第 29 卷 第 5 期
2008 年 5 月

半 导 体 学 报
JOURNAL OF SEMICONDUCTORS

Vol. 29 No. 5
May, 2008

Optimization and Application of SRAM in 90nm CMOS Technology*

Zhou Qingjun† and Liu Hongxia

(*Key Laboratory of Ministry of Education for Wide Band-Gap Semiconductor Materials and Devices, School of Microelectronics, Xidian University, Xi'an 710071, China*)

Abstract: This paper presents an optimized SRAM that is repairable and dissipates less power. To improve the yield of SRAMs per wafer, redundancy logic and an E-FUSE box are added to the SRAM and an SR SRAM is set up. In order to reduce power dissipation, power on/off states and isolation logic are introduced into the SR SRAM and an LPSR SRAM is constructed. The optimized LPSR SRAM64K×32 is used in SoC and the testing method of the LPSR SRAM64K×32 is also discussed. The SoC design is successfully implemented in the Chartered 90nm CMOS process. The SoC chip occupies 5.6mm×5.6mm of die area and the power dissipation is 1997mW. The test results indicate that LPSR SRAM64K×32 obtains 17.301% power savings and the yield of the LPSR SRAM64K×32s per wafer is improved by 13.255%.

Key words: optimization; LPSR SRAM; redundancy logic; power on/off states
EEACC: 2570A; 2570D; 2575F
CLC number: TN402 **Document code**: A **Article ID**: 0253-4177(2008)05-0883-06

1 Introduction

Embedded memories constitute a significant portion of silicon area in today's integrated circuits. Both the number of embedded memories and their average size are increasing steadily. The International Technology Roadmap for Semiconductors (ITRS) indicates that current embedded memories occupy more than 70% of system-on-chip (SoC) area, and this number is likely to increase to 94% by the year 2014[1]. Power consumption of embedded memories plays a key role in the total power consumption of SoC and defects in memories can therefore significantly degrade the yield of embedded memory per wafer. In such a setting, low power and repairable embedded memories are desirable because they help reduce power dissipation of embedded memories and improve the yield of embedded memory per wafer.

Many redundancy mechanisms have been used to improve the yield of embedded memories. BISR (built-in self-repair) techniques with hierarchical redundancy architecture were proposed for word-oriented embedded memories. Without an electric fuse (E-FUSE) box to store addresses of faulty units, a memory built-in self test (MBIST) first detects faulty units in each power on state. The area overhead of the BISR circuit for a 2M-bit Static Random Access Memory (SRAM) is 2.56%[2]. The main objective of Ref. [3] is to provide a treatment of yield related techniques and a detailed assessment of the yield of compiler-based embedded SRAMs. According to the assessment in Ref. [3], the yield of a 2M-bit SRAM is improved by about 16% and the area overhead is improved by 1.3%. In this paper, an E-FUSE box is used for storing addresses of faulty units and redundancy logic is used for the replacement of faulty units. The E-FUSE box and the redundancy logic can be power off to reduce power consumption. Furthermore, for a 2M-bit SRAM, the testing results show that the yield gain is 13.255% and the area overhead is 0.983%.

Several methods of reducing power consumption for embedded memories have been proposed. Segmented virtual grounding (SVGND) is used to reduce power consumption[4]. The leakage current in memory cells is reduced by using a source-body bias not exceeding the value that guarantees safe data retention and low leaking nonminimum length transistors[5]. These methods are used for each memory cell and mission mode. In this paper, to reduce power consumption and facilitate tests, many operation modes are introduced into the SRAM and the power on/off states are set according to the requirements of each sub block of the SRAM. Moreover, redundancy mechanisms and low power techniques are combined into one embedded SRAM. The testing results prove that the Low Power Self Repair SRAM64K × 32 (LPSRM) obtains 17.301% power savings and the yield gain is 13.255%.

