

Fabrication of a magnetic tunnel junction-based 240-tile nonvolatile field-programmable gate array chip skipping wasted write operations for greedy power-reduced logic applications

Daisuke Suzuki^{1a)}, Masanori Natsui^{1,2}, Akira Mochizuki¹,
Sadahiko Miura³, Hiroaki Honjo³, Keizo Kinoshita¹,
Hideo Sato¹, Shoji Ikeda^{1,2,5}, Tetsuo Endoh^{1,4,5}, Hideo Ohno^{1,2,5,6},
and Takahiro Hanyu^{1,2,5}

¹ Center for Spintronics Integrated Systems, Tohoku University,
2-1-1, Katahira, Aoba-ku, Sendai 980-8577, Japan

² Research Institute of Electrical Communication, Tohoku University,
2-1-1, Katahira, Aoba-ku, Sendai 980-8577, Japan

³ Green Platform Research Laboratories, NEC Corporation,
34, Miyukigaoka, Tsukuba 305-8501, Japan

⁴ Graduate School of Engineering, Tohoku University,
6-6, Aramaki Aza Aoba-ku, Sendai 980-8579, Japan

⁵ Center for Innovative Integrated Electronic Systems, Tohoku University,
468-1, Aramaki Aza Aoba-ku, Sendai, Miyagi 980-0845, Japan

⁶ WPI-Advanced Institute for Materials Research, Tohoku University,
2-1-1, Katahira, Aoba-ku, Sendai 980-8577, Japan

a) show-you@ngc.riec.tohoku.ac.jp

Abstract: A nonvolatile field-programmable gate array (NVFPGA) test chip with 240 tiles (the basic components) in a 12×20 2D-array is fabricated by 90 nm CMOS and 70 nm magnetic tunnel junction (MTJ) technologies. Since not only circuit configuration data but also temporal data are still remained in the MTJ devices even when the power supply is cut off, standby power dissipation is completely eliminated by utilizing tile-level power gating. Power reduction is further accelerated by skipping wasted write operations of nonvolatile flip-flops (NVFFs) for storing temporal data when the temporal data and the stored one are the same. As a typical application, a motion-vector prediction function is implemented on the proposed NVFPGA, which results in a write power reduction of 77% compared to that of a conventional MTJ-based NVFPGA and a total power reduction of 70% compared to that of an

SRAM-based FPGA.

Keywords: field-programmable gate array, magnetic tunnel junction device, nonvolatile logic-in-memory architecture, power-gating

Classification: Electron devices, circuits, and systems

References

- [1] I. Kuon and J. Rose: IEEE Trans. Very Large Scale Integr. (VLSI) Syst. **19** [1] (2011) 71.
- [2] E. Pop: Nano Res. **3** (2010) 147.
- [3] J. Greene, S. Kaptanoglu, W. Feng, V. Hecht, J. Landry, F. Li, A. Krouglyanskiy, M. Morosan and V. Pevzner: Proc. ACM Int. Symp. FPGA (2011) 87.
- [4] M. Iida, M. Koga, K. Inoue, M. Amagasaki, Y. Ichida, M. Saji, J. Iida and T. Sueyoshi: IEICE Trans. Electron. **E94-C** (2011) 548.
- [5] Y. Y. Liauw, Z. Zhang, W. Kim, A. E. Gamal and S. S. Wong: ISSCC Dig. Tech. Papers (2012) 406.
- [6] M. Miyamura, M. Tada, T. Sakamoto, N. Banno, K. Okamoto, N. Iguchi and H. Hada: IEDM Tech. Dig. (2012) 247.
- [7] S. Ikeda, K. Miura, H. Yamamoto, K. Mizunuma, H. D. Gan, M. Endo, S. Kanai, J. Hayakawa, F. Matsukura and H. Ohno: Nature Materials **9** (2010) 721.
- [8] S. Yamamoto, Y. Shuto and S. Sugahara: Jpn. J. Appl. Phys. **51** (2012) 11PB02.
- [9] D. Suzuki, M. Natsui, S. Ikeda, H. Hasegawa, K. Miura, J. Hayakawa, T. Endoh, H. Ohno and T. Hanyu: Symp. VLSI Circuits Dig. Tech. Papers (2009) 80.
- [10] D. Suzuki, M. Natsui, T. Endoh, H. Ohno and T. Hanyu: Jpn. J. Appl. Phys. **51** (2012) 04DM02.
- [11] T. Tuan, A. Rahman, S. Das, S. Trimberger and S. Kao: IEEE Trans. Computer-Aided Design Integr. Circuits Syst. **26** (2007) 296.
- [12] M. Natsui, D. Suzuki, N. Sakimura, R. Nebashi, Y. Tsuji, A. Morioka, T. Sugibayashi, S. Miura, H. Honjo, K. Kinoshita, S. Ikeda, T. Endoh, H. Ohno and T. Hanyu: ISSCC Dig. Tech. Papers (2013) 194.
- [13] M. E. Sinangil and A. P. Chandrakasan: ISSCC Dig. Tech. Papers (2013) 318.

