



International Journal of Engineering Research and Science & Technology

ISSN : 2319-5991
Vol. 1, No. 1
March 2015



*National Conference on "Recent Trends in Communication
& Information Technologies" NCRTCIT 2015*

Organized by

Dept. of ECE, Indra Ganesan College of Engg., Trichy, Tamil Nadu, India.



www.ijerst.com

Email: editorijerst@gmail.com or editor@ijerst.com

Research Paper

OPTIMAL REPAIR RATE AND AREA REDUCTION FOR EMBEDDED MEMORIES BY USING BIRA & AMT

V Geetha^{1*}, X Susan Christina¹ and X Susan Christina¹

*Corresponding Author: V Geetha ✉ geeta93selvi@gmail.com

As the capacity and density of embedded memories have rapidly increased, the probability of memory faults will increase. That results in yield drops and quality degradation. Yield improvement of embedded memories have become very important. Yield refers to the percentage of good die on the wafer. For embedded memories Built-In-Redundancy-Analysis (BIRA) is used to achieve optimal repair rate and yield improvement. In this Project the proposed scheme generates set of test patterns with weights 0, 0.5 and 1 using an accumulator-based-3 weighted test pattern generation. In a Built-In Redundant analysis (BIRA) scheme, it performs exhaustive search to find and repair the faults in RAM. It uses part of spare memory as an AMT (Address Mapping Table) to reduce the faulty cell address to its logical address which consumes less area. And by using tiling technique in spare memory permanent faults were handled. These reduced addresses are stored in CAM (Content Addressable Memory) which is used during the Redundant Analysis procedure.

Keywords: Addressable Memory (CAM), Built in redundancy analysis (BIRA), Built in self test (BIST), Accumulator based 3-weighted test pattern generation.

INTRODUCTION

In modern system-on-chip (SOC) designs, embedded memories occupy a significant portion of the chip area. Embedded memory plays a vital role in many electronic devices such as set-top boxes, sensors, industrial meters and mobile phones due to its small size. Embedded memory mainly focuses on the reliability rather than large size, so it is tested for faults using BIST. As the capacity and density of semiconductor memories have rapidly increased with the

technological development of semiconductor manufacturing, the probability of memory faults has increased. This causes yield drops and quality degradation. As the chip area increases yield is reduced. In order to improve the yield, testing is done by the embedded system. Circuit manufacturer must thoroughly test their products before delivering them to the customers. Built-in Redundancy-repair (BISR) techniques are one main yield enhancement technique for embedded RAMs with redundancy.

¹ M.I.E.T Engineering College, Trichy, Tamilnadu, India

Embedded memory is tested by using a built-in self-test (BIST) module or external automatic test equipment (ATE) before RA is performed. According to the testing method, there are two methods to repair faulty memories. They are BIRA module and an external ATE. Most RA algorithms have been based on 2-D redundant cells since the 1980s. Furthermore, various RA algorithms for 2-D redundancy have been developed. They are Repair most (RM), CRESTA, LRM, and ESP are the most well known RA algorithms for built-in redundancy analysis (BIRA). **RM** (Repair most) is a simple algorithm. It counts the number of faults in each line and allocates the spare lines to a faulty line in descending order of the number of faults. Although its repair rate is high, it is not optimal, because an exhaustive search is not performed. **CRESTA** (Comprehensive Real-time Exhaustive Search Test and Analysis) focuses on the optimal repair rate and fast analysis. But the hardware cost is high. **LRM** (Linear Repair Most) and **ESP** (Essential Spare Pivot) mainly focus on minimizing the area overhead in terms of storage requirements with a simple RA algorithm. However, the repair rates of the two algorithms are not optimal since faulty information is omitted. **IS** (Intelligent Solve) and **ISF** (Intelligent Solve First) algorithms achieve both low area overhead and an optimal repair rate. However, the two algorithms take a lot of time to complete the RA in cases with complex fault distributions.