---

* Project supported by the National Natural Science Foundation of China (No. 60206006), the National Defense Pre-Research Foundation of China (No. 51308040103) and the Xi'an Applied Materials Innovation Fund (No. XA-AM-200701)

† Corresponding author. Email: zhqj76@126.com

Received 27 October 2007, revised manuscript received 8 January 2008

©2008 Chinese Institute of Electronics

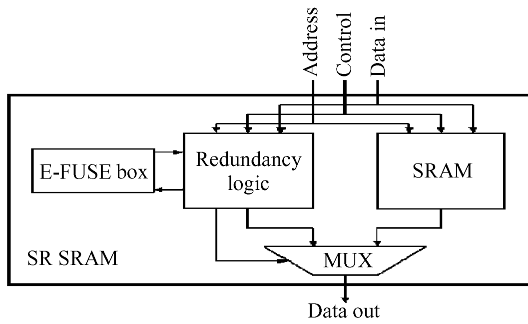

Fig.1 Connection of SRAM, redundancy logic, MUX, and E-FUSE box

The remainder of this paper is organized as follows. Section 2 presents the optimization of SRAM for improving the yield of SRAMs per wafer by adding redundancy logic and an E-FUSE box. Section 3 reports the low power optimization of Self Repair SRAM (SR SRAM) by introducing power on/off states and isolation logic. Section 4 describes the application of the LPSRM in SoC. Testing results are discussed in Section 5 and conclusions are drawn in Section 6.

2 Optimization of SRAM for improving the yield of SRAMs per wafer

2.1 SR SRAM block

The SR SRAM, shown in Fig. 1, is an assembly of one or more SRAMs, one E-FUSE box, one MUX, and the redundancy logic featuring real-time replacement of faulty memory units. Faulty memory units are detected during tests and their addresses are permanently stored into the E-FUSE box, which is composed of many E-FUSEs. The number of E-FUSEs depends on the memory addressing space and the number of redundant registers provided[6]. The redundancy logic, which includes redundant address registers (RARs) and redundant data registers (RDRs), may significantly improve the yield of SR SRAMs per wafer, but the redundancy logic and the E-FUSE box increase overall area and power consumption[7,8]. For any given SRAM, the number of optimum redundant registers is calculated according to the results of yield calculators. With the power on, the addresses of faulty units are loaded into RARs from the E-FUSE box and each address at the inputs of SR SRAM is compared with all addresses stored into the RARs. Whenever a match occurs, the redundancy logic automatically switches the data towards the RDR, which is logically associated with the address of the faulty unit.

The E-FUSE box is used for storing addresses of faulty units. When the addresses have been loaded into the RAR box, the E-FUSE box is power off to reduce power consumption by setting efc_isolate = “1”. When the addresses outside match the addresses stored into RAR box, the data are written into the RDR box and SRAM without using a MUX. The data are read from the RDR box by using a MUX, so the hardware is saved. In Ref. [2], without an E-FUSE box to store addresses, MBIST is first done to detect faulty units in each state with the power on. Two MUXes are used to read and write. The repair procedure is very complicated, so the area overhead is greater.

2.2 Redundancy logic diagram

Figure 2 shows the redundancy logic, which consists of five blocks: the RAR box, RDR box, address comparison block, RDR read block, and RDR write