1 Introduction

The SRAM-based field-programmable gate array (FPGA) has become one of the key digital circuit implementation media owing to nano-scale CMOS technology [1]. However, standby power dissipation due to leakage current has become a critical issue for such nano-scale FPGAs [2]. A nonvolatile FPGA (NVFPGA), where both circuit configuration data and temporal data of flip-flops (FFs) are retained in nonvolatile devices, is a promising way of reducing standby power. Since all the data are retained in the nonvolatile devices even if the power supply is cut off, standby power dissipation can be eliminated by power gating where the basic building blocks (tiles) in the idle state are temporarily turned off.

To realize such an NVFPGA, it is very important that the nonvolatile storage function is implemented with no area overhead. It is also important that the nonvolatile device has unlimited endurance since the temporal data of FFs must be stored in nonvolatile devices every time before power-off. Recently, some NVFPGAs have been presented using nonvolatile devices such as floating-gate MOS (FGMOS) transistors [3], ferroelectric capacitors [4], resistive RAM (ReRAM) devices [5], atom switches [6], and magnetic tunnel junction (MTJ) devices [8, 9, 10]. FGMOS transistors and ferroelectric capacitors are well-known nonvolatile devices, but their limited endurance and low compatibility with logic gates due to the complex process flow restricts their applications. ReRAM devices and atom switches are suitable for storing circuit configuration data with small chip area since they are fabricated on the metal layer by back-end of line (BEOL) processing, but their endurance is also limited.

In this paper, we propose an NVFPGA using MTJ devices. Since MTJ devices have unlimited endurance [7] and 3D-stacking capability with BEOL processing, they are suitable for realizing ultra low-power NVFPGA with power gating. To fully utilize such features of MTJ devices, a nonvolatile lookup table (NVLUT) circuit, which is one of the fundamental logic components of the NVFPGA, is implemented using a logic-in-memory (LIM) structure where logic and nonvolatile storage functions are compactly merged [9, 10]. Since any arbitrary logic functions can be performed by using small difference of current signal through MTJ devices [10], the LIM-based NVLUT circuit is implemented using only one sense amplifier, which results in reduction of leakage current path. To realize the MTJ-based NVFPGA, it is also important to reduce power dissipation due to current-induced magnetization switching for storing temporal data in an MTJ-based nonvolatile FF (NVFF). When the temporal data is the same as the stored data in the MTJ device or the output of the NVFF is not used for operation, the write operation is wasted. Thus, the write power can be reduced by detecting these two cases and skipping write operations.

A test chip that includes 240 tiles in a 12×20 2D array is fabricated using 90 nm CMOS/70 nm MTJ technologies and the nonvolatile capability of the proposed NVFPGA is demonstrated. As a typical application, a motion-vector prediction function is mapped to the proposed NVFPGA. By skipping all the wasted write operations, the write power for storing temporal data of NVFFs is reduced by 77% compared to that of a conventional MTJ-based NVFPGA, and the total power is reduced by 70% compared to that of an SRAM-based one.