Recently, many tree-based RA algorithms have also been introduced that make an effort to lessen the search space. These algorithms are based on the branch and bound (B&B) algorithm. The B&B algorithm is a simple fault-driven approach. To overcome the restriction of these approaches, SFCC and BRANCH analyzer were

introduced. **SFCC** (Selected fault count comparison) achieves a low area overhead and optimal repair rate, and reduces the search space by building a search tree based on the line faults. **BRANCH** analyzer analyzes all nodes concurrently within a branch for combinations of 2-D spares. In order to analyze all nodes concurrently and achieve an optimal repair rate, BRANCH analyzer and SFCC store all fault information into fault-storing content-addressable memories (CAMs). Most BIRA approaches require extra hardware overhead in order to store and analyze faults in the memory. Existing approaches do not utilize spare memories during the redundancy analysis (RA) procedure. Proposed system uses part of the spare memory as an Address Mapping Table (AMT) in order to reduce the area overhead created by Built-In Redundancy Analysis (BIRA) scheme.

RELATED WORK

A. Fault Classification

A fault in a system is some deviation from the expected behaviour of the system: a malfunction. Faults in embedded memory may be due to variety of factors, including hardware failure, operator (user) error. Faults can be classified into one of three categories:

Transient Faults

These occur once and then disappear. For example, a network message doesn't reach its destination but does when the message is retransmitted.

Intermittent faults

Intermittent faults are characterized by a fault occurring, then vanishing again, then reoccurring, then vanishing. These can be the most annoying

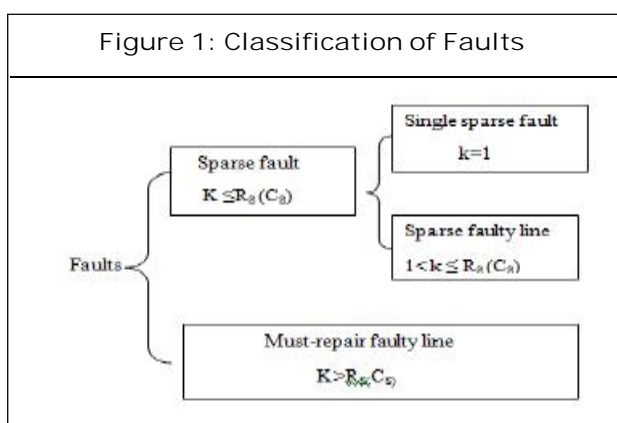
of component faults. A loose connection is an example of this kind of fault.

Permanent Faults

This type of failure continues to exist until the faulty component is repaired or replaced. Examples of this fault are disk head crashes, software bugs, and burnt-out power supplies.

The **failure pattern** is classified into six types: word line fault, bit line fault, cluster fault, continuous fault, twin fault, and single fault. Word (bit) line fault is column (row) failures. Cluster fault is formed of various shapes by distribution of faulty cells. Continuous fault is a line fault which has greater than two faults, i.e., faulty cells are continuously distributed. Twin fault is formed by a line fault or a diagonal fault by two faults. A single fault does not share a row and column address with other faults. Since a memory with 2-D spare architecture is repaired by line replacement, faults (failure pattern) were classified into three types:

- single fault
- sparse line fault,
- must-repair line fault.



Diagonal fault of twin fault and single fault can be defined as a single fault. Sparse line fault is a faulty row or column line.

The number of spare rows and columns are R_s and C_s , respectively. A faulty row (column) line has more than one and less than or equal to C_s (R_s) faults. Cluster fault, continuous fault, and line fault of twin fault can be defined as a sparse line fault. A must-repair faulty row (column) line has greater than C_s (R_s) faults. Word line fault, bit line fault, part of cluster fault, and continuous fault can be defined as must-repair line fault. Because the cluster faults contain many cross line faults, the cross line faults generate complicated cases for RA.

B. Spare Memory

Spare memory is a division of memory. Embedded memory is divided into two channels. The first channel is mirrored to the second channel, creating a redundant copy of memory. If a fault occurs within the memory of one channel, the memory controller shifts to the paired channel without disruption, and the channels can re-synchronize when repairs are completed. In sparing mode, the trigger for failover is a preset threshold; when the threshold is reached, the content is copied to its spare counterpart and it is activated for use.