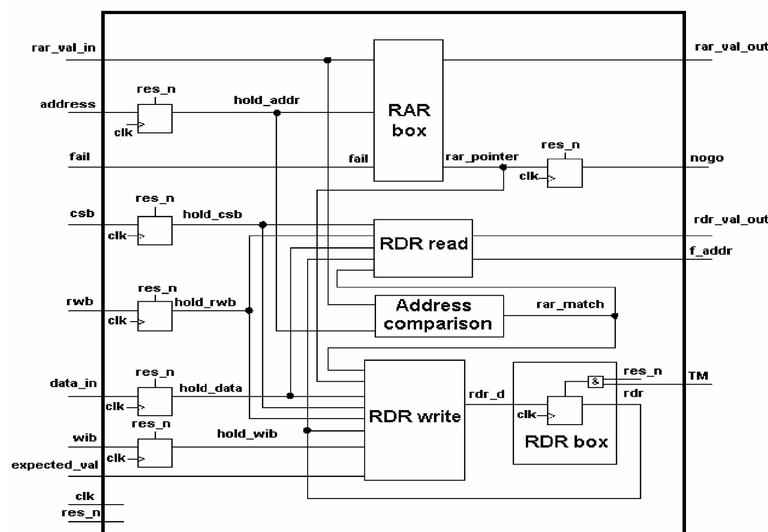

Fig.2 Redundancy logic

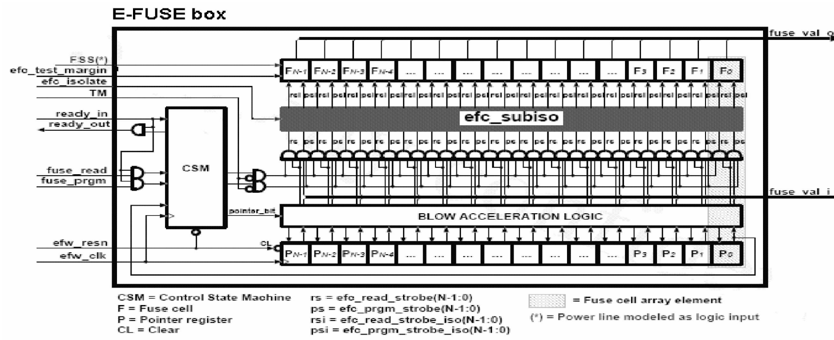

Fig.3　E-FUSE box

block. The RAR box is composed of all RARs, and the RDR box is composed of all RDRs. Each RAR can store one address of faulty units and each RDR can replace one faulty unit of SRAM. The fail signal coming from the MBIST controller will go high when a faulty unit of SRAM is detected. The addresses of faulty units are loaded into RARs through rar_val_in from the E-FUSE box and can be output through rar_val_out for analysis. Each address from outside is compared with all addresses in RARs. As soon as a match occurs, the signal rar_match will go high and data through expected_val will be written into RDR box and SRAM. But a read operation will be done from the RDR box[9]. As a result, the faulty unit is repaired. If there is no free RAR to store the residuary address of faulty units, the signal nogo will go high, indicating that the SRAM is not repairable.

2.3　E-FUSE box

The E-FUSE box consists of a bank of electrically programmable elements (E-FUSE cells) that allow permanent storage of RAR states through fuse_val_i. Programming/sensing sequences are performed by the local Control State Machine (CSM) connected with the pointer register P, as shown in Fig. 3. E-FUSEs need to be sensed in every power on state of the chip or the E-FUSE box and flip-flops in the RAR box are mapped one-to-one from E-FUSE cells. For example, RAR element k is associated with E-FUSE cell Fk. During fuse sensing, inputs clk and efw_clk must be driven by the same clock source. Signal ready_out going high for at least one efw_clk cycle indicates that fuse programming/sensing has finished. During fuse programming, a block of blow acceleration logic allows the skipping of all fuses that do not have to be programmed, thus reducing the programming time. After loading fuse states into the RARs, the E-FUSE box can be electrically switched off by setting efc_isolate = "1", which reduces the overall leakage current of the chip[10].

3　Low power optimization of SR SRAM

3.1　Power off mode

In previous SRAMs, each block is in either the power on state or power off state simultaneously, increasing the power consumption. In order to reduce power consumption and facilitate tests, we introduce many modes of operation and set each block in either state according to its actual requirements[11]. Before MBIST test mode, scan test mode must be done to test RARs, RDRs, E-FUSEs, and surrounding logic. MBIST mode can be separated into debug mode and repair mode. In debug mode, faulty units are detected but not repaired and their addresses are output for analysis. In repair mode, addresses of faulty units are loaded into RARs from the E-FUSE box, and the faulty units are repaired by RDRs unless there is no free RAR to store the residuary address. In mission mode, data is read from the memory or the redundancy logic depending on whether the match occurs between the current address and an address stored in the RARs. When LPSR SRAM is not set in test mode, the MBIST controller can be switched off. Table 1 shows the combinations of power on/off states in each operation mode.

3.2　Isolation logic

When a sub block of LPSR SRAM is turned off, the outputs float. In order to prevent the floating lines from creating paths between V_{dd} and V_{ss} at the inputs of sub blocks or units, isolation logic must be inserted

Table 1　Combinations of power on/off states

Mode of operation	Mbist controller	SRAM + RDR	RAR	E-FUSE	Surrounding logic
Scan test	On	On	On	On	On
MBIST debug test	On	On	On	On	On
MBIST repair test	On	On	On	Off	On
Mission	Off	On	On	Off	On
Power down	Off	Off	On	Off	Off

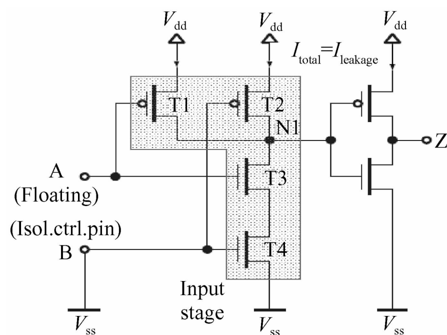

Fig.4 Isolation logic

to force these lines to be biased in the low state. Isolation logic is implemented by means of AND cells and their input circuit structure allows one or more inputs to be floated, while at least one input is forced into a low state[12,13]. In the 2-input AND circuit shown in Fig. 4, a low state at isolation control input B switches T4 off and T2 on. As a consequence, the voltage at node N1 that drives the inverter is tied to V_{dd} regardless of the states of T1 and T3, which forces output Z to a low state. In this case, no current flows through isolation gates except the total leakage current.