2 MTJ-based nonvolatile FPGA architecture

Fig. 1 (a) shows a symbol of an MTJ device. Since binary information (M) is programmed as a resistance value (R_M) by the current-induced magnetization switching, it can be regarded as a variable resistor. Fig. 1 (b) shows the measured $R - I$ curve of the fabricated MTJ device. I_{W0} and I_{W1} are

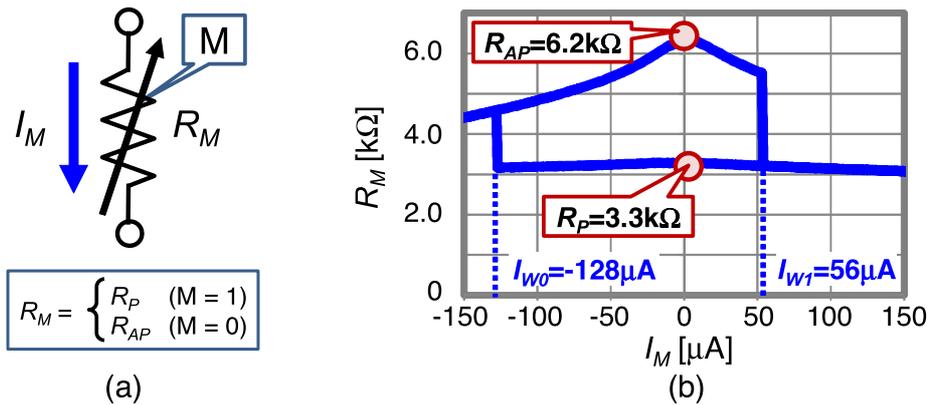


Fig. 1. MTJ device: (a) Symbol. (b) Measured $R - I$ curve of a fabricated MTJ device.

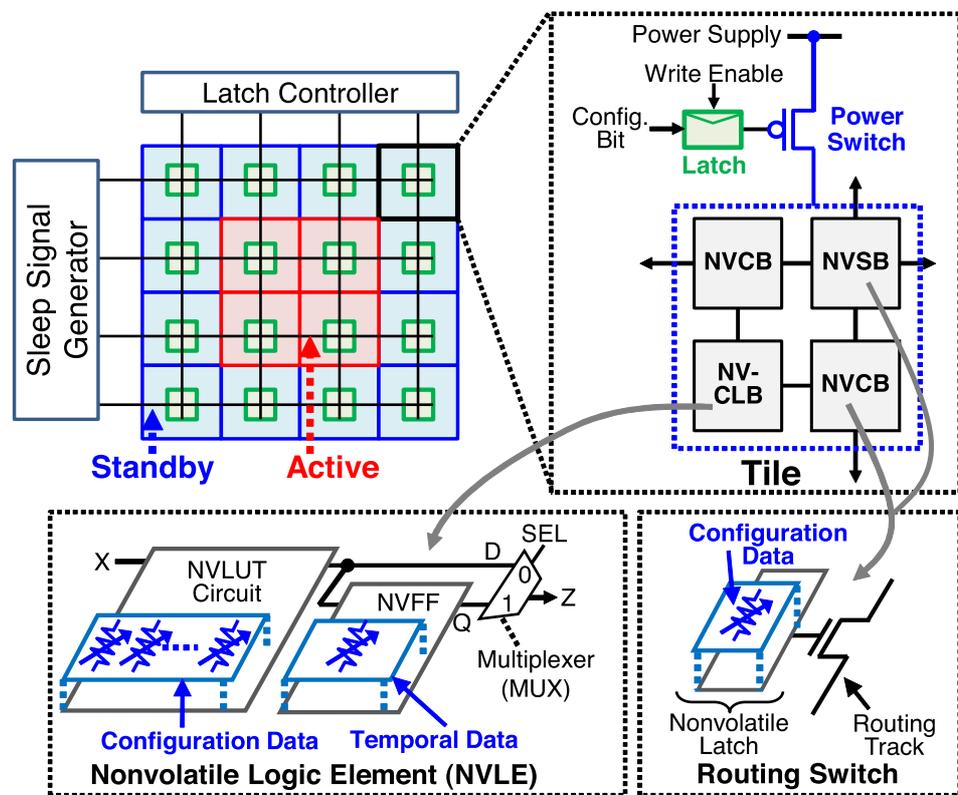


Fig. 2. Structure of the proposed NVFPGA. Since all the data are remained in each tile, tile-level power gating is easily performed by adding power switches together with latches for storing power-on/off state.

