A readout circuit with cell output slew rate compensation for 5T single-ended 28 nm CMOS SRAM

Deng-Shian Wang, Yu-Hsun Su, Chua-Chin Wang*
National Sun Yat-Sen University, Department of Electrical Engineering, 70 Lian-Hai Rd., Kaohsiung, Taiwan

ARTICLE INFO

Keywords:
Single-ended SRAM cell
Leakage sensor
Slew rate compensation circuit
AVD
Disturb-free

ABSTRACT

A readout circuit with slew rate compensation for nano-scaled SRAMs (Static Random-Access Memory) is proposed in this study. Since the leakage current will dramatically increase to jeopardize memory read/write performance when the nano-scaled SRAM is operated at high system voltages, the proposed AVD (Adaptive Voltage Detector) is utilized to detect the system voltage variation to resolve this issue. When AVD generates an enable signal to WLBC (Word Line Boost Circuit), WLBC will boost wordline voltage to drive the SRAM cell such that the output slew rate is increased. Therefore, the active power of the SRAM will be reduced. A prototypical SRAM is implemented using TSMC 28 nm CMOS logic low power technology. By measurement results at 0.8 V system voltage, the proposed compensation design reduces 17.2% of the power dissipation, and enhances the output slew rate by 46.5% at the expense of 6.6% area overhead. The energy per access is measured to be 0.414 pJ given a 0.8 V system voltage and 100 MHz system clock.

1. Introduction

Statistically, the number of sold mobile phones has overtaken that of sold personal computers worldwide since 2014. According to the Semico Research report, the area of memory in the SoC die of mobile phones is over 50% after 2008 [1]. Undoubtedly, SRAM has occupied a significant percentage thereof. So does SRAM with respect to energy consumption of the entire SoC. Villa et al. reported that up to 70% of the entire SRAM active power is dissipated in charging and discharging the bitlines during read and write operations [2]. To reduce the active power, a single-ended SRAM cell has been reported by Izumikawa et al. [3]. The active power can be significantly decreased because of the number of the transistor has been reduced and the cell only needs to charge or discharge one bitline. However, single-ended SRAM cell encounter a serious problem: poor noise margins [4].

Besides, the straightforward way to reduce the active power is to take advantage of advanced processes. The active power on the bitline is known as follows.

\[ P_{\text{active}} = CV^2 \times f \]  

where \( f \) is the operating frequency, \( V \) is the system voltage, and \( C \) is the equivalent capacitance on the bitline. Advanced processes provide not only smaller transistor layout to reduce the parasitic capacitance on the bitline, but also lower threshold voltage of MOS devices so that the SRAM can be operated at a low system voltage. Although the advanced process can effectively reduce the active power, another serious problem is that the leakage current increases with the technology scaled down. The leakage results in high power consumption and low operation speed. What even worse is that the data stored in the SRAM cells may be compromised by the leakage current. Another approach to reduce the SRAM leakage power by changing memory cell design was reported in Ref. [5]. Besides, to speed up the operating frequency and prevent the data destroyed by the leakage current, many other compensation methods were proposed.

- Sense amplification [6,7]: Since the leakage current increases with the technology scaled down, the read operation speed is decreased. The sense amplification method is used to amplify the voltage (or current) difference between bitlines to pre-determine the output result during the read operation. Thus, following digital circuitry can operate in stable state earlier such that the active power can be reduced. Besides, the output delay of the SRAM cell is secured since it is isolated from the bitline parasitic capacitance. However, since the sense amplification needs to detect the difference between the bitlines, it is not meant for single-ended SRAM cells.
- Bitline boosting [8]: Daeyeon et al. reported a compensation method which uses PMOS as the access transistors for SRAM cells [8]. To increase the access transistor current during the write operation, the...
secondary supply voltage higher than the system voltage is required to boost the bitline voltage. However, since the bitline voltage has been boosted, the leakage from the bitline is also increased. As mentioned before, the increasing leakage will damage the data state stored in the SRAM cell.