4 Application of LPSRM in SoC

An LPSRM composed of 8 LPSR SRAM8K×32s was used in SoC design. Figure 5 shows the layout of the floorplan. Because there are many modules in this project, a test control unit, which includes MBIST control logic and a finite state machine, implements all test modes[12,14]. In the SoC design, there are many test modes, such as MBIST test mode, scan test mode, AW test mode, and USB2 test mode.

The entry timing sequence of test modes is shown in Fig.6. When mbist_test_en is active (high) and the other test_ens are inactive (low), the chip will enter

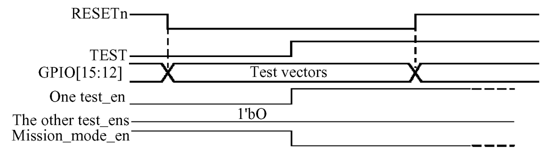

Fig.6 Entry timing sequence of test modes

MBIST test mode. After entering MBIST test mode, we can set mem_sel [4:0] to select the testing memory. In order to test the LPSRM in the SoC exhaustively, the redundancy logic and the E-FUSE box must be tested first and this test is performed by means of automatically generated scan patterns, which are shifted in and out through the redundancy logic in combination with the E-FUSE box[15]. Then the SRAM64K×32 (M) is scanned by a dedicated MBIST controller according to the MBIST algorithms. When a fault is detected in MBIST repair mode, the corresponding memory address is written into a free RAR by activating mbist_fail. As a result, the faulty unit is replaced by the RDR and the faulty unit is repaired. If no RAR is available to store the residuary address, mbist_nogo will go high, indicating that the memory is not repairable[6,9,16].

The SoC design has been successfully implemented in a Chartered 90nm CMOS process. The SoC chip occupies 5.6mm×5.6mm of die area and the power dissipation is 1997mW. The package type is LFBGA205. The metal layers are composed of 6 layers of Cu and 1 layer of Al. The summary of the SoC chip, including the LPSRM, is shown in Table 2.

5 Test results and discussion

All the 2061 SoC chips are tested in one wafer, the diameter of which is 300mm. Each SoC chip includes one LPSRM. Self repair logic, the E-FUSE

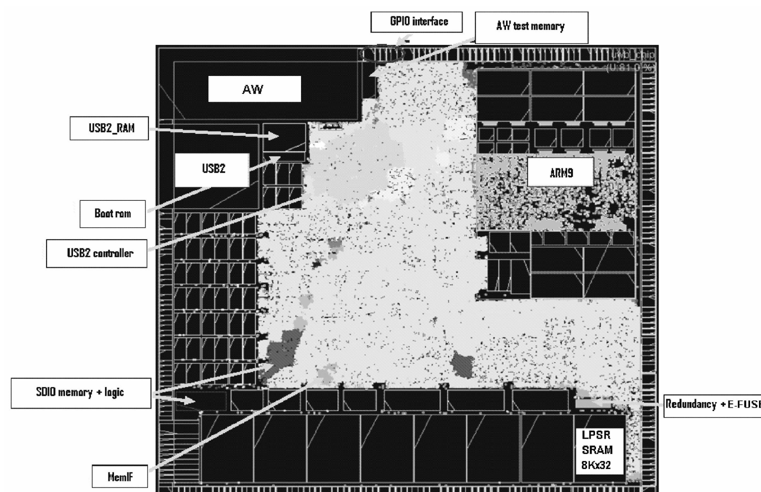

Fig.5 Layout of floorplan

Table 2　Summary of the SoC chip including the LPSRM

Technology	90nm CMOS
Package	LFBGA205
Metal layer	6 layers of Cu + 1 layer of Al
Polysilicon	SALICIDE
Supply of digital core	1.2V
Supply of IO	3.3V
Max frequency of core	132MHz
Pad number	205
Die size	5.6mm × 5.6mm = 31.36mm^2
Power consumption	1997mW

Table 3　Test results of M and SRM

	Area/mm^2	Power consumption/mW
M	3.661	30.229
SRM	3.697	30.704
Increment/%	0.983	1.571

box, and the redundancy logic can be bypassed by setting SR_bypass = "1". Also, low power logic can be bypassed by setting LP_bypass = "1". During the test pattern generation, if SR_bypass = "1" and LP_bypass = "1" are set, the pattern is used for testing M; if SR_bypass = "0" and LP_bypass = "1" are set, the pattern is used for testing SR SRAM64K × 32 (SRM); if SR_bypass = "0" and LP_bypass = "0" are set, the pattern is used for testing LPSRM. The total power consumption equals the sum of active power and leakage power[3,10].

