

Testing Embedded MRAM IP for SoCs

Author

Faisal Goriawalla

Embedded Test & Repair
Product Marketing Manager,
Synopsys

Introduction

The challenges of embedded memory test and repair are well known, including maximizing fault coverage to prevent test escapes and using spare elements to maximize manufacturing yield. With the surge in availability of promising non-volatile memory architectures to augment and potentially replace traditional volatile memories, a new set of SoC level memory test and repair challenges are emerging. With momentum building for Spin Transfer Torque MRAM (STT-MRAM) as the leading flavor of embedded MRAM technology, this white paper focuses on unique test challenges for STT-MRAM on-chip memory while considering needs for automotive applications. To select an appropriate memory test and repair solution for embedded MRAM, designers need to consider factors such as the special needs of performing trimming during production test, augmented memory fault detection algorithms specific to MRAM architectures, and maximizing manufacturing yield of the process sensitive MTJ (Magnetic Tunnel Junction) bit cell.

What Is STT-MRAM?

Embedded memory IP options include STT-MRAM, Phase-Change Memory (PCM), Resistive RAM (ReRAM) and Ferroelectric RAM (FRAM). Each emerging memory technology is different and suited for specific application(s) but STT-MRAM appears poised to go mainstream.

STT-MRAM is a resistive memory technology in which the change in magnetic spin of electrons in the material produces a measurable change in resistivity. Conceptually, each cell consists of two magnets: one that is stationary, and one that can flip. When the magnets are parallel to each other, resistance is low; when the second magnet flips and reverses direction, the resistance is high.

STT-MRAM technology enjoys low power coupled with low cost due to the ability of the magnetic tunnel junction (MTJ) device to be embedded at the back-end of line (BEOL) interconnect layer of the chip with only three additional masks. STT-MRAM enablement is accelerating in commercial foundries, with GlobalFoundries, Intel, Samsung, TSMC and UMC all having publicly announced offerings for SoC designers at 28nm/22nm technologies.

System architects are adopting STT-MRAM technology for low-power MCU designs, such as IoT wearables, that can benefit from smaller die sizes. STT-MRAM often replaces embedded flash for these early adopters. For autonomous vehicle radar SoCs, the data retention and density of STT-MRAM are significant advantages. In the near future, STT-MRAM will be used to replace SRAM in end applications such as hyperscale computing, in-memory computing, artificial intelligence and machine learning.

Maximizing Manufacturing Yield For STT-MRAM IP

Foundries need new equipment not used in conventional CMOS manufacturing such as ion beam etching while improving MTJ bit cell reliability to support the large (1Mbit~256Mbit) memory array densities that some applications require.

While STT-MRAM technology has adequate endurance and read/write latencies, susceptibility to process variation can cause reliability issues. One of the drawbacks of MTJs bit cell is the small read window, i.e., the difference between high and low resistance states is typically just 2-3X. As a result, sensing the value of an MTJ bit-cell is much more difficult than an SRAM bit-cell.

STT switching is a stochastic process. This means that while reducing write current improves energy efficiency, it increases the probability of write errors with degraded yield. To meet an acceptable yield and maintain in-field reliability, designers need to implement a sophisticated ECC solution. Relying solely on redundant elements such as extra rows or columns incurs high area overhead and reduces the density advantages of MRAM. So, unlike traditional CMOS memory technologies, a combination of ECC and redundancy mechanisms are the best approach to overcome the unique stochastic and process variation related manufacturing challenges of MRAM.

ECC math demonstrates that to achieve a certain chip failure rate (CFR), the memory bit failure rate (BFR) that foundries must achieve becomes increasingly stringent at larger array sizes. Assuming random defects for a 64Mb memory array size, an application targeting the most stringent automotive ASIL-D level (equivalent to SoC level FIT rate of 10) would need at least a DECTED (Double Error Correct, Triple Error Detect) level of ECC with foundry achievable levels of BFR for the MTJ bit cell today. While the ECC scheme can be more relaxed (e.g., SECTED—Single Error Correct, Double Error Detect) for consumer applications and/or smaller array sizes, larger array sizes will necessitate an even more complicated ECC mechanism to meet acceptable overall levels of defective parts per million (DPPM) for the end user. Table 1 shows the types of hard and soft errors that two commonly implemented ECC mechanisms can correct.

Type Of Correctable Error(s) / ECC Scheme	SECTED	DECTED
One soft error OR one hard error	Yes	Yes
Two hard errors	No	Yes
One soft error AND one hard error	No	Yes
Two soft errors	No	Yes

Table 1: Comparison of ECC schemes

To maximize manufacturing yield, the memory BIST solution must therefore utilize extra redundant elements in the memory array as well as provide a sophisticated ECC solution (supporting DECTED) to protect larger MRAM macros on the chip.

Support For Optimal Test Algorithms

In addition to regular stuck-at, transition, coupling, address decoder and hundreds of other fault types observed in traditional memories, testing embedded STT-MRAM memory IP also needs to account for architecture specific faults such as Program/Erase Mask and Sector/Chip Erase Faults. Hence embedded STT-MRAM specific faults need to be detected by an extended class of March based algorithms with user flexibility to specify multiple background patterns (e.g., solid, checkerboard) as well as various addressing modes (e.g., fast column, fast row) to ensure highest test coverage. Due to the large embedded MRAM macro sizes, fast algorithms with lower complexity need to be available in the BIST engine to have acceptable ATE production/manufacturing diagnostic test time.

For automotive environments, the SoC designer will need to have the flexibility to run additional customizable algorithms in-field to match system operational constraints. Table 2 shows examples of how different algorithms of varying complexity may need to be executed to match system constraints in different test phases.

Memory test algorithm	SoC test phase	Number of clock cycles in parallel memory test scenario (run time)	Number of clock cycles in serial memory test scenario (run time)
Test algo1 (Low complexity - 8N)	Mission mode	33,000 (0.066 ms)	80,000 (0.16 ms)
Test algo2 (Medium complexity - 16N)	Power on/off	57,000 (0.114 ms)	123,000 (0.246 ms)
Test algo3 (High complexity - 55N)	ATE production/ manufacturing test	193,000 (0.386 ms)	383,000 (0.766 ms)

Table 2: Capability to select memory test algorithm is important for applications such as automotive*

*Source ITC 2017: Advanced Functional Safety Mechanisms for Embedded Memories and IPs in Automotive SoCs

The BIST engine must therefore support extended classes of algorithms to test fault types specific to MRAM, offer flexibility to run different back ground patterns and addressing modes as well as allow user configurability to execute different algorithms during multiple test phases for the SoC.

On-Chip Trimming/Calibration Support

To maximize the (relatively small) MTJ bit cell read and write windows, embedded MRAM IP suppliers typically recommend trimming reference values to be performed during production/manufacturing test. The purpose of this “calibration” is to compute optimal parameters such as read bias and write bias values of the STT-MRAM array based on actual silicon characteristics. The user is expected to load a default set of values in to the embedded MRAM and then perform a series of supplier recommended test sequences by erasing, programming, and reading from the memory. The intent is to store the bias calculation results in another NVM element such as OTP/efuse. These optimal values are then read out of OTP/e-fuse at subsequent chip power-on to bias the MRAM.

These sequences of steps including the bias calculations are well suited to be built into the on-chip memory BIST engine. The advantages of an “on-chip approach are reduced ATE test time and faster time to market, as it negates the need for the user to develop an additional logic/test controller). The user can send the necessary instructions to the BIST engine via a JTAG IEEE1149.1 TAP interface to initiate and manage the MRAM test modes, change the order of bias sequences, select applicable address ranges within the macro to perform tests, and control other operations to mitigate any unexpected process challenges associated with manufacturing ramp of an emerging technology.

After successful trimming, the memory array needs to be tested (and, if necessary, repaired). Faulty bit cells will need to be swapped for redundant rows or columns. The repair signature is computed for the defective memories by the BIST engine followed by a test utilizing the repair resources. These steps are repeated across all voltage domains in the SoC followed by cycling through the process, voltage and temperature (PVT) conditions to capture the cumulative repair signature. On successful test and repair, the memory BIST engine can finally program the OTP/e-fuse with the combined trimming and repair data. A high-level flowchart of this process is shown in Figure 1. The “Fail” outcome simply indicates the status flag from the BIST engine to the user and does not necessarily imply that the part needs to be discarded.

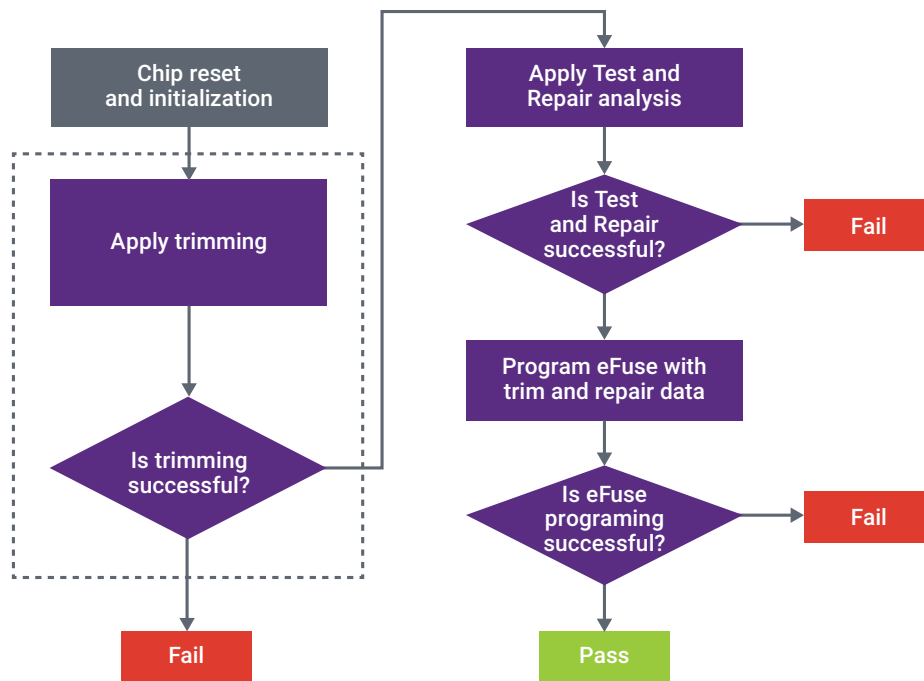


Figure 1: Flowchart showing sequence of operations to be performed by BIST engine for eMRAM

In case of MRAM, the BIST engine must therefore take on the additional task of performing calibration/trimming of reference parameters in addition to the classical built-in repair analysis (BIRA) and built-in self repair (BISR) during production test.

DesignWare STAR Memory System eMRAM Trimming, Test, Repair And Diagnostics

Synopsys' DesignWare® STAR Memory System® (SMS) solution tests, repairs and diagnoses both on-chip memories (single/dual/two/multiport RAM/Register File/ROM including CPU and GPU cache, CAM, eflash) and off chip memories (DDR/LPDDR/HBM). By collaborating with leading foundries, Synopsys has augmented the SMS eMRAM to include algorithms specific to MRAM architectures with trimming/calibration capabilities. Synopsys also offers an ISO 26262 certified STAR ECC solution that can be leveraged to improve manufacturing yield of MRAM as well as improve in-field reliability for memories in application areas such as automotive and MilAero. The SMS eMRAM solution is silicon validated and offers the following capabilities:

Embedded Test and Repair

- At-speed test
- High test coverage with March algorithms with programmability via JTAG. Linear, ping-pong, diagonal, single addressing modes
- Support for eMRAM functional operations and test modes

Calibration

- On-chip programmability of trimming and timing parameters
- On-chip trimming, redundancy test and analysis fuse data generation and fuse programming

Diagnostics

- Provides comprehensive bit-fail information
- Test and repair status report via status register, additional test report (failing IO, Failing Bit Count) via diagnostics chains
- Unified software for ATE pattern generation and bit fail diagnostics and analysis

Advanced Modes

- Test with ECC bypassed and ECC enabled modes
- DFT specific: memory synchronous bypass, additional observability flops enhancing scan coverage
- Optional serial access (via JTAG) to memory static pins, test and power
- Automated SMS generation and verification flow for given memory configurations

Summary

STT MRAM enjoys scaling, power, memory density and cost benefits. The rapid enablement at five leading edge foundries and growing end user adoption supports a critical push from limited technology innovators to early mass market users. However, there are new test related challenges such as optimizing manufacturing yield for large array sizes, detecting different fault types specific to MRAM architectures, and the need to perform on-chip bias calculation of reference values. The memory BIST engine selected by the SoC designer needs to solve these challenges via combination of hard and soft repair capabilities, an extended class of test algorithms, and the inclusion of a programmable sequence of steps to perform calibration/trimming.

Synopsys' DesignWare SMS eMRAM provides a comprehensive solution for trimming, test, repair and diagnostics of STT-MRAM. It is a silicon-validated, low risk solution that is available now. This solution addresses the STT-MRAM test challenges via automation transparent to the end user and with minimal impact to the SoC design.