C. Content Addressable Memory

CAM is designed to search its entire memory in a single clock cycle, it is much faster than RAM in virtually all search applications. CAMs are widely used wherever fast parallel search operations are required. CAM is a storage array designed to find the location of a particular stored value. The main difference between a Content Addressable Memory and a Random Access Memory (RAM) is that, in a RAM the user supplies the address and gets back the data and in a CAM the user supplies the data and gets back the address. In BIRA the cam can be used to store the

reduced addresses given by AMT for analyzing the fault address using RA procedure.

Redundancy Analysis (RA)

After the fault collection, Redundancy Analysis(RA) procedure is done by the embedded system to find repair solutions. In order to achieve an optimal repair rate, an RA algorithm, which is an exhaustive search method, is used. Optimal repair rate is the rate of occurrence of failure incidences for a repairable system. It is defined as the ratio of memories repaired by redundancy analysis to all repairable of the total tested memories.

Optimal repair rate =

$$\frac{\text{Memories repaired by BIRA}}{\text{All repairable memories of the total tested memories}}$$

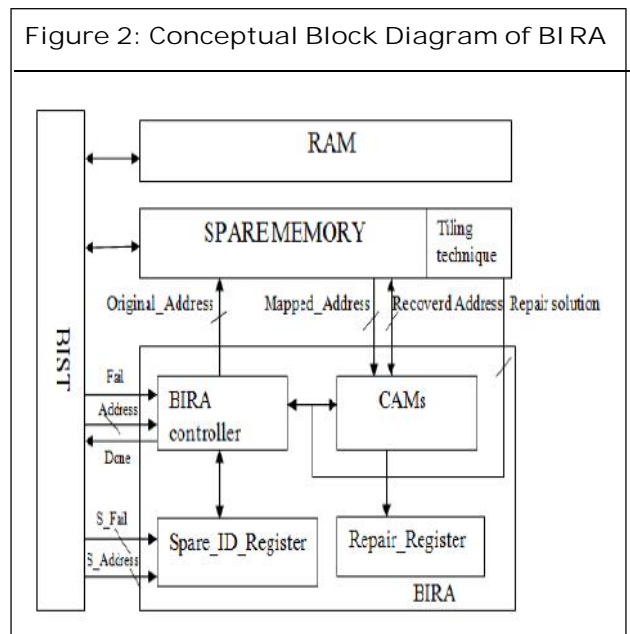
During Redundancy Analysis, all the solution candidates are applied to the faults according to the number of faults. If there is one solution at least, then the memory is repairable. If there is no solution after all the solution candidates are applied to the faults, then the memory is irreparable. If the number of faults is less than the sum of the available spare memories, there may be no solution by applying solution candidates to the row and column combinations of the faults.

PROPOSED WORK

The conceptual block diagram of a built-in redundancy analysis (BIRA) is shown below, which is composed of BIST and BIRA. Here the spare memory is used as an AMT, so it is tested by BIST before the RAM is tested. This BIRA scheme has two modes, 1) Test mode 2) Normal mode.

In the test mode, BIST detects the faults in the RAM and sends the fault information to BIRA.

BIRA can then change the fault addresses into short mapped logical addresses. The BIST tests the RAM and the BIRA determines the redundancy allocation for efficiently replacing the defective elements of the RAM.



The proposed BIRA consists of BIRA_Controller, fault storing CAMs, Spare_ID_register, and Repair_register. The controller performs several procedures: it compares faulty addresses, stores the original addresses in the AMT, stores the mapped addresses in the CAMs, and recovers the mapped addresses to the original addresses.

The mapped addresses are stored in the fault storing CAMs. The faulty addresses of the faulty spare memories are stored in the Spare_ID_register. The faulty spare memory is not suitable to use for the AMT as also to replace faulty memory lines. and by using the tiling technique in the spare memory the permanent faults are handled. If detect any permanent faults in memory the testing control avoid the prohibited zone to apply the process so it will avoid usage of the faulty zone. It will reduce the processing time and

increases the processing speed. That is, BIRA schemes perform RA procedures taking into account the amount of remaining good spare memory. The repair registers store the repair solutions, which are the memory addresses from the BIRA controller. Once the RA procedure is completed, the spare memories are allocated according to the information in the Repair register.