the currents to program $M = 0$ and $M = 1$, respectively.

Fig. 2 shows the structure of the proposed NVFPGA where MTJ devices are distributed over the 2D array of tiles. Since all the data are remained in each tile, the power supply can be turned on/off individually. A tile comprises of a nonvolatile configurable logic block (NVCLB) with a cluster of nonvolatile logic elements (NVLEs) [10], two nonvolatile connection blocks (NVCBs) for interfacing the NVCLB to routing tracks, and a non-

volatile switch block (NVSB) for signal routing between the routing tracks. An NVLE consists of a nonvolatile look-up table (NVLUT) circuit and a non-volatile flip-flop (NVFF) to create a structure capable of implementing either combinational logic or sequential logic. The routing function is determined by the configuration of the nonvolatile latch in each routing switch.

Meanwhile, one of the key power-gating architecture design decisions is the granularity of the smallest block that can be independently power gated since there is a trade-off between area overhead and standby power dissipation [11]. In the proposed NVFPGA, tile-level power gating [11] is applied, and the power-on/off state of each tile is controlled by its configuration bit. Therefore, the area overhead in each tile is caused by just a power switch together with a latch. Since these latches are distributed over the array of tiles, they are configured by a latch controller and a sleep signal generator; thus, write enables are applied by the latch controller, and sleep-control signals are applied by the sleep signal controller. In this way, the tile-level power gating is implemented.

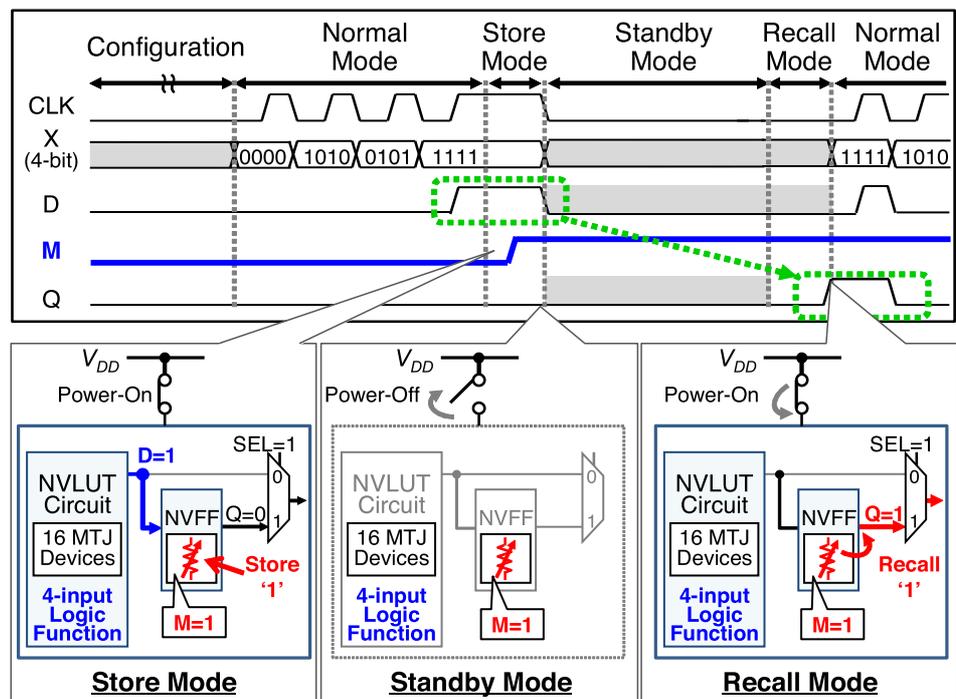


Fig. 3. Timing diagram of an NVLE in the proposed NVFPGA. Since all the data in the NVFPGA are remained in MTJ devices even if power supply is cut off, no external nonvolatile memory access is required.