- **Body bias boosting [8]:** Daeyeon et al. also reported another compensation method in the same paper. According to the body effect theory, by boosting the body voltage of the PMOSs in SRAM cell to increase the threshold voltage, the leakage can be effectively reduced. Notably, the bitline boosting approach only needs to boost the voltage during the read or write operation, but the body boosting method needs to maintain a high voltage all the time. Moreover, since the PMOS in SRAM cells needs an additional Nwell to boost the body bias, it increases the area quite significantly.

- **Slew rate compensation [9]:** We proposed a compensation method for the 5T single-ended SRAM cell. The compensation method consists of a leakage current sensor and a compensation circuit. When the leakage of the SRAM cell is too high, the compensation circuit will be enabled by the leakage current sensor to elevate the slew rate of the shared inverter on the bitline. Therefore, the following digital circuitry has been speeded up to a stable state such that the active power is decreased. However, this slew rate compensation only works at the read 0 operation.

Although we have already presented a 4T single-ended SRAM cell [10], we added an access transistor between the SRAM cell and the bitline to become a 5T SRAM cell, where the noise margin is increased [11]. In this study, we utilized a gate voltage boosting compensation method demonstrated in Ref. [12] for the 5T single-ended SRAM cell.
By boosting the gate voltage of the access transistor, the slew rate of the SRAM cell can be increased so that both the read and write operations can be accelerated. Although the boosting voltage will increase the leakage current flowing to ground in the access transistor, the additional leakage current will help the 5T SRAM cell to preserve the data state because of the fact that the 5T single-ended SRAM cell utilizes the hidden self-refreshing path to preserve the data stored in the SRAM cell. The proposed SRAM with Readout Slew Rate Compensation Circuit is realized using TSMC 28 nm low power CMOS technology. PMOS transistors with normal threshold voltage, Vth, are used to consist of a latch-like storage. To increase the reading and writing current, low-Vth NMOS transistors are used as access switches. Besides, a low-Vth NMOS M305 is added between the SRAM cell and the BLb to reduce the interference from the BLb. Figs. 3 and 4 show the read and write timing diagrams of the proposed SRAM cell, respectively. To prevent the electric charge on the BLb from destroying the data stored in the accessed SRAM cell, M306 will be turned on at the beginning of the read operating to discharge BLb to ground. Referring to Fig. 4, according to the input data, Data_in, the control circuit will turn on M303 (when Data_in is logic 0) or M304 (when Data_in is logic 1) to record the data in the SRAM cell. Besides, unlike the traditional SRAM cells, the proposed single-ended SRAM cell doesn’t have a path to ground node Q and Qb, a hidden-refreshing path is utilized to maintain the data status stored in the SRAM cell [11].

Fig. 5 demonstrates the R/W timing diagram and the voltage at each node of the proposed SRAM cell by simulations using HSPICE and CADENCE tools.

2.2. Readout Slew Rate Compensation Circuit

Referring to Fig. 1 again, the proposed Readout Slew Rate Compensation Circuit consists of an AVD and a WLBC. AVD is utilized to detect the system voltage. Since the power consumption starts to exponentially rise at 0.7 V, the proposed AVD will generate a flag signal, BST_en, to turn on the following WLBC when the system voltage reaches 0.7 V. Then, WLBC will generate a driving signal higher than system voltage to drive the access transistor, M305, in the SRAM cell, such that the read and write operations can be speeded up. Details of AVD and WLBC will be introduced in the following subsections.