The spare memory is tested before the RAM is tested. Because the RAM cannot be repaired by a faulty spare memory and we do not use the faulty spare memory as an AMT, the address of faulty spare memory should be stored in the Spare_ID_register. When BIST detects a fault from the spare memory, the signal S_Fail is asserted, and the faulty addresses are sent to the Spare_ID_register through the port S_Address. The BIRA_Controller counts the available spare memories and checks faulty spare memory to perform an adequate RA procedure.

After the test for the spare memory is done, the RAM is tested. When BIST detects a fault, it pauses and the signal Fail is asserted. The faulty addresses are then sent to BIRA through the port Address. If the fault is the first one, the BIRA_Controller stores the fault address to the AMT by avoiding faulty spare memory in the Spare_ID_register. Otherwise, the BIRA_Controller compares the addresses with already stored addresses in the AMT through the port Original_Address.

If the address is the same, its mapped addresses are sent to the BIRA_Controller through the port Mapped_Address and are stored in the CAMs. Otherwise, the address is stored in the AMT and mapped to short logical addresses. These mapped addresses are sent to the BIRA_Controller through the port Mapped_

Address and stored in the CAMs. Once the mapped addresses are stored in the CAMs, the BIRA_Controller asserts the signal Done, and BIST continues the test.

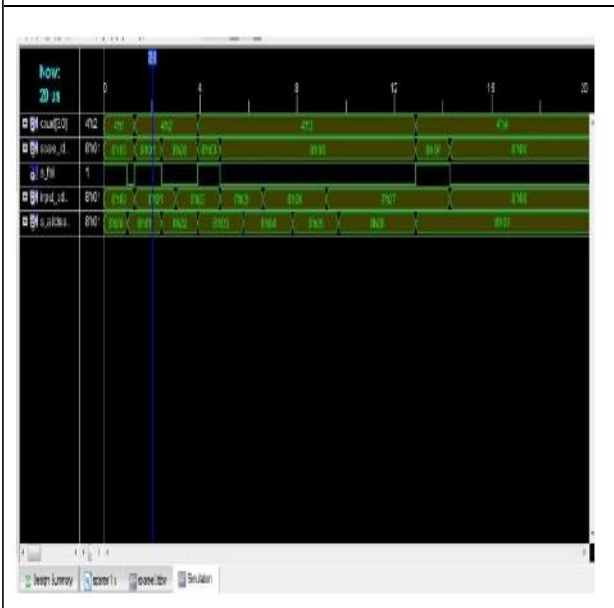
After the test is completed, BIRA finds the repair solution by analyzing the fault information of the CAMs using the proposed RA algorithm. Because the repair solutions are composed of the mapped addresses, they are recovered to the original addresses through port Recover_Address. Finally, the repair solutions, which are composed of the original addresses, are stored in the Repair_register through the port Repair_Solution to replace the faulty lines of RAM with spare memories.

Proposed BIRA uses the Content Addressable Memory to store the logical address of the faulty cells in RAM memory. When a fault is detected during the BIST procedure, the BIRA controller stops the testing and stores the corresponding faulty cell address in the spare memory and it is mapped to logical address. This logical address is then stored in CAM. When CAM receives an input faulty address, it searches the already stored data for match. If the incoming address is not available in its stored data, then it stores the faulty incoming address. This CAM content is used during the repairing procedure in order to recover the original address.

SIMULATION RESULTS

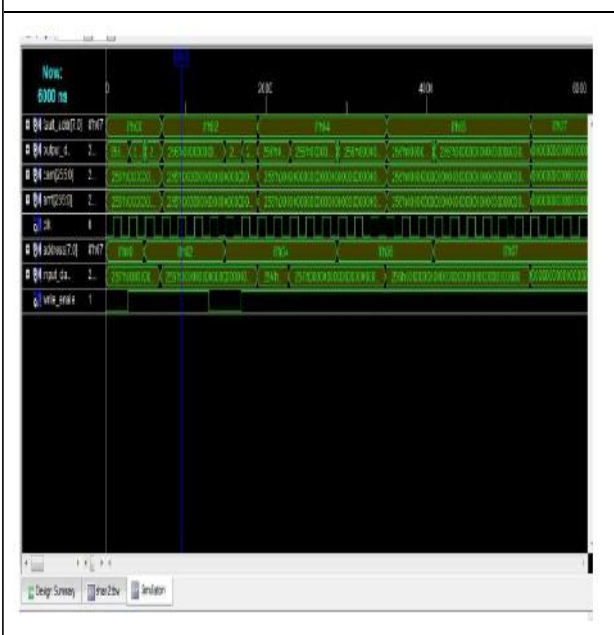
Spare memory is tested before RAM, because faulty spare cells cannot be used as Address Mapping Table. Input_address and the s_address are the inputs. The input_address and the spare cell address i.e. s_address are compared in order to find the Address non-uniqueness type fault. If two addresses are same then it is considered as fault because, two cells

