Design of A Highly Reliable and Low-Power SRAM With Double-Node Upset Recovery for Safety-critical Applications
Aibin Yan, Jing Xiang, Zhengfeng Huang, Tianming Ni, Jie Cui, Patrick Girard, Xiaoqing Wen

To cite this version:
Aibin Yan, Jing Xiang, Zhengfeng Huang, Tianming Ni, Jie Cui, et al.. Design of A Highly Reliable and Low-Power SRAM With Double-Node Upset Recovery for Safety-critical Applications. ITC-Asia 2023 - IEEE International Test Conference in Asian, Sep 2023, Matsue, Japan. lirmm-04241173

HAL Id: lirmm-04241173
https://hal-lirmm.ccsd.cnrs.fr/lirmm-04241173
Submitted on 13 Oct 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Design of A Highly Reliable and Low-Power SRAM With Double-Node Upset Recovery for Safety-critical Applications

Aibin Yan\textsuperscript{1,2}, Jing Xiang\textsuperscript{1}, Zhengfeng Huang\textsuperscript{2}, Tianming Ni\textsuperscript{3}, Jie Cui\textsuperscript{1}, Patrick Girard\textsuperscript{4}, and Xiaqing Wen\textsuperscript{5}

\textsuperscript{1}School of Computer Science and Technology, Anhui University, Hefei, China
\textsuperscript{2}School of Microelectronics, Hefei University of Technology, Hefei, China
\textsuperscript{3}School of Integrated Circuits, Anhui Polytechnic University, Wuhu, China
\textsuperscript{4}Laboratory of Informatics, Robotics and Microelectronics of Montpellier, University of Montpellier / CNRS, Montpellier, France
\textsuperscript{5}Graduate School of Computer Science and Systems Engineering, Kyushu Institute of Technology, Fukuoka, Japan

Abstract—For high-speed operations, low power consumption and small silicon area, transistors are being scaled aggressively. Meanwhile, circuit reliability is facing greater challenges in advanced technologies. In this paper, a highly reliable and low-power SRAM with double-node-upset (DNU) recovery, namely HRLP16T, is proposed for safety-critical fields. HRLP16T can recover from single-node-upset (SNU) at all the sensitive nodes, and it has eight node pairs recover from DNUs. Simulation-based evaluation results demonstrate advantages in terms of delay and power consumption over epical existing SRAM cell designs.

Index Terms—Radiation, circuit reliability, sensitive node, soft error, double-node-upset.

I. INTRODUCTION

In aerospace, radiative particles striking the sensitive nodes of an SRAM can cause soft errors. Radiative particles include protons, heavy ions, muons, electrons, etc. \cite{1} The scenario that the stored value of a sensitive node changes because of the striking is called an SNU while the scenario that the stored values of two sensitive nodes change because of the striking is called a DNU. SRAMs are widely used in aerospace, military, communications and other safety-critical fields, so that the reliability of SRAMs is very important. In addition, the tradeoff of reliability, overhead and power consumption has also become a vital research content. It is known that the error correcting code (ECC) method can correct errors in large scale integration circuits \cite{2}. However, since it needs coding and decoding circuits, large power, delay and area penalties are in ECCs. To address the above issues, we focus on novel structure designs of SRAMs based on the radiation-hardening-by-design (RHB) approach to make SRAMs radiation-hardened with balanced overhead.

In order to mitigate soft errors, researchers have proposed a series of hardened SRAM cells. A conventional 6T cell consists of cross-coupled inverters. If the value of a node changes, the values of other nodes change due to the positive feedback formed by the cross-coupled inverters. Since a 6T cell cannot tolerate SNU, researchers have proposed many improved schemes, including RHBD10T \cite{3}, RHM12T \cite{4}, S4P8N \cite{5}, S8P4N \cite{5}, RHMD10T \cite{6}, SAR14T \cite{7}, SEA14T \cite{8}, Lin12T \cite{9}, and SRRD12T \cite{10}. However, these SRAM cells have their own shortcomings, (For example, RHBD10T and RHM12T cannot recover from SNU, Lin12T, S4P8N and S8P4N can recover from SNU, but they have fewer node pairs can recover from DNUs.) SAR14T, SRRD14T and SEA14T have eight node pairs that can recover from DNUs; however, SAR14T has higher power consumption, SRRD14T has a large write delay, and SEA14T has a large read delay and higher power consumption.