Table 3 shows the test results for M and SRM. The area overhead of the redundancy logic together with the E-FUSE box is only 0.983%, which is less than the area overhead in Refs.[2,3]. The test results for SRM and LPSRM are listed in Table 4. The power savings of the LPSRM is [(30.704 − 25.392) / 30.704] × 100% = 17.301% and the area increment is [(3.704 − 3.697) / 3.697] × 100% = 0.19%.

Figure 7 presents the simulation results of the LPSRM in the SoC design with ModelSim. Figure 8 shows the testing results of the LPSRM in one SoC chip with tester J750. Figure 8 fits Fig. 7 very well, which proves that the real circuitry in the SoC chip is coincident with that in the SoC design. Figures 7 and 8

Table 4　Testing results of SRM and LPSRM

Memory	Voltage /V	Frequency /MHz	Area /mm^2	Power consumption /mW	Power saving /%
SRM	1.20	132	3.697	30.704	0
LPSRM	1.20	132	3.704	25.392	17.301

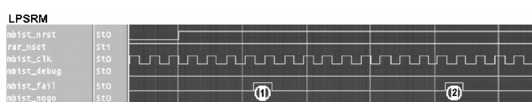

Fig.7　Simulation results of the LPSRM in the SoC design with ModelSim

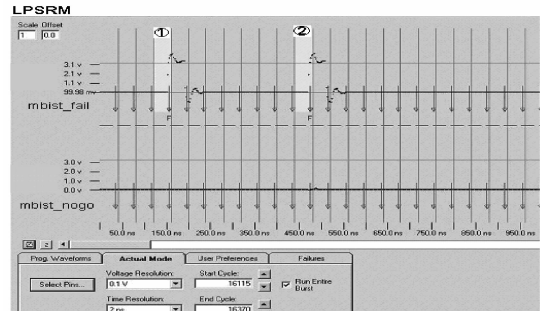

Fig.8　Test result of the LPSRM in one SoC chip with tester J750

Table 5　Testing summary of yield

MBIST test mode	LPSRMs in all	Good	Repaired	Yield gain/%
MBIST debug mode	2061	1441	0	0
MBIST repair mode	2061	1632	191	13.255

are both in MBIST repair mode. mbist_fail goes two times high and mbist_nogo is always low, which indicates that two faulty units in the M were repaired by the redundancy logic. In MBIST debug mode, faulty units are detected but not repaired.

The testing summary of the yield is shown in Table 5. After repair mode, 191 faulty Ms have been repaired so the yield gain (%) is (191/1441) × 100% = 13.255%. In Ref.[3], the yield calculated is improved by about 16%.

The defect density function is denoted as D. A uniform distribution function with magnitude $1/D_n$ between 0 and D_n for the defect density is assumed. The area of the chip without redundancy is denoted as A. The yield (Y_0) without redundancy is computed as:

$$Y_0 = \frac{1}{D_n}\int_0^{D_n} e^{-AD}\,dD \tag{1}$$

After adding redundancy, the yield (Y_1) is computed as:

$$Y_1 = Y_0 + (1 - Y_0)P_R \tag{2}$$

where P_R is the probability of memories repaired successfully by redundancy. The more redundancies that are added, the larger P_R becomes. So, the yield gain increases. On the other hand, the redundancy logic and the E-FUSE box increase the chip area, the number of SoC chips per wafer decreases, and the cost per chip increases. Tradeoffs must be made between yield and cost.

6　Conclusion

A new type of embedded SRAM is presented in this paper. It dissipates less power due to power on/off states and isolation logic, and improves the yield of LPSRMs per wafer by means of redundancy logic together with an E-FUSE box. Furthermore, the test

method of the LPSRM has been described. The test results show that the LPSRM obtains 17.301% power savings and the yield gain is 13.255%.