Figure 3: Simulation Result for Fault Detection in Sparememory



have same address i.e. address is not unique one and it is declared as fault. Whenever a fault occurs, the s_fail signal is asserted and the corresponding faulty address is stored in spare_id_register. Number of spare faults is counted by the variable count.

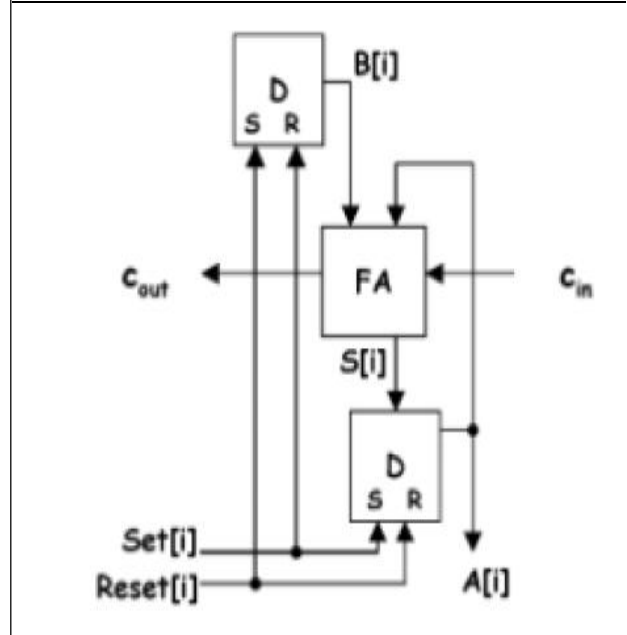
Figure 4: Simulation Result for BIRA Scheme



FURTHER IMPLEMENTATION

BIST is a design-for-testability technique that places the testing functions physically with the circuit under test (CUT). The basic BIST architecture requires the addition of three hardware blocks to a digital circuit: a test pattern generator, a response analyzer, and a test controller. The test pattern generator generates the test patterns for the CUT. An accumulator-based 3-weight test pattern generation scheme generates set of patterns with weights 0, 0.5, and 1. Since accumulators are commonly found in current VLSI chips, this scheme can be efficiently utilized to drive down the hardware of built in self test pattern generation, as well. The implementation of the 3-weighted-test pattern generation scheme is based on the full adder. Therefore, in order to transfer the carry input to the carry output. The logic module provides the Set and Reset signals that drive the Set and R inputs of the Register A and Register B inputs.

Figure 5: Accumulator cell test vector Generation Operation of the cell



- I. For $A[i] = 1$, We give $set[i]=1$ and $reset[i]=0$ and hence $A[i]=1$ and $B[i]=0$. Then the output is equal to 1, and Cin will be equal to Co . Cin is transferred to the $Cout$.
- II. For $A[i] = 0$, We give $set[i]=0$ and $reset[i]=1$ and hence $A[i]=0$ and $B[i]=1$. Then the output is equal to 0, and here Cin is equal to $Cout$. Cin is transferred to the $Cout$.
- III. For $A[i] = "-"$, $set[i] = 0$ and $reset[i] = 0$. The D input of the flip-flop of register B is driven by either 1 or 0, depending on the value that will be added to the accumulator inputs in order to generate satisfactorily random patterns to the inputs of the CUT.

The signals that drive the S inputs of the flip-flops of Register A , also drive the R inputs of the flip-flops of Register B and vice versa. The implementation of the proposed weighted pattern generation scheme is based on the accumulator cell is given in Figure 1, which consists of a Full Adder (FA) cell and a D-type flip-flop with asynchronous set and reset inputs whose output is also driven to one of the full adder inputs. This technique is used without loss of generality, that the set and reset are active high signals. The respective cell of the driving register $B[i]$ is also shown.