Fig. 3 shows a timing diagram of the 4-input NVLE where a 4-input logic function is programmed in 16 MTJ devices in the NVLUT circuit and the output of the NVFF is selected (*SEL* is stored in a nonvolatile latch). Once the circuit configuration data is programmed in the MTJ devices, it remains even if the power supply is cut off, and it does not change unless another

configuration is performed. In the store mode, $D = 1$ is stored in an MTJ device in the NVFF; thus, it is also remained during the power-off state. Just after power-on, $M = 1$ is immediately recalled into the NVFF and Q becomes '1.' In this way, power supply can be turned on/off with no external nonvolatile memory access.

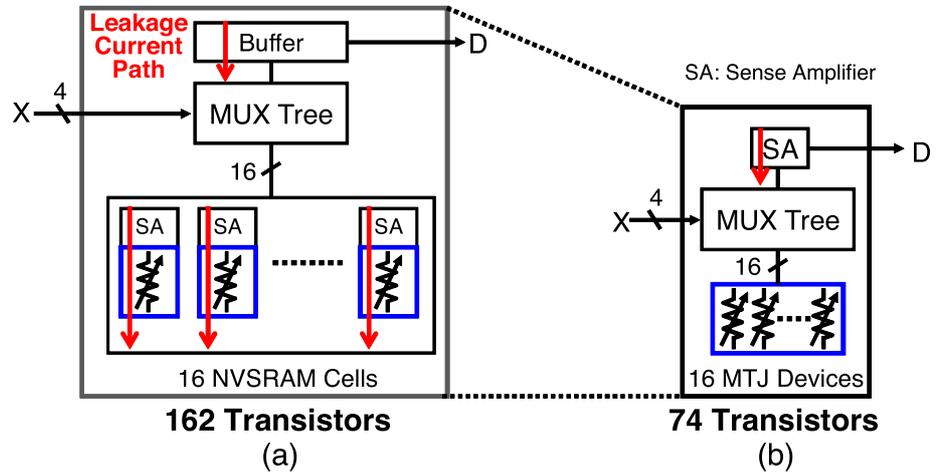


Fig. 4. Block diagrams of 4-input NVLUT circuits using MTJ devices: (a) CMOS-based (conventional) [8]. All the SRAM cells are replaced by nonvolatile ones. Each NVSRAM cell has a sense amplifier to read data from an MTJ device. (b) LIM-based (proposed) [10]. Since just only one sense amplifier is shared among all the MTJ devices in the NVLUT circuit, both transistor counts and number of leakage current paths are greatly reduced compared to those of the CMOS-based one.

Figure 4 shows block diagrams of 4-input NVLUT circuits with MTJ devices. A simple way to implement an NVLUT circuit is a CMOS-based implementation where SRAM cells are replaced by nonvolatile ones [8], as shown in Fig. 4 (a), where transistor counts are large since the sense amplifier together with read/write control transistors are indispensable for each configuration bit to read data from the MTJ device. In the LIM-based NVLUT circuit, in contrast, the current signal through the MUX tree and the corresponding MTJ device is directly evaluated as the result of logic operation. Therefore, any arbitrary 4-input logic function can be performed just a sense amplifier as shown in Fig. 4 (b) [10]. As a result, the number of transistors as well as the number of leakage current path are smaller than those of the CMOS-based one.