The SRAM cell proposed in this paper, namely HRLP16T cell, is composed of 16 transistors. Among these transistors, six PMOS transistors and six NMOS transistors constitute the storage module, and the remaining four NMOS transistors constitute the access operation part. The proposed HRLP16T cell shows excellent SNU self-recovery. For DNU self-recovery, a HRLP16T cell has eight pairs of self-recovery nodes. Compared to low-reliability SRAMs, i.e., RHBD10T, RHM12T, Lin12T, S4P8N, and S8P4N, a HRLP16T cell has higher reliability. Moreover, the write/read access time (WAT/RAT) of a HRLP16T cell are equal to 4.02ps/22.41ps, the power of a HRLP16T cell is equal to 40.68nW, the area of a HRLP16T cell is 11.08×10^{-3} nm^2. Compared to high-reliability SRAMs, i.e., SAR14T, SRRD12T, and SEA14T, a HRLP16T cell has balanced overhead.

The rest of the paper is organized as follows: Section II introduces the schematic diagram and working principle of the proposed HRLP16T cell. Section III describes the comprehensive comparison and evaluation results. Section IV concludes this paper.

II. PROPOSED HRLP16T SRAM CELL

A. Schematic and Normal Operations

Fig. 1 shows the structure of the proposed HRLP16T cell. It can be seen from Fig. 1 that the HRLP16T cell is composed of 16 transistors. Among these transistors, six PMOS transistors and six NMOS transistors constitute the storage module, and the remaining four NMOS transistors constitute the access operation module. The four NMOS transistors of the access operation module are connected to the four storage nodes, i.e., S0, Q, QB, and S1. In addition, the gates of four NMOS transistors are connected to the word line WL. These four NMOS transistors are used to connect the bit line BL or BLN to the storage nodes of the cell, respectively. We use word line WL and bit lines (i.e., BL and BLN) to implement read/write and hold functions.
Fig. 2 shows the error-free operations of the proposed HRLP16T cell. It can be seen that reading 0, reading 1, writing 0 and writing 1 can be operated correctly. Note that, the cell was implemented in the 22nm commercial CMOS technology at the room temperature and with the 0.8V supply voltage.

(1) Hold Mode: When WL = 0, N7~N10 are all OFF, the cell is in hold mode. A feedback loop (S0 → S1 → Q → QB → Q → S0) is formed in the cell to maintain the stored value.

(2) Write Mode: When WL = 1, BL = 0, and BLN = 1, the proposed HRLP16T cell performs write 0 operation. At this time, N7~N10 are all ON, nodes Q and S1 are discharged to 0 by BL. BLN charges nodes QB and S0, and then the operation of writing '0' is completed. When WL = 1, BL = 1, and BLN = 0, the HRLP16T cell performs the write '1' operation. Similarly, N7~N10 are all ON, nodes QB and S0 are discharged to 0 by BLN. BL charges nodes Q and S1, and then the operation of writing '1' is completed.

(3) Read Mode: Before the read operation, both BL and BLN are preset to 1. When the cell stores 0, i.e., Q = S1 = 0, the BL discharges through the connected storage nodes. When the cell stores 1, i.e., QB = S0 = 1, the BLN discharges through the connected storage nodes. Then, sense amplifier detects the differential voltage between BL and BLN to determine the reading value.

**B. Self-Recovery Principle**

It is known that a particle strike on a semiconductor device generates a transient voltage pulse, which may flip the stored value at the sensitive node. The sensitive area in a semiconductor device is strongly reverse-biased [4]. Therefore, the induced transient current flows from the N-type diffusion to the P-type diffusion. Fig. 3 shows the model of soft errors. As shown in Fig. 3(a), when a radiation particle strikes a PMOS transistor, only a positive transient pulse is generated, i.e., either 1 → 1 or 0 → 1 based on the initial node value. As shown in Fig. 3(b), while in the case of an NMOS, it generates only a negative transient pulse, i.e., 1 → 0 or 0 → 0.

It can be seen from Fig. 1 that nodes QB and Q are connected to the drain of PMOS transistors, so that only positive transient pulses can be generated at Q and QB, i.e., 1 → 1 or 0 → 1. When HRLP16T cell stores value 0, i.e., Q = S1 = 0 and QB = S1 = 1, the Q, S0 and S1 are sensitive nodes. Similarly, When HRLP16T cell stores value 1, i.e., Q = S1 = 1 and QB = S1 = 0, the QB, S0 and S1 are sensitive nodes because the sensitive nodes of an integrated circuit are those in the surroundings of the reverse-biased drain junctions of an OFF transistor [11]. Due to the symmetry of the HRLP16T cell, we only analyze the case that the cell stores 0. When HRLP16T cell stores '0', i.e., Q = S1 = 0 and QB = S1 = 1, QB is an insensitive node so that we consider nodes Q, S0 and S1.