CONCLUSION

A Built-in Redundancy Analysis scheme for embedded memories is proposed. Fault collection in spare memory followed by RAM, address mapping and repair analysis are some of the steps in BIRA. Faults due to address nonuniqueness are considered here and fault addresses are collected using BIRA controller. Collected faulty addresses are mapped using Address Mapping Table which resides in spare memory. Mapped

faulty addresses are stored finally in Content Addressable Memory. And by using the tiling technique in spare memory permanent faults were handled. If detect any permanent faults in memory the testing control avoid to apply the process to that zone. So It will reduce the processing time so speed will increase. An accumulator based 3 weighted test pattern generation will be used for generating test sequence. It can be implemented using any adder design and then it does not require any modification of the adder; hence it does not affect the operating speed of the adder. By comparing existing system this proposed system is advanced for embedded memories.

REFERENCES

1. Chang D.M. Li J.F. and Tseng T.-Wand (2006), "A built-in redundancy-analysis scheme for RAMs with 2-D redundancy using 1-D local bitmap," in Proc. Design, Autom. Test Eur., pp. 53–58, Mar.
2. Cho H. and Kang W. et al (2010), "A built-in redundancy analysis with a minimized binary search tree," ETRI J., vol.32, no. 4, pp. 638–641, Aug.
3. Chen T.-J. and Li J.-F. et al (2012), "Cost-efficient built-in redundancy analysis with optimal repair rate for RAMs," IEEE Trans. Comput.- Aided Design Integr. Circuits Syst., vol. 31, no. 6, pp. 930–940, Jun.
4. Han, T.W. and Jeong et al (2010), "An advanced BIRA for memories with an optimal repair rate and fast analysis speed using a branch analyzer," IEEE Trans. Comput.- Aided Design Integr. Circuits Syst., vol. 29, no. 12, pp. 2014–2026, Dec.
5. Jeong W., Jin and Kang I. et al (2009), "A Fast

- Built-in Redundancy Analysis for Memories With Optimal Repair Rate Using a Line-Based Search Tree", *IEEE Trans Very Large Scale Integr. (VLSI) Systems.*, vol 17, no 12, pp. 1665-1678, Dec.
6. Jeong W., Kang I., Jin K. and Kang S. (2009), "A fast built-in redundancy analysis for memories with optimal repair rate using a line-based search tree," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 17, no. 12, pp. 1665–1678, Dec.
 7. JooHwan Lee, Kihyun Park and Sungho Kang (2012), "High-efficiency BIRA for embedded memories with a high repair rate and low area overhead", *Journal of semiconductor technology and science*, vol. 12, no. 3, pp. 515–517, sep.
 8. Lee J., Park K. et.al (2009), "An area-efficient built-in redundancy analysis for embedded memories with optimal repair rate using 2-D redundancy," in *Proc. Int. SoC Design Conf.*, , pp. 353–356, Nov.
 9. Li J.-F, Tseng T.W. et.al (2010), "ReBISR: A Reconfigurable Built-In Self-Repair Scheme for Random Access Memories in SOCs" , *IEEE Trans Very Large Scale Integr. (VLSI) Systems.*, vol 18, no 6, pp. 921-937, Jun.
 10. Pagiamtzis K and Sheikholeslami A. (2006), "Content addressable memory (CAM) circuits and architecture: A tutorial and survey," *IEEE J. Solid-State Circuits*, vol. 41, no. 3, pp. 712–727, May.
 11. Wooheon Kang, Hyungjun Cho, JooHwan Lee, and Sungho Kang, Member, IEEE., "A BIRA for Memories With an Optimal Repair Rate Using Spare Memories for Area Reduction", *IEEE transactions on very large scale integration (vlsi) systems*, vol. 22 pp. 2336-2349 October 2013



International Journal of Engineering Research and Science & Technology

Hyderabad, INDIA. Ph: +91-09441351700, 09059645577

E-mail: editorijerst@gmail.com or editor@ijerst.com

Website: www.ijerst.com

