Re-Configurable Built In Self Repair scheme in Ram for Yield Improvement

Tessy M John¹, Dhanya Oommen²

¹P G Scholar, VLSI and Embedded Systems, Department of ECE, T K M Institute of Technology, Kollam, India
²Assistant Professor, Department of ECE, T K M Institute of Technology, Kollam, India

Abstract: The usage of embedded memories is more than half of the die area for a typical system on chip (SOC). Due to the complexity of memory architectures, the possibility of occurring manufacturing defects is more. Hence memory testing is necessary. Built In Self Test (BIST) has been proven to be most cost-effective and widely used solutions for memory testing. It is a mechanism that allows a machine or a circuit to test itself. Built In Self Repair (BISR) techniques are widely used for repairing embedded memories. The earlier repairing methods have low repair rate and takes lot of time to repair. BISR technique is used to reduce repair time and to increase the repair rate. The design architecture is simulated using ModelSim and Xilinx ISE 13.1 tools. Most of the repairing are based on allocating some redundancy to memory elements. For allocating redundancy, redundancy analysis is necessary. An efficient redundancy analysis method called Re-Configurable Built In Redundancy Analysis is used to improve repair rate and to reduce repair time.

Keywords: SOC, BIST, BISR, BIRA, Test Pattern Generator, FPGA.

1. Introduction

Very Large Scale Integration (VLSI) had a dramatic impact on the growth of digital technology. VLSI has not only reduced the size and the cost but also increased the complexity of the circuits. These improvement have resulted in significant increase in performance in circuits. But it create some potential problems, which may retard the effective use and growth of future VLSI technology. Development in VLSI technology results in continuously increasing density of memory chips. The exponential increase in density creates yield improvement and testing issues. As the feature size of component shrinks, the sensitivity to faults also increases. Built In Self Test (BIST) can solve the memory testing problems, which increases the predictability. Test patterns generated by a BIST controller can be either deterministic or pseudo random. Built In Self Repair (BISR) is the most efficient technique, that is used to repair faults and to improve the yield of memory. To increase memory yield, many manufactures use incorporated redundancy to replace faulty cells.

Embedded random access memory (RAM) is one key component in modern complex system-on-chip (SOC) designs. Typically, many RAMs with various sizes are included in an SOC, and they occupy a significant portion of the chip area. Furthermore, RAMs are subject to aggressive design rules, such that they are more prone to manufacturing defects. That is, RAMs have more serious problems of yield and reliability than any other embedded cores in an SOC. Keeping the RAM cores at a reasonable yield level is thus vital for SOC products. Built-in self-repair (BISR) technique has been shown to improve the RAM yield efficiently. Built-in redundancy-analysis (BIRA) algorithm is one key component of a BISR scheme, and it is responsible for allocating redundancies of memory under test. Thus, the BIRA circuit has heavily influence on the repair efficiency of the BISR scheme. If a memory has only spare rows or spare columns, i.e., 1-D redundancy, then the redundancy allocation is simple and straightforward.

In this work, BIST circuitary generate test patterns and it is given to BIRA. Based on a suitable redundancy analysis algorithm BIRA circuit allocate redundancy and repair RAM under test. The coding can be synthesized by the Xilinx ISE Design Suite 14.2, simulated using ModelSim simulator and can be implemented using Spartan 3E FPGA.

2. Related Works

Many works has been carried out in the field of testing and repairing memory. Built-in redundancy-analysis (BIRA) algorithm is one key component of a BISR scheme, and it is responsible for allocating redundancies of RAMs under test. Thus, the BIRA circuit has heavily influence on the repair efficiency of the BISR scheme.