To realize an ultra low-power FPGA, it is also important to reduce power dissipation for storing temporal data of each NVFF. The key idea for the power reduction is to skip wasted write operations. Let us consider two cases as shown in Fig. 5 (a). In case I, Q is not used for the logic operation since $SEL = 0$ and the NVLE performs a combinational logic operation. In

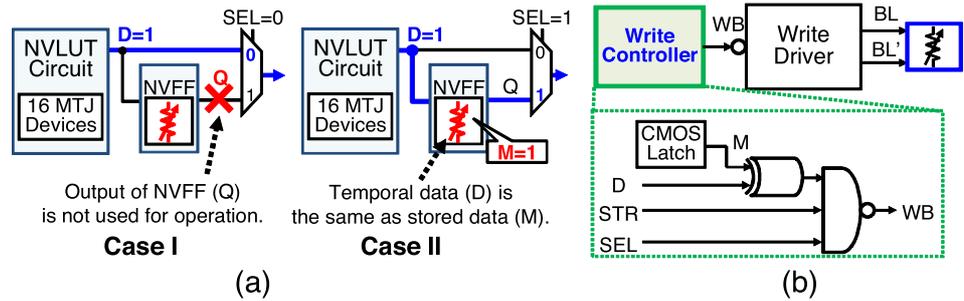


Fig. 5. Proposed method for skipping wasted write operations: (a) Two cases of wasted write operations: (a) Two cases of wasted write operation. In case I, Q is not used for logic operation. In case II, M does not change since $M = D = 1$. (b) Block diagram. The write driver is activated only when $(STR, SEL) = (1, 1)$ and $D \neq M$.

case II, M does not change since $M = D = 1$. Therefore, the wasted write operations can be skipped by detecting these two cases. Fig. 5 (b) shows the block diagram of the proposed write controller with a write driver for an NVFF. The write driver applies a bi-directional write current to the MTJ device when $WB = 0$. The initial state of M is stored into a CMOS latch just after power-on and it is kept during normal mode. In the store mode, STR becomes '1' and WB becomes '0' only when $SEL = 1$ and $D \neq M$. In this way, all the wasted MTJ write operations are completely skipped.

3 Evaluation

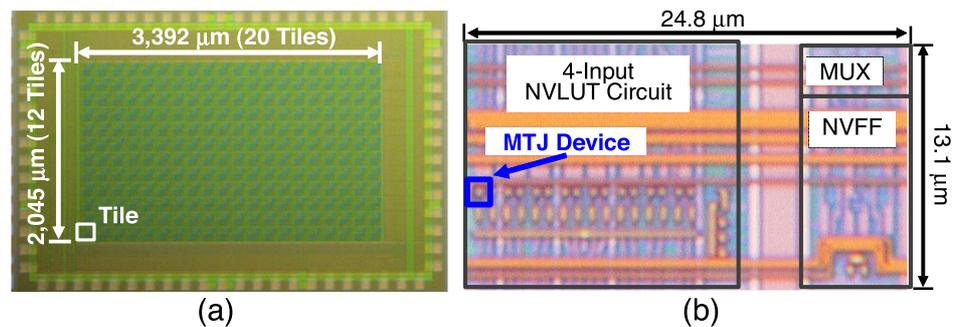


Fig. 6. Die photo of the fabricated test chip using 90 nm CMOS/70 nm MTJ technologies : (a) Overall structure. (b) 4-input NVLE.

We have fabricated a test chip which has 240 tiles in a 12×20 2D-array using 90 nm CMOS/70 nm MTJ technologies with five metal layers as shown in Fig. 6 (a). The MTJ devices are stacked over the 4th metal layer. The tile includes four routing tracks and four 4-input NVLEs whose effective area is $325 \mu\text{m}^2$ as shown in Fig. 6 (b).