2.2.1. Adaptive Voltage Detector (AVD)

Fig. 6 shows the schematic of the proposed AVD, which consists of three stages: detect stage, compare stage, and latch stage. CS (chip select) is an input enable signal. When CS is logic 1, the first stage of AVD, detect stage, will be activated, where M407 acts as a diode. If the system voltage is lower than the turn-on voltage (threshold voltage) of M407, the output of detect stage, Vo6, will almost equal to the system voltage, as shown in Fig. 7. Once the system voltage is high enough to turn on M407, the system voltage will be divided by R401 and the equivalent resistance of M407 generate the output Vo6. Referring to the compare stage, the inverter composed of M407 and M411, will be enabled by the CLK = logic 1. The equation of the inverter transfer

\[
\text{By boosting the gate voltage of the access transistor, the slew rate of the SRAM cell can be increased so that both the read and write operations can be accelerated. Although the boosting voltage will increase the leakage current flowing to ground in the access transistor, the additional leakage current will help the 5T SRAM cell to preserve the data state because of the fact that the 5T single-ended SRAM cell utilizes the hidden self-refreshing path to preserve the data state stored in the SRAM cell. The proposed SRAM with Readout Slew Rate Compensation Circuit is realized using TSMC 28 nm low power CMOS technology. PMOS transistors with normal threshold voltage, Vth, are used to consist of a latch-like storage. To increase the reading and writing current, low-Vth NMOS transistors are used as access switches. Besides, a low-Vth NMOS M305 is added between the SRAM cell and the BLb to reduce the interference from the BLb. Figs. 3 and 4 show the read and write timing diagrams of the proposed SRAM cell, respectively. To prevent the electric charge on the BLb from destroying the data stored in the accessed SRAM cell, M306 will be turned on at the beginning of the read operating to discharge BLb to ground. Referring to Fig. 4, according to the input data, Data_in, the control circuit will turn on M303 (when Data_in is logic 0) or M304 (when Data_in is logic 1) to record the data in the SRAM cell. Besides, unlike the traditional SRAM cells, the proposed single-ended SRAM cell doesn’t have a path to ground node Q and Qb, a hidden-refreshing path is utilized to maintain the data status stored in the SRAM cell [11].

Fig. 5 demonstrates the R/W timing diagram and the voltage at each node of the proposed SRAM cell by simulations using HSPICE and CADENCE tools.

2.2. Readout Slew Rate Compensation Circuit

Referring to Fig. 1 again, the proposed Readout Slew Rate Compensation Circuit consists of an AVD and a WLBC. AVD is utilized to detect the system voltage. Since the power consumption starts to exponentially rise at 0.7 V, the proposed AVD will generate a flag signal, BST_en, to turn on the following WLBC when the system voltage reaches 0.7 V. Then, WLBC will generate a driving signal higher than system voltage to drive the access transistor, M305, in the SRAM cell, such that the read and write operations can be speeded up. Details of AVD and WLBC will be introduced in the following subsections.

2.2.1. Adaptive Voltage Detector (AVD)

Fig. 6 shows the schematic of the proposed AVD, which consists of three stages: detect stage, compare stage, and latch stage. CS (chip select) is an input enable signal. When CS is logic 1, the first stage of AVD, detect stage, will be activated, where M407 acts as a diode. If the system voltage is lower than the turn-on voltage (threshold voltage) of M407, the output of detect stage, Vo6, will almost equal to the system voltage, as shown in Fig. 7. Once the system voltage is high enough to turn on M407, the system voltage will be divided by R401 and the equivalent resistance of M407 generate the output Vo6. Referring to the compare stage, the inverter composed of M407 and M411, will be enabled by the CLK = logic 1. The equation of the inverter transfer
The transfer voltage is proportional to the system voltage when the transfer voltage is higher than 0 V, as shown in Fig. 7. If the proposed SRAM is operated at high system voltage is known as follows [13].

\[
V_{trf} = \frac{V_m + \sqrt{\frac{\beta_p}{\beta_n}} (V_{DD} - |V_{p0}|)}{1 + \sqrt{\frac{\beta_p}{\beta_n}}}
\]  

\[ (2) \]

Fig. 5. The simulation results of write and read timing diagram.

Fig. 6. Schematic of AVD.