If a RAM has spare rows/spare columns or spare rows/spare IOs, i.e., 2-D redundancy, the redundancy allocation becomes an NP-complete problem [20]. To achieve the best repair efficiency, some BISR schemes use exhaustive search algorithms to allocate 2-D redundancy [17]. But, either the area cost or the test/repair time of these schemes is very high. In [14] and [9], the authors use parallel BIRA modules to allocate redundancy optimally. However, the number of BIRA modules is drastically increased with respect to the number of redundancies of a RAM under test. This results in very high area cost. In [18], Oehler et al. use a single BIRA module to allocate redundancy optimally. The BIRA module is programmable and one repair strategy is programmed to allocate redundancies each time. Therefore, if the programmed repair strategy cannot repair the memory under test, then another repair strategy is programmed in the BIRA module and the memory under test is retested. This process is repeatedly until a successful repair strategy is found or all possible repair strategies are tried. Although the area cost of
the BIRA circuit is reduced, this results in very high time cost of test and repair. Therefore, some of existing BISR schemes use heuristic redundancy analysis algorithms to allocate the redundancy of RAMs under test. These heuristic analysis algorithms attempt to optimize the repair efficiency and minimize the area cost and the time of test/repair. For example, the BIRA module of the BISR schemes reported in [19], and [20] performs heuristic redundancy analysis algorithms to allocate the redundancy. Due to the importance of redundancy analysis algorithms, furthermore, some research works focus on the development of efficient heuristic redundancy analysis algorithms [12]. To boost the reliability of a RAM in life time, some BISR schemes integrating online repair functions have been proposed in [10].

3. Methodology

The various components of BISR scheme are BIST circuitry, redundancy analyzer, Re-configuration mechanism, and a repirable RAM.

3.1 System Overview

The basic block diagram of BISR scheme is shown in figure 1.

![Figure 1: Block Diagram of Repair scheme](image)

The overall repair flow is roughly described as follows. Firstly, the BIST tests the repairable memory. If a fault is detected, then the fault information is stored in the BIRA (Built In Redundancy Analyzer) circuit. Then, the BIRA circuit allocates redundancies to replace defective elements. If the number of available redundancies is less than the number of faults in memory, then the BIST is stopped and memory is ir-repairable. If the number of available redundancies is greater than the number of faults in memory, then the defective cells are replaced by spare elements.

3.2 BIST

BIST is a mechanism that permits a machine to test itself. The memory BISR (MBISR) concept contains an interface between memory BIST (MBIST) logic and redundancy wrapper for storing faulty addresses. Test Pattern Generator (TPG) Responsible for generating the test vectors according to the desired technique (i.e. depending upon the desired fault coverage and the specific faults to be tested for) for the CUT. Linear feedback shift register (LFSR) and pseudo random pattern generator (PRPG) are the most widely used TPGs. The block diagram for the BIST is shown in the Figure 2.

![Figure 2: Block Diagram of BIST](image)

3.3 Dual Port RAM

In this paper, we have to check that whether there is an error or not in RAM under test. A Dual Port RAM is considered. The block diagram for the Dual Port RAM is shown in the Figure 3.

![Figure 3: Block diagram of Dual Port RAM](image)

Dual-ported RAM (DPRAM) is a type of Random Access Memory that allows multiple reads or writes to occur at the same time, or nearly the same time, unlike single-ported RAM which only allows one access at a time. A RAM should be able to store many words, one per address, read the word that was saved at a particular address and Change the word that is saved at a particular address. An address will specify which memory value we are interested in. Each value can be a multiple-bit word (e.g., 32 bits). A Chip Select, CS, enables or disables the RAM. ADDRESS specifies the address or location to read from or write to. WR selects between reading from or writing to the memory. To read from memory, WR
should be set to 0. OUT will be the n-bit value stored at ADDRESS. DATA is the n-bit value to save in memory.

3.4 Built In Self Repair scheme

A BISR circuit for an RAM with 2-D redundancy (i.e., spare rows and spare columns) typically consists of a built-in self test (BIST) module, a built-in redundancy-analysis (BIRA) module, and a reconfiguration mechanism. The BIST tests the RAM and the BIRA determines the redundancy allocation for efficiently replacing the defective elements of the RAM.

A typical BISR flow for a 2-D RAM is described as follows. When the power of the RAM with BISR circuit is turned on, the BIST circuit generates test patterns to check whether the RAM under test is defective. If a fault is detected, the fault is stored in the BIRA circuit which performs redundancy allocation on the fly. During the BIRA process, if the number of available redundancies is exhausted, then the RAM is irreparable and the BIST is stopped. On the other hand, if the number of available redundancies is not exhausted, then the BIST and BIRA are performed until the testing of the RAM is completed and then the BIST is stopped.

3.5 Redundancy Analysis Algorithm

For allocating redundancy redundancy analysis is necessary. Various redundancy analysis algorithms are used. An efficient algorithm called Range Checking First Algorithm (RCFA) is used in this method. The flow chart of algorithm is shown in figure 4.