Fig. 7 shows the measured waveforms of the fabricated test chip. As shown in Fig. 7 (a), both combinational logic and sequential logic operations

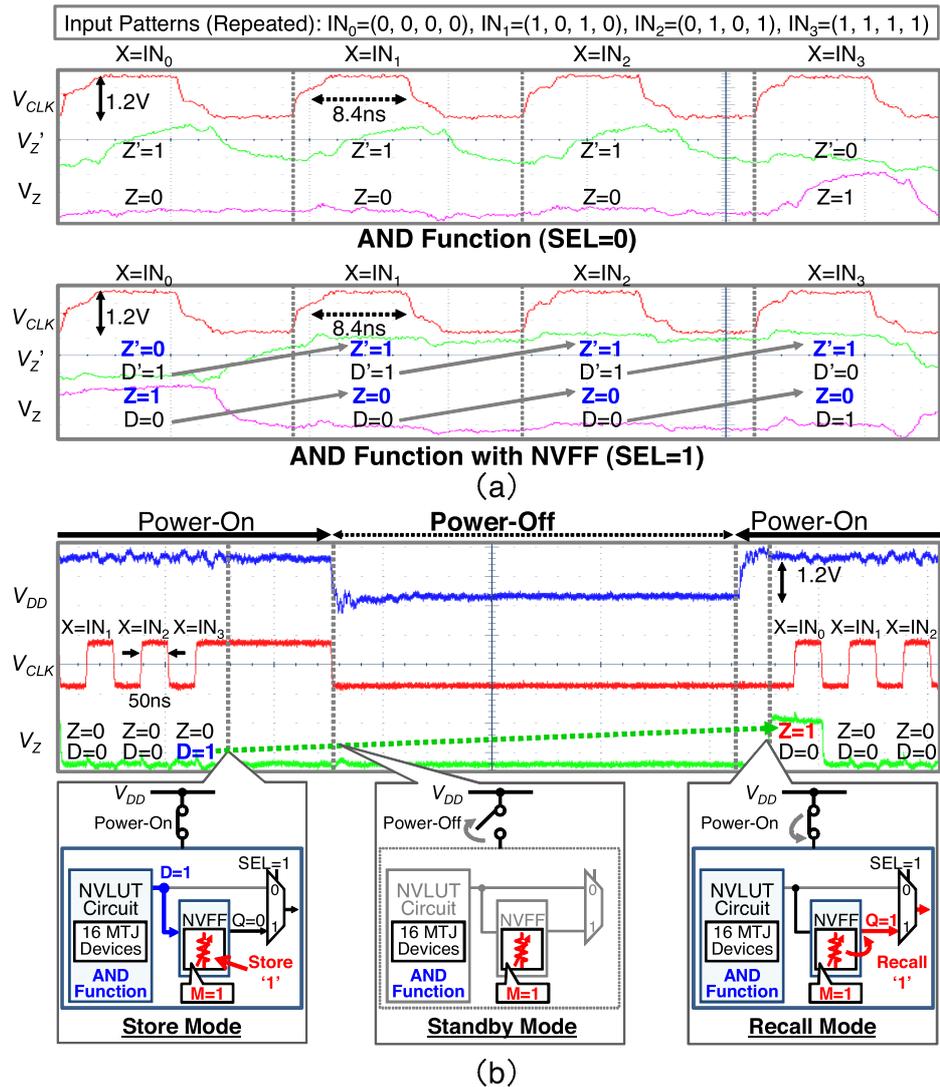


Fig. 7. Measured waveforms of the fabricated test chip: (a) Basic logic operations. Both combinational logic and sequential logic operations are confirmed by changing SEL . (b) Power gating behavior. $D = 1$ is stored in the MTJ device as $M = 1$ before power-off, and it is recalled to the NVFF after power-on.

are confirmed by changing SEL . In Fig. 7 (b), we can also confirm that $D = 1$ is stored in the MTJ device as $M = 1$ before power-off and it is recalled to the NVFF after power-on. As a result, the power-gating behavior without losing both circuit configuration data and temporal data is successfully confirmed by the test chip.

