wang, C.-C., et al.: A leakage compensation design for low supply voltage SRAM. *IEEE Trans. VLSI Syst.* 24(5), 1761–1769 (2016). <https://doi.org/10.1109/TVLSI.2015.2484386>
 17. Al-Harbi, S.M., Gupta, S.K.: An efficient methodology for generating optimal and uniform march tests. In: Proceedings of the 19th IEEE VLSI Test Symposium (VTS 2001); 2001 Apr 29– May 3; Marina Del Rey, CA, USA, pp. 231–237. IEEE, New York (2001). <https://doi.org/10.1109/VTS.2001.923444>
 18. Wang, C.-C., et al.: A single-ended 28-nm CMOS 6T SRAM design with read-assist path and PDP reduction circuitry. *J. Circ. Syst. Comput.* 29(6), 2050095 (2020). <https://doi.org/10.1142/S0218126620500954>
 19. Yoshimoto, S., et al.: A 40-nm 0.5-V 20.1- μ W/MHz 8T SRAM with low-energy disturb mitigation scheme. In: Proceedings of the 2011 Symposium on VLSI Circuits - Digest of Technical Papers; 2011 Jun 15–17; Kyoto, Japan, pp. 72–73. IEEE, New York (2011). <https://ieeexplore.ieee.org/document/5986220>
 20. Mori, H., et al.: A 298-fJ/writecycle 650-fJ/readcycle 8T three-port SRAM in 28-nm FD-SOI process technology for image processor. In: Proceedings of the 2015 IEEE Custom Integrated Circuits Conference (CICC); 2015 Sep 28–30; San Jose, CA, USA, pp. 1–4. IEEE, New York (2015). <https://doi.org/10.1109/CICC.2015.7338360>
 21. Lee, J., et al.: A 17.5-fJ/bit energy-efficient analog SRAM for mixed-signal processing. *IEEE Trans. VLSI Syst.* 25(10), 2714–2723 (2017). <https://doi.org/10.1109/TVLSI.2017.2664069>
 22. Shin, K., Choi, W., Park, J.: Half-select free and bit-line sharing 9T SRAM for reliable supply voltage scaling. *IEEE Trans Circuits-I.* 64(8), 2036–2048 (2017). <https://doi.org/10.1109/TCSI.2017.2691354>
 23. Surana, N., Mekie, J.: Energy efficient single-ended 6-T SRAM for multimedia applications. *IEEE Trans. Circuits-I* 66(6), 1023–1027 (2019). <https://doi.org/10.1109/TCSII.2018.2869945>
 24. Do, A.-T., Zeinolabedin, S.M.A., Kim, T.T.-H.: Energy-efficient data-aware SRAM design utilizing column-based data encoding. *IEEE Trans. Circuits-II* 67(10), 2154–2158 (2020). <https://doi.org/10.1109/TCSII.2019.2958668>
 25. Chen, J., et al.: Analysis and optimization strategies toward reliable and high-speed 6T compute SRAM. *IEEE Trans. Circuits-I* 68(4), 1520–1531 (2021). <https://doi.org/10.1109/TCSI.2021.3054972>
 26. Wang, C.-C., Kuo, C.-P.: 200-MHz single-ended 6T 1-kb SRAM with 0.2313 pJ energy/access using 40-nm CMOS logic process. *IEEE Trans. Circuits-II* 68(9), 3163–3166 (2021). <https://doi.org/10.1109/TCSII.2021.3091973>
 27. Wang, C.-C., Sangalang, R.G.B., Tseng, I.-T.: A single-ended low power 16-nm FinFET 6T SRAM design with PDP reduction circuit. *IEEE Trans. Circuits-II* 68(12), 3478–3482 (2021). <https://doi.org/10.1109/TCSII.2021.3123676>

How to cite this article: Wang, C.-C., et al.: A 1.0 fJ energy/bit single-ended 1 kb 6T SRAM implemented using 40 nm CMOS process. *IET Circuits Devices Syst.* 1–13 (2023). <https://doi.org/10.1049/cds2.12141>