Fig. 7. Vp0 vs. Vtrf.
voltage, the transfer voltage of the inverter will be higher than the output voltage of the detect stage. Then, logic 1 will be delivered to the next latch stage. As the level of CLK is dropped from 1 to 0, the compare stage is disabled and the latch stage locks the output of the compare stage. Overall, if the system voltage is too high to cause the leakage seriously affecting the entire SRAM operation, the AVD will generate an enable signal to activate WLBC.

Notably, the inverter in the compare stage can be used as a corner detector. Since the proposed 5T SRAM cell utilizes the hidden-refreshing path to keep the data stored in the SRAM cell, it may be easily disturbed by the leakage during the read or write operation. It is particularly serious when the SRAM cell is operated at fast PMOS and typical (or slow) NMOS corner. The boosting voltage which drives the access transistor, M305, 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 in Fig. 2, namely M301 and M302, leakage. Referring to Eqn. (2), the slope of the inverter transfer voltage is increased by high hole mobility, $\mu_p$, at fast PMOS corner, shown as the red line in Fig. 7. Therefore, the boundary voltage of the compensation requirement will be mitigated.

2.2.2. Word Line Boost Circuit (WLBC)

Fig. 8 shows the schematic of the proposed WLBC, where the timing diagrams is shown in Fig. 9. As CS is logic 0, WLBC is operated at shut down mode. After CS is pulled up to logic 1, AVD begins to detect the system voltage and WLBC waits for the AVD’s output, namely waiting mode. As long as AVD has not completed the system voltage detection, the output of AVD, BST_en, is kept at logic 0. Therefore, when WLBC is in the waiting mode, node b of the Cbst (top plate) will be pulled
down to ground by inv3 and the other side of the Cbst (bottom plate), Vbst, will be pulled up to VDD via M413. If the system voltage is higher than the boundary voltage, AVD will pull BST_en high. At this moment, WLBC enters standby mode. Once the control signal WL pulls high, which means one of the SRAM cells start to write or read, WLBC will enter boosting mode such that M413 is turned off and then node b is pulled high. To estimate the boosting voltage on Vbst, the equivalent circuit with respect to each capacitance on Vbst is shown in Fig. 10. According to Kirchhoff circuit laws, the currents at from the node Vbst are summarized as follows.

\[
\begin{align*}
V_{DD} - (V_R - V_{DD})C_{bst} + [V_{DD} - (V_R - V_{DD})]C_{413} &= [0 - (0 - V_R)]C_{414} + 31 \times [0 - (V_{DD} - V_R)]C_{414} \\
&= (3)
\end{align*}
\]

where the \(C_{413}\) is the parasitic capacitance between the drain and the gate of M413, and \(C_{414}\) is the parasitic capacitance between the source and the gate of M414. Then, the boosting voltage \(V_B\) can be derived from Eqn. (3) to be the following equation.
Fig. 13. Simulation results of SNM by (a) all-PT-corner (b) Monte Carlo.

Table 1
The comparisons of slew rate and the power consumption.