To evaluate the advantage of the proposed NVFPGA, a motion-vector prediction function, where a large number of adders and registers are used, is implemented as shown in Fig. 8. It consumes a large amount of standby power during sum of absolute differences (SAD) calculation due to its massively parallel structure [12]. The proposed NVFPGA is suitable for such an application since both the circuit configuration data of the PE and reference

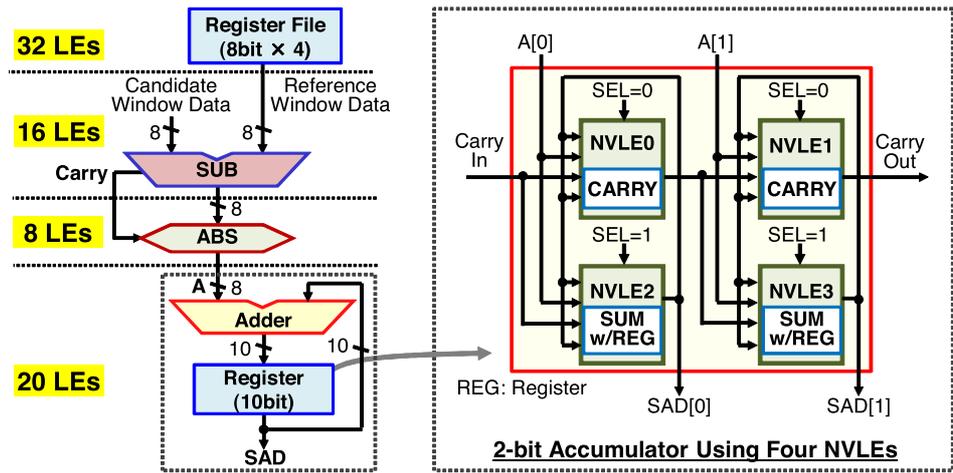


Fig. 8. Mapping of a motion-vector prediction function to the proposed NVFPGA whose search area size is 4×4 pixels and reference window size is 2×2 pixels, respectively. A full adder is implemented using two 4-input NVLEs.

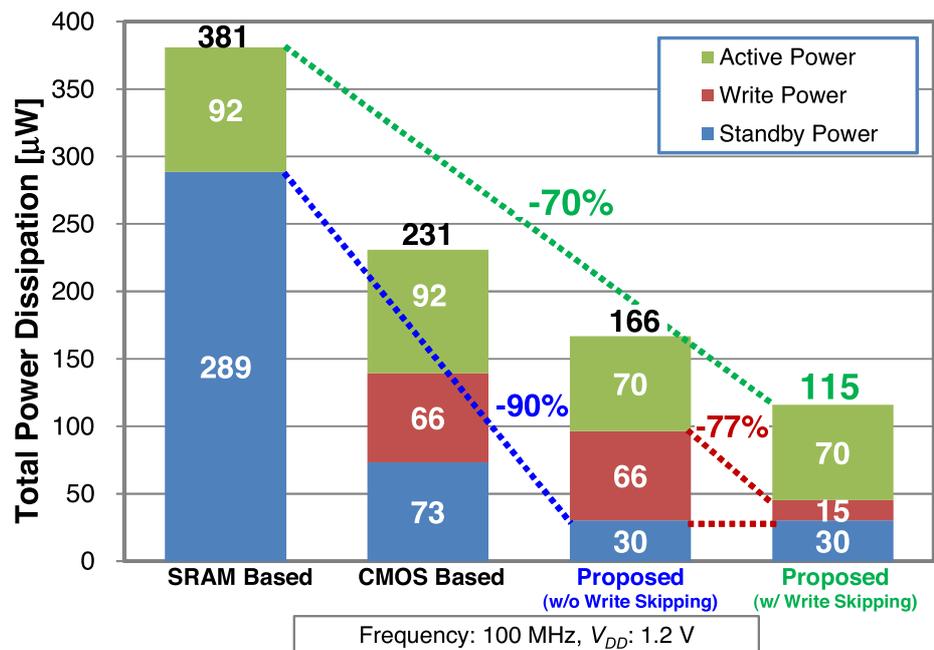


Fig. 9. Comparison of total power dissipation of motion-vector prediction function. Search area size and reference windows size are 16×16 pixels and 8×8 , respectively. Total power is reduced by 70% compared to that of the SRAM-based FPGA.

window data are remained once they are stored in MTJ devices even if power supply is cut off.

Fig. 9 summarