Industrial Engineering Journal
ISSN: 0970-2555
Volume : 54, Issue 3, No.1, March : 2025

ENHANCING BUILT-IN-SELF-REPAIR (BISR) IN SRAM USING MACHINE LEARNING

Kanteti Himabindu, Research Scholar, Department of Electronics and Communication Engineering, JNTU-GV
Dr. K.Babulu, Professor & Director of evaluation and Admissions, Department of Electronics and Communication Engineering, JNTU-GV.
Dr.G.N.Swamy, Professor & Head of EIE, VRSEC, Vijayawada.

ABSTRACT

To increase fault tolerance and overall operational efficiency of embedded memories in SOC, it is essential to use Built in Self Repair (BISR) processes. Many redundancy algorithms have evolved to improve the yield of the memories. Each algorithm has its own targeted parameter improvement methodology. In order to get optimal repair efficiency and repair time the designer has to do trade off with more Chip Area by which the cost increases. In order to meet the optimal values for multi parameters ,this paper discusses the need of Machine Learning technique along with conventional Built in self repair. Based on recent research, we understand that by imposing ML models we can achieve better metrics, including improvement in testing and repair time ,smaller chip Area, power etc.

Key Words : Built in Self Repair – Fault Tolerance - Machine Learning- SRAM in SOC.

INTRODUCTION :

Static Random Access Memory (SRAM), especially in relation to System-on-Chip (SoC), has been under enormous amounts of strain because of the dense structure. Due to these dense structures memories are highly prone to defects and failures because of manufacturing procedures. These defects can be both temporary and permanent which leads to memory failures, because of which system reliability reduces. To mitigate these failures, strong fault-tolerance methods have become essential to improve the system operation and yield. Built-in self repair (BISR) is the most common repair technique used to diagnose and repair, to improve operational performance. Many researchers proposed different methodologies to get optimal repair rate , speed and less area overhead .But almost all methodologies are being traded off with one or the other parameter. an unconscious implementation leads to large area overhead and increase in test and repair time, which directly affects the cost. To address this problem , the advancements of Machine learning (ML) in the field of Computer vision, have paved the path for its potential application in this area also. ML has become an important area of research with several promising opportunities at various levels of hardware design. We can understand the significance of ML in VLSI to improve system performance, increase the reliability and speed up the design process and drive the advancement of chip manufacturing processes.

2.METHODS & DISCUSSION:

2.1 BISR :

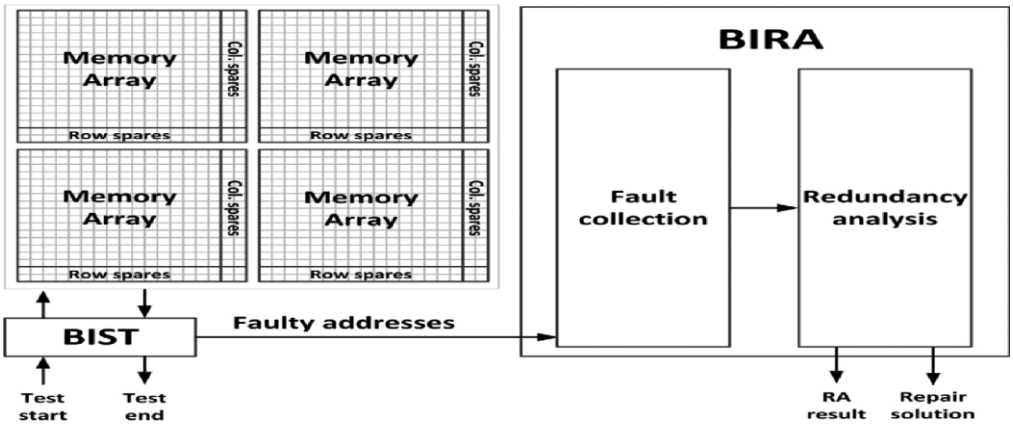
The testing and repair analysis of Memories are performed by automatic test equipment (ATE) or by BISR. ATE is external test equipment, most of the commercial memory chips are tested directly. ATE transmits the test patterns to the chip and receives the data. After receiving the faulty information, it analyses and finds the repair solution. then the solution is used to replace the faulty cell with spare cells of the chip. but where as in SOCs accessing the Memories with ATE is very difficult. So they are tested and repaired by BISR. An additional hardware logic is used as BISR along with memories which leads to extra area overhead for the Chip. Many researchers proposed different Heuristic and Exhaustive algorithms to solve the spare allocation problems. In [2,3] authors surveyed different algorithms proposed by different researchers over a period of time to repair the memories efficiently.