If the BIST circuit detects a fault, then the fault information is exported to the ReBIRA circuitry, and then the ReBIRA performs redundancy allocation on the fly using the rules of the implemented redundancy algorithm. The redundancy algorithm implemented in this scheme is Range Checking First Algorithm (RCFA). The ReBIRA allocating redundancy on the fly means that the redundancy allocation process and the BIST process are performed concurrently. The proposed ReBIRA scheme uses a local bitmap (i.e., a small bitmap) to store fault information of the faults detected by the BIST circuit.

Once the local bitmap is full, the BIST is paused and the ReBIRA allocates redundancies according to the faulty information. After the ReBIRA allocates a redundancy to repair a corresponding faulty row or column, the local bitmap is updated and the BIST is resumed. This process is iterated until the test and repair process is completed. The repair signatures from the ReBIRA circuit are then sent to the repair registers that are present in the repairable RAMs. Repair signatures include repair register data (defective row/column address), repair register address (the address location in the row/column repair registers for storing the defective row/column address) and repair register write signal. The repairing procedure involves the entering of the repair register data in the repair registers.

3.6 Re-BISR Scheme

A typical SOC contains multiple RAM. If each repairable RAM uses one self contained BISR circuit (Dedicated BISR scheme), then the area cost of BISR circuits in an SOC becomes high. This results in converse effect in the yield of RAMs. This paper presents a reconfigurable BISR (ReBISR) scheme for repairing RAMs with different sizes and redundancy organizations.

The overall RAM ReBISR flow is described as follows. Before the BIST circuit and the ReBIRA circuit start testing and repairing the RAMs, the RAM configurations (details) should be known by these two circuits. This is done by the FSM, where it generates the necessary control signals that are required for sending the RAM configurations from RAM details table to BIST and ReBIRA circuits. The process of entering the RAM configurations into the RAM details table is described as follows.

When reset is high, all the locations in the RAM details table are filled with zeroes. Else, if the signal program_ram_details is high, the RAM details are entered into RAM details table through the pin ram_details using the write pointer. As the details are entered one by one, the write pointer is incremented by 1. Once the RAM details table is full, the write pointer stops incrementing and holds the value. Once the RAM details table is full, the BIST and ReBIRA circuits start the testing and repairing processes of RAMs one by one.

4 Results and Discussion

The modules are modelled using VHDL in Xilinx ISE Design Suite 13.1 and the simulation of the design is performed using Modelsim SE 6.2c to verify the functionality of the design. A structural model of Built In Self Test is developed.
The Built In Self Test Module contains modules such as test pattern generator, circuit under test and a response analyzer. Coding for Test Pattern Generator, Circuit Under Test and output response analyzer is done. The simulation results are follows. The simulation result of dual port RAM is shown in the Figure 6.

Figure 6: Dual Port RAM waveform

Figure 6 shows the simulation result of Dual port memory. It consist of two ports. The input pins are address, data, chip select (CS), write enable (WE), output enable (OE). Chip select (CS) controls the overall operation of the chip. When write enable (WE)="1", write operation is takes place. When write enable(WE)="0", read operation is takes place. Write means, particular data is write to memory and Read means, particular data is read from memory. For both read and write operation output enable (OE) pine is set to '1'.

The simulation result of dual port RAM with fault is shown in figure 7. Example: address 0= 11100011; data 0= 00011100 when cs=11, we= 1, oe= 1, write operation takesplace. when cs= 1,we= 0, oe= 1, stuck=0, read operation takesplace.

Figure 7: Simulation Result of Dual Port RAM with fault

5 Conclusion

Very Large Scale Integration (VLSI) has not only reduced the size and the cost but also increased the complexity of the circuit. As the scale of integration grows memory testing and repairing become a major issue. If each RAM in an SOC has individual BISR circuit, then the total area of BISR circuits for the RAMs in the SOC is large, since an SOC usually has many RAMs. To reduce the area cost of BISR circuits, a ReBISR scheme for RAMs in an SOC is used. A ReBISR circuit can be shared by multiple RAMs such that the total area cost of BISR circuits in an SOC can be drastically reduced. A reconfigurable BISR scheme for repairing multiple repairable RAMs with different sizes and different numbers of redundancies has been presented.

References