Here we discuss SNU self-recovery principle in detail. (1) When Q is affected by an SNU, i.e., the value of node Q changes from 0 to 1, it can cause P1 and P3 to be OFF and N1 and N6 to be ON temporarily. The value of S1 does not change immediately, so that N2 remains OFF and P5 remains ON. Since N2 and P1 are both OFF, node S0 goes to a high impedance state. Normally, a high impedance state does not affect the logic value of the node, so that P6 remains OFF and N5 remains ON. Similarly, both P3 and P6 are OFF, node QB remains to be 1, so that P2 is OFF and N3 is ON. Since P5 also keeps ON, node Q returns to its original value of 0. Therefore, Q can self-recover from SNU.

(2) When S0 is affected due to the SNU, i.e., the value of node S0 changes from 1 to 0, it can cause N5 to be OFF and P6 to be ON temporarily. The value of Q does not change.
immediately, so that N6 and N1 remains OFF and P1 and P3 remains ON. Since P1 remains ON and N1 remains OFF, node S0 returns to its original value of 1. Hence, N5 is ON. Since P3 keeps ON and N6 keeps OFF, node QB remains to be 1. Hence, P4 keeps OFF and N4 keeps ON. Node S1 keeps its original value. Therefore, S0 can self-recover from SNU.

(3) When S1 is affected due to the SNU, i.e., the value of node S1 changes from 0 to 1, it can cause N2 to be ON and P5 to be OFF temporarily. The value of QB does not change immediately, so that P2 and P4 remains OFF and N3 and N4 remains ON. Since both P2 and P5 are OFF, node Q goes to a high impedance state and Q remains to be 0. Hence, P1 keeps ON and N1 keeps OFF, so that S0 keeps its original value of 1. Then, N5 keeps ON. Since N4 and N5 are ON, and P4 is OFF, node S1 returns to its original value of 0. Therefore, S1 can self-recover from SNU. In summary, every node of the HRLP16T cell can recover from SNU. Fig. 4 shows the SNU self-recovery simulation results of the HRLP16T cell. It can be seen that it has excellent SNU self-recoverability.

![Fig. 4. SNU self-recovery simulation results of the HRLP16T cell.](image)

Next, let us analyze the principle of self-recovery from DNUs. When the cell stores 0, QB is an insensitive node so that node pair <S0, QB>, <Q, QB>, and <S1, QB> can self-recover from DNUs because every node of the HRLP16T cell can recover from SNU. In addition, node pair <S0, S1> can also self-recover from DNUs when the cell stores 0. Similarly, Q is an insensitive node when the cell stores 1, so that node pairs <S0, Q>, <S1, Q>, and <QB, Q> can self-recover from DNUs. Moreover, node pair <S0, S1> can also self-recover from DNUs. Due to symmetry of the HRLP16T cell, we only analyze the case that the cell stores 0.

When the errors are injected into nodes S0 and S1 simultaneously to simulate a DNU, the value of S0 changes from 1 to 0, and the value of S1 changes from 0 to 1. Thus, N5 and P5 are OFF, and N2 and P6 are ON. Since Q and QB are the outputs of a cross-coupled inverters, their values cannot be changed without any external interference (e.g., accessing the nodes by the bit lines during a write operation). Therefore, Q and QB remain to be original values. Hence, P1 and N4 keep ON, N1 and P4 keep OFF. Node S0 returns to its original value 1 because P1 is ON and N1 is OFF, and then N5 is ON. Since N5 and N4 are ON, and P4 is OFF, node S1 returns to its original value of 0. Therefore, node pair <S0, S1> can self-recover from DNU.

When the DNUs are injected into node pairs <QB, S0> or <QB, S1>, the stored values of HRLP16T can be changed due to the deposited charge. The value of QB changes from 1 to 0, and the value of S0 (S1) changes from 1 (0) to 0 (1). Thus, P2, P4 and P6 (N2) are ON, and N3, N4 and N5 (P5) are OFF. Since P4 is ON and N4 is OFF, the value of S1 flips. Similarly, since P2 is ON and N3 (P5) is OFF, the value of Q is flipped. Thus, the stored value of HRLP16T is flipped. In order to avoid the occurrence of DNUs at node pairs <QB, S0> and <QB, S1>, we can design the layout of HRLP16T cell with sufficient separation between node QB and node S0 (S1).