In order to repair the faulty cells of Memories, BISR consists of two steps Built in Self-test (BIST) and Built in repair analysis (BIRA).structure of BISR is shown in Fig1.BIST generates the test patterns to test the memories and send the information to BIRA in order to analyze the faulty

information with the help of redundancy analysis after analyzing the faulty information a repair solution is generated by BIRA. Repair solution gives the information of replacement of faulty cells, faulty rows and columns with available spare cells, spare rows and spare columns in an optimal way.

The performance criteria to assess BIRA is repair rate, speed and area overhead.

Repair rate indicates capability of repairing the memories and obtaining the proper repair solution. Generally, the repair rate is the ratio of the number of memories repaired by the BIRA procedure to the total number of tested memories. The tested memories include all the reparable and irreparable memories.



1. Structure of Memory BISR

To get optimal repair rate and repair speed some BIRA methodologies sacrifice large area overhead by which the cost also increases. In some applications where area overhead is critical Repair rate and Speed have to be sacrificed. Heuristic BIRA algorithms are unable to reach high repair efficiency. whereas Exhaustive BIRA algorithms can guarantee complete memory repair. But these algorithms require

Long execution time with large area overhead. Balancing these three parameters is one of the difficult tasks for the hardware designers. Though many Redundancy algorithms are proposed, redundancy allocation becomes an NP-complete problem. Therefore, there are trade-offs among the three performance criteria,and improving the performance of BIRA algorithms is the focus of current research in this field.

2.2 Machine Learning:

Recent developments in Machine learning (ML) are receiving enormous attention in most fields due to the capability of ML in modelling complex systems with the help of historical data. Machine Learning uses advanced models-based algorithms for prediction. These models include predictive models and neural network-based models that are used for systems to make necessary decisions**.**

3. Scope of application of ML on SRAM:

From a decade Hardware developers and researchers are exploring different possibilities of utilising the ML to solve the existing problems of the VLSI field in a better way. For instance, Performance and power estimation with the help of ML before implementation is the main focus of paper (4). The proposed work is based on Regression tree which is a very popular algorithm in ML. For size optimization Principal component analysis (PCA) and Evolutionary algorithms are used in paper (5). several ML algorithms, namely Gradient Tree Boosting (GTB), Random Forest Regressor (RF), Decision Tree Regressor (DT), Extra Tree Regressor (XT), and Extreme Gradient Boosting Regressor (XGBR) are used to perform the stability Analysis of SRAM (6).This paper presents a design space exploration framework for an FPGA-based soft processor that is built on the estimation of power and performance metrics using algorithm and architecture parameters. The proposed framework is based on regression trees, a popular machine learning technique, that can capture the relationship of low-level soft-processor parameters and high-level algorithm parameters of a specific application domain, such as image compression. In doing this, power and execution time of an

algorithm can be predicted before implementation and on unseen configurations of soft processors. For system designers this can result in fast design space exploration at an early stage in design.

In (7) authors applied ML techniques to detect faults and anomalies. Several ML techniques are exploited to detect faulty wafers in semiconductor manufacturing. usage of ML with both supervised and unsupervised algorithms used to identify intermittent failures in post silicon validation is proposed in(8). There are certain challenges, such as the NP-complete problem, that make ML quite challenging to find the best solution.

In paper (9) author proposed a novel approach to predict the area of hardware components from specification . This approach successfully predicted the area of real life hardware components such as control and status register interfaces which are very important in embedded systems . with this approach they are able to identify the area 600 times faster than conventional flow.