<table>
<thead>
<tr>
<th>Corner</th>
<th>Temp. (°C)</th>
<th>Without Comp.</th>
<th>With Comp.</th>
<th>Improvement</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>slew rate (V/ns)</td>
<td>power (µW)</td>
<td>slew rate (V/ns)</td>
</tr>
<tr>
<td>TT</td>
<td>0</td>
<td>0.0988</td>
<td>3.126</td>
<td>0.1228</td>
</tr>
<tr>
<td></td>
<td>25</td>
<td>0.1001</td>
<td>5.791</td>
<td>0.1258</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>0.1014</td>
<td>5.119</td>
<td>0.1289</td>
</tr>
<tr>
<td></td>
<td>75</td>
<td>0.1027</td>
<td>9.575</td>
<td>0.1319</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>0.1039</td>
<td>18.29</td>
<td>0.1331</td>
</tr>
<tr>
<td>FF</td>
<td>0</td>
<td>0.1124</td>
<td>16.12</td>
<td>0.1488</td>
</tr>
<tr>
<td></td>
<td>25</td>
<td>0.1125</td>
<td>33.23</td>
<td>0.1494</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>0.1118</td>
<td>67.41</td>
<td>0.1496</td>
</tr>
<tr>
<td></td>
<td>75</td>
<td>0.1125</td>
<td>118.6</td>
<td>0.1507</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>0.1121</td>
<td>186.2</td>
<td>0.1510</td>
</tr>
<tr>
<td>SS</td>
<td>0</td>
<td>0.0828</td>
<td>0.988</td>
<td>0.0971</td>
</tr>
<tr>
<td></td>
<td>25</td>
<td>0.0853</td>
<td>1.400</td>
<td>0.1012</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>0.0877</td>
<td>1.664</td>
<td>0.1050</td>
</tr>
<tr>
<td></td>
<td>75</td>
<td>0.0899</td>
<td>2.189</td>
<td>0.1093</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>0.0919</td>
<td>3.443</td>
<td>0.1128</td>
</tr>
</tbody>
</table>

Avg. 26.74 18.39
To ensure the boosting voltage can maintain a read or write cycle, \( C_{\text{bst}} \) is selected to be ten times of \( C_{413} \). The size of \( M_{413} \) and \( M_{414} \) are equal such that \( C_{413} \) is equal to \( C_{414} \). Notably, the highest system voltage of proposed SRAM is 1 V and the voltage drop limitation between the drain and source of the 28 nm MOS is 1.25 V, which means that \( V_B \) needs to lower than 1.25 V, while 32 SRAM cell modules share a WL control signal in the proposed SRAM. Finally, the WLBC replicate the input WL to the output WL_c with voltage of logic “1” higher than the system voltage.

3. Implementation and measurement

3.1. Simulation and analysis

Fig. 11 (a) is the layout of the proposed SRAM where the chip area is 0.72 × 0.75 mm² (with pad) and the core area is 0.09 × 0.13 mm². Besides, the area of the AVD and WLBC is 8.5 × 7.4 and 39.5 × 10 μm², respectively. The die photo of the proposed SRAM design is shown in Fig. 11 (b). Since it is covered by dummy metals, the core is not visible in this photo.

Referring Fig. 12 (a), the DNM (Dynamic noise margin) of the proposed SRAM cell by all-PT-corner simulations is justified (system voltage is 0.8 V). Each node includes 25 simulation results ([TT, FF, SS, SF, FS] × [0 °C, 25 °C, 50 °C, 75 °C, 100 °C]). The Fail node means that the data is flipped in one of the simulation results. When the system voltage of the SRAM is 0.8 V, the DNM is around 50 ps @ 0.6 V, which means that the data state stored in the proposed SRAM cell will not be flipped when the amplitude of the noise is lower than 0.6 V and the length of time is shorter than 50 ps. Fig. 12 (b) shows the DNM Monte Carlo simulation results. Each node includes 1000 times simulation results. The Fail node means that the data is flipped more than 10 times out of 1000 times. Apparently, Fig. 12 (a) and (b) are almost the same.

To find out the noise rejection limitation during the operation, the all-PT-corner read SNM (Static noise margin) simulations of the proposed SRAM cell is demonstrated in Fig. 13 (a) (when system voltage is 0.8 V), where the worst case of SNM is 249 mV at FF and 100 °C corner. Fig. 13 (b) shows 10,000 times SNM Monte Carlo simulation.
results. The standard deviation is 32 mV and the range of the 3-σ deviations is 0.367 ~ 0.557 V.

Table 1 shows the comparisons of slew rate and the power consumption between the SRAM with and without the Readout Slew Rate Compensation Circuit. In summary, the mean average power dissipation is reduced by 18.39% and the mean average slew rate is enhanced by 26.74%. Notably, the overhead of the core area is only 6.6%.