![Fig. 5. DNU self-recovery simulation results of the HRLP16T cell.](image)

Fig. 5 shows the DNU self-recovery simulation results of the HRLP16T cell. It can be seen from Fig. 5 that node pair <S0, S1> can self-recover from DNUs whether HRLP16T stores 0 or 1. Besides, when the cell stores 0, QB is an insensitive node so that node pairs <S0, QB>, <Q, QB>, and <S1, QB> can avoid to occur DNUs. When the cell stores 1, Q is an insensitive node so that node pairs <S0, Q>, <S1, Q>, and <QB, Q> can prevent DNUs from occurring. In summary, there are eight node pairs of HRLP16T that can self-recover or cannot suffer from DNUs.

**C. Simulation**

We adopt the double exponential current model to simulate that a particle strikes a sensitive node as an SNU, so that a transient of the correct polarity can be created at a sensitive node (i.e., a positive (negative) transient is created at the drain of PMOS (NMOS) transistors) (see Fig. 6) [12]. The expression of the double exponential current is shown as follows:

\[
I(t) = \frac{Q}{\tau_a - \tau_p} \left( e^{\frac{-t}{\tau_a}} - e^{\frac{-t}{\tau_p}} \right)
\]

In Eq. (1), \(Q\) represents the total collected charge, the \(\tau_a (= 13 - \text{e ps})\) represents the collection time constant and the \(\tau_p (= 12 - 3\text{e ps})\) symbolizes the time constant of initial ion track establishment.

![Fig. 6. Simulation of particle-strike. (a) positive transient pulse. (b) negative transient pulse.](image)
positive edge and end with the time that intersect the voltages. Begin with the WL reaches 50% of its full swing at the power consumption. Compared to other SRAMs, the power consumption (static and dynamic) for SAR14T has the highest power consumption (static and dynamic) for comparison. It can be seen from Table I that the average power consumption of other nodes immediately. SRRD12T has the largest silicon area. Because of the structure of cross-coupled inverters, the value change of one node can influence the values of other nodes immediately. SRRD12T has the largest silicon area. Therefore, compared to other SRAMs, the HRLP16T is moderate in terms of WAT.

Finally, we analyze the silicon area. Since HRLP16T uses sixteen transistors, it can be seen from Table I that the silicon area of HRLP16T is slightly larger than that of the other SRAMs except S4P8N. HMD10T has the smallest silicon area because of its fewest number of transistors.

C. PVT Estimations
Process-voltage-temperature (PVT) variation impacts on SRAMs are discussed. Fig. 7 shows the estimation results of supply voltage/temperature variation impacts on RAT, WAT, and power for SRAM designs. We can observe from Fig. 7-(a)(b)(c) that RAT and WAT decrease and power consumption increases with supply voltage increases. By observing the slope of the broken lines in Fig. 7-(a), we can know that RAT of SEA14T, SAR14T, and HMD10T is lower than the WAT of HRLP16T. Because of the structure of cross-coupled inverters, the value change of one node can influence the values of other nodes immediately. SRRD12T has the largest silicon area. Therefore, compared to other SRAMs, the HRLP16T is moderate in terms of WAT.

Now we discuss overhead comparisons. First, we analyze the power. Power consumption during signal switching accounts for a large proportion of total power consumption. In order to ensure fairness, we use average power consumption (static and dynamic) for comparison. It can be seen from Table I that HMD10T and Lin12T have lower power consumption. SAR14T has the highest power consumption because of the leakage current of inverters. Compared to other SRAMs, the HRLP16T has moderate power consumption.

Second, we analyze the RAT. We use the average time that begin with the WL reaches 50% of its full swing at the positive edge and end with the BL/BLN reaches 100mV below the supply voltage as RAT. It can be seen from Table I that the S4P8N, S8P4N, and SAR14T have lower RAT than HRLP16T. SEA14T shows the largest RAT because it has only two read/write transistors. Compared to other SRAMs, HRLP16T has lower RAT except SAR14T, S8P4N and S8P4N.

Now we discuss overhead comparisons. First, we analyze the power. Power consumption during signal switching accounts for a large proportion of total power consumption. In order to ensure fairness, we use average power consumption (static and dynamic) for comparison. It can be seen from Table I that the S4P8N, S8P4N, and SAR14T have lower power consumption. Compared to other SRAMs, the HRLP16T has the highest power consumption. Because of the structure of cross-coupled inverters, the value change of one node can influence the values of other nodes immediately. SRRD12T has the largest silicon area. Therefore, compared to other SRAMs, the HRLP16T is moderate in terms of WAT.

Finally, we analyze the silicon area. Since HRLP16T uses sixteen transistors, it can be seen from Table I that the silicon area of HRLP16T is slightly larger than that of the other SRAMs except S4P8N. HMD10T has the smallest silicon area because of its fewest number of transistors.

C. PVT Estimations
Process-voltage-temperature (PVT) variation impacts on SRAMs are discussed. Fig. 7 shows the estimation results of supply voltage/temperature variation impacts on RAT, WAT, and power for SRAM designs. We can observe from Fig. 7-(a)(b)(c) that RAT and WAT decrease and power consumption increases with supply voltage increases. By observing the slope of the broken lines in Fig. 7-(a), we can know that the RAT of SEA14T, SAR14T, and HMD10T is lower than the WAT of HRLP16T. Because of the structure of cross-coupled inverters, the value change of one node can influence the values of other nodes immediately. SRRD12T has the largest silicon area. Therefore, compared to other SRAMs, the HRLP16T is moderate in terms of WAT.
We can observe from Fig. 7-(c) that the power of SAR14T is sensitive to supply voltage variation mostly, and the power of RHM12T is less sensitive to variation supply voltage. Therefore, the HRLP16T is moderate in terms of power sensitivity to supply voltage variations. We can observe from Fig. 7-(d)(e)(f) that RAT, WAT and power consumption increase with temperature increases. By observing the slope of the broken lines in Fig. 7-(d), we can see that SRRD12T is more sensitive to temperature variation. S4P8N is less sensitive to temperature variation. Therefore, the HRLP16T is moderate in terms of RAT sensitivity to temperature variations.

It can be seen from Fig. 7-(e) that the WAT of RHBD10T is sensitive to temperature variation mostly. The WAT of S8P4N is less sensitive to temperature variation. Therefore, the HRLP16T is moderate in terms of WAT sensitivity to temperature variations.

Fig. 8 shows the estimation results of threshold-voltage/LEEF variation impacts on RAT, WAT, and power for the SRAM designs. We can observe from Fig. 8-(a)(b)(c) that RAT and WAT increase and power consumption decreases with threshold-voltage increases. By observing the slope of the broken lines in Fig. 8-(a), we can see that SRRD12T and SAR14T are more sensitive to threshold-voltage variation. S8P4N is less sensitive to threshold-voltage variation. Therefore, the HRLP16T is moderate in terms of RAT sensitivity to threshold-voltage variations. It can be seen from Fig. 8-(b) that the WAT of RHM12T is sensitive to threshold-voltage variation mostly. The WAT of SEA14T is less sensitive to threshold-voltage variation. Therefore, the HRLP16T is moderate in terms of WAT sensitivity to threshold-voltage variations.

We can observe from Fig. 8-(c) that the power of HRLP16T is sensitive to threshold-voltage variation mostly, and the power of RHMD10T is less sensitive to threshold-voltage variation. We can observe from Fig. 8-(d)(e)(f) that RAT, WAT and power consumption increase with LEEF increases. By observing the slope of the broken lines in Fig. 8-(d), we can see that SRRD12T is more sensitive to LEEF variation. S4P8N is less sensitive to LEEF variation. Therefore, the HRLP16T is moderate in terms of RAT sensitivity to LEEF variations.

It can be seen from Fig. 8-(e) that the WAT of RHBD10T is most sensitive to LEEF variation. The WAT of S8P4N is...
less sensitive to LEEF variation. Therefore, the HRLP16T is moderate in terms of WAT sensitivity to LEEF variations. We can observe from Fig. 8-(f) that the power of RHMD10T is sensitive to LEEF variation mostly, and the power of Lin12T is less sensitive to LEEF variation. Therefore, the HRLP16T is moderate in terms of power sensitivity to LEEF variations. In summary, the overhead of the proposed SRAM has moderate sensitivity to PVT variations.

IV. CONCLUSIONS

SRAMs is becoming severely susceptible to radiative particle-striking-induced soft errors, such as SNU and DNU, as the CMOS technology aggressively scales down. This paper has proposed a novel SRAM cell design, namely HRLP16T, with better trade-off between reliability and overhead. The proposed cell can not only recover from any SNU but also many DNU. The proposed SRAM cell also has moderate sensitivity to PVT variations. Therefore, the proposed HRLP16T SRAM can be feasibly applied to aerospace applications where high reliability and better performance are required.

REFERENCES