AnQi Zhang proposes the utilization of CNT SRAM memory (10). Modifications are made to the NeuroSim tool to analyze how different fault patterns affect the memory units after recovering from stuck-at faults. The paper also presents mechanisms for fault recovery that aim to improve system accuracy. By leveraging ML applications to emerging technology like CNT SRAM, it becomes possible to effectively address high fault rates and enhance overall system accuracy.

By observing the research developments of Machine learning in semiconductor technologies, SRAM faulty cell testing and obtaining a high faulty rate , there is a revolutionary chance to improve BISR performance with ML . Early identification and repair of weak or failing SRAM cells is made feasible by the use of machine learning (ML) algorithms to examine performance indicators, real-time monitoring data, and historical failure patterns. The occurrence of faults in contemporary semiconductor manufacturing processes can be addressed by ML-driven systems, which offer adaptive, accurate, and productive oversight of defects in contrast to traditional BISR approaches.

CONCLUSION

The potential of integrating conventional Built-in Self-Repair (BISR) methods with Machine Learning (ML) to improve the fault tolerance and efficiency of embedded memories in SoCs is investigated in this work. Designers can enhance repair processes and reduce testing time, chip area, power consumption, and repair time by incorporating machine learning models. With ML-driven techniques, yield could be improved more efficiently than with traditional redundancy algorithms that require trade-offs between repair efficiency and cost. Although the benefits of machine learning in redefining BISR techniques are highlighted in this paper, more research is required to create workable implementations. Enhancing real-time flexibility, scalability, and repair accuracy in embedded memory architectures should be the main goal of future research on machine learning models.

REFERENCES

1. C.-T. Huang, C.-F. Wu, J.-F. Li, and C.-W. Wu, “Built-in redundancy analysis for memory yield improvement,” *IEEE Transactions on Reliability*, vol. 52, no. 4, pp. 386–399, Dec. 2003, doi: 10.1109/TR.2003.821925

2. K. Cho, W. Kang, H. Cho, C. Lee, and S. Kang, “A survey of repair analysis algorithms for memories,” *ACM Computing Surveys*, vol. 49, no. 3, pp. 1–41, Sep. 2017, doi: 10.1145/2971481.

3. K. Hima Bindu, DR.K. Babulu, G.N. Swamy, “Survey on Heuristic Built in Redundancy Algorithms, Journal of Adv Research in Dynamical & Control Systems, Vol. 11, No. 1, 2019, ISSN 1943-023X.

4.Adam Powell, Christos Savvas-Bouganis, and Peter Y. K. Cheung. “High level power and performance estimation of fpga-based soft processors and its application to design space exploration”. 59(10):1144–1156, November 2013.

5. T. Pessoa, N. Lourenc¸o, R. M. Martins, R. P´ovoa, and N. Horta. “Enhanced analog and rf ic sizing methodology using pca and nsga-ii optimization kernel”. pages 1–4, March 2018

6. Jihene Bouhlila1,Felix Last,Rainer Buchty, Mladen Berekovic (2024), Machine Learning for SRAM Stability Analysis, IEEE International Symposium on Circuits and Systems (ISCAS), DOI: 10.1109/ISCAS58744.2024.10558564.

7. Seungyong Doh. Machine learning-based novelty detection for faultywafer detection in semiconductor manufacturing. Expert Systems with Applications, 39(4):4075–4083, 2012.

8.Andrew DeOrio, Qingkun Li, Matthew Burgess, and Valeria Bertacco. Machine learning-based anomaly detection for post-silicon bug diagnosis. In Proceedings of the Conference on Design, Automation and Test in Europe, pages 491–496. EDA Consortium, 2013.

9.Elena Zennaro,Lorenzo Servadei1, Keerthi kumara Devarajegowda,Wolfgang , A Machine Learning Approach for Area Prediction of Hardware Designs from Abstract Specifications,21st Euromicro Conference on Digital System Design,2018, DOI 10.1109/DSD.2018.00076

10. An Qi Zhang; Amr M.S. Tosson; Lan Wei, "Error Resilience and Recovery of Process Induced Stuck-at Faults in MLP Neural Networks using Emerging Technology", IEEE/ACM International Symposium on Nanoscale Architectures (NANOARCH), 2021