3.2. Chip measurement

Fig. 14 shows the function test of the proposed SRAM. The testing process is (write 1 → read 1 → write 0 → read 0) in one cycle for one SRAM cell. The BLb in the SRAM array is discharged to logic 0 during the write operation to pull high the output.

Figs. 15 and 16 show the rise edge and fall edge of the outputs at 0.8 V system voltage, respectively, where Data_out_t is the output of the SRAM array with the Compensation Circuit, and Data_out is the output of the uncompensated counterpart. The rise time is reduced from 37.3 ns to 29.1 ns and the fall time is reduced from 24.33 ns to 5.94 ns. Fig. 17 shows the maximum operating clock frequency at different supply voltages. More importantly, the energy per access (average of read and write access) is reduced from 0.5 pJ to 0.414 pJ given the 0.8 V system voltage and 100 MHz system clock.

Table 2 tabulates the performance comparison of the proposed design with several recent works. A figure-of-merit (FOM) is used to tell the overall performance,

\[
\text{FOM} = \frac{\text{Energy/ access}}{\text{Normalized cell area}} \times 10^3
\]

The FOM in Eqn. (5) the smaller, the better. The proposed design has the best energy per access and the best operating frequency besides the second best FOM.

Table 2

<table>
<thead>
<tr>
<th>Technology</th>
<th>Cell Architecture</th>
<th>Capacity (kb)</th>
<th>Frequency (MHz)</th>
<th>Energy/access (pJ)</th>
<th>Core Area (µm²)</th>
<th>FOM (x10³)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TCAS-I′14 [15]</td>
<td>12T</td>
<td>4</td>
<td>11.5</td>
<td>16</td>
<td>134 × 132</td>
<td>2.7</td>
</tr>
<tr>
<td>TCAS-I′14 [16]</td>
<td>8T</td>
<td>512</td>
<td>200</td>
<td>N/A</td>
<td>947 × 2810</td>
<td>3.17</td>
</tr>
<tr>
<td>TVLSI′15 [9]</td>
<td>5T</td>
<td>4 + 1</td>
<td>54</td>
<td>0.9411</td>
<td>136.51 × 181.16</td>
<td>3.01</td>
</tr>
<tr>
<td>TVLSI′15 [17]</td>
<td>5T</td>
<td>32</td>
<td>N/A</td>
<td>2.7</td>
<td>N/A</td>
<td>2.19</td>
</tr>
<tr>
<td>This work</td>
<td></td>
<td>1 + 1</td>
<td>210</td>
<td>0.414</td>
<td>N/A</td>
<td>7.28</td>
</tr>
<tr>
<td>Normalized cell area a</td>
<td>2.7</td>
<td>3.17</td>
<td>3.01</td>
<td>2.19</td>
<td>5.91</td>
<td>3.01</td>
</tr>
</tbody>
</table>

\[ a \text{ Normalized cell area = 1000 × Core Area/(Capacity×Tech.²).}\]
4. Conclusion

An SRAM with Readout Slew Rate Compensation Circuit for 5T SRAM is presented in this paper. Two SRAM arrays are realized in one single chip to explicitly demonstrate the performance of the proposed compensation design, where one array is with the proposed compensation design and the other without. The proposed compensation design is composed of an AVD and a WLBC. AVD is utilized to detect the system voltage to trigger the following WLBC which generates a boosting voltage to speed up the read and write operations. By the measurement results, given 0.8 V system voltage, the power dissipation is reduced by 17.2%, and the output slew rate is improved by 46.5% at the expense of 6.6% area overhead. The energy per access is measured to be 0.414 pJ with 800 mV power supply and 100 MHz clock frequency.

Acknowledgement

This investigation is partially supported by Ministry of Science and Technology under grant MOST 104-2622-E-006-040-CC2, MOST 105-2218-E-110-006, and MOST 105-2221-E-110-058. The authors would like to express their deepest gratefulness to Chip Implementation Center of National Applied Research Laboratories, Taiwan, for their thoughtful chip fabrication service and EDA tool support.

References