References

[1] Marinissen E J, Prince B, Keitel-Schulz D, et al. Challenges in embedded memory design and test. Proceedings of the Design, Automation and Test in Europe Conference and Exhibition, 2005: 722

[2] Lu S K, Yang C, Lin H W. Efficient BISR techniques for word-oriented embedded memories with hierarchical redundancy. Proceedings of the 5th IEEE/ACIS International Conference on Computer and Information Science, 2006: 355

[3] Wang Xiaopeng, Ottavi M, Meyer F J, et al. Estimating the manufacturing yield of compiler-based embedded SRAMs. IEEE Trans Semicond Manuf, 2005, 18(3): 412

[4] Sharifkhani M, Sachdev M. Segmented virtual ground architecture for low-power embedded SRAM. IEEE Trans Very Large Scale Integr (VLSI) Syst, 2007, 15(2): 196

[5] Cserveny S, Sumanen L, Masgonty J M, et al. Locally switched and limited source-body bias and other leakage reduction techniques for a low-power embedded SRAM. IEEE Trans Circuits Syst II, 2005, 52(10): 636

[6] Wang C W, Cheng K L. Fault pattern oriented defect diagnosis for memories. Proceedings International Test Conferences, 2003: 29

[7] Shoukourian S, Vardanian V, Zorian Y. SoC yield optimization via an embedded-memory test and repair infrastructure. IEEE Design & Test of Computers, 2004, 21(3): 200

[8] Hobson R F. A new single-ended SRAM cell with wire-assist. IEEE Trans Very Large Scale Integr (VLSI) Syst, 2007, 15(2): 174

[9] Abramovici M, Breuer M A, Friedman A D. Digital systems testing and testable design. Beijing: Tsinghua University Press, 2003

[10] Levacq D, Dessard V, Flandre D. Low leakage SOI CMOS static memory cell with ultra-low power diode. IEEE J Solid-State Circuits, 2007, 42(3): 690

[11] Roy K, Mukhopadhyay S, Mahmoodi-Meimand H. Leakage current mechanisms and leakage reduction techniques in deep-submicrometer CMOS circuits. Proc IEEE, 2003, 91(2): 305

[12] Yu Yang, Zhao Qian, Shao Zhibiao. Low power SRAM/SOI memory cell design. Chinese Journal of Semiconductors, 2006, 27(2): 318

[13] Min K S, Choi H D, Chio H Y, et al. Leakage-suppressed clock-gating circuit with zigzag super cut-off CMOS (ZSCCOMS) for leakage-dominant sub-70-nm and sub-1-V-VDD LSIs. IEEE Trans Very Large Scale Integr (VLSI) Syst, 2006, 14(4): 430

[14] Ruan Xiaoyu, Katti R S. Data-independent pattern run-length compression for testing embedded cores in SoCs. IEEE Trans Comput, 2007, 56(4): 545

[15] Zeng Gang. Concurrent core testing for SOC using merged test set and scan tree. IEICE Trans Inform Syst, 2006, E89-D(3): 1157

[16] Lu S K, Tsai Y C, Hsu C H. Efficient built-in redundancy analysis for embedded memories with 2-D redundancy. IEEE Trans Very Large Scale Integr (VLSI) Syst, 2006, 14(1): 34

90nm CMOS 工艺 SRAM 的优化及应用*

周清军† 刘红侠

(西安电子科技大学宽禁带半导体材料与器件国家重点实验室，西安 710071)

摘要： 提出了一种优化的 SRAM，它的功耗较低而且能够自我修复. 为了提高每个晶圆上的 SRAM 成品率，给 SRAM 增加冗余逻辑和 E-FUSE box 从而构成 SR SRAM. 为了降低功耗，将电源开启/关闭状态及隔离逻辑引入 SR SRAM 从而构成 LPSR SRAM. 将优化的 LPSR SRAM64K×32 应用到 SoC 中，并对 LPSR SRAM64K×32 的测试方法进行了讨论. 该 SoC 经 90nm CMOS 工艺成功流片，芯片面积为 5.6mm×5.6mm，功耗为 1997mW. 测试结果表明：LPSR SRAM64K×32 功耗降低了 17.301%，每个晶圆上的 LPSR SRAM64K×32 成品率提高了 13.255%.

关键词： 优化；低功耗自我修复 SRAM；冗余逻辑；电源开启/关闭状态

EEACC： 2570A；2570D；2575F

中图分类号： TN402 **文献标识码：** A **文章编号：** 0253-4177(2008)05-0883-06

* 国家自然科学基金(批准号：60206006)，国防预研基金(批准号：51308040103)以及西安应用材料创新基金(批准号：XA-AM-200701)资助项目

† 通信作者. Email：zhqj76@126.com

2007-10-27 收到，2008-01-08 定稿

©2008 中国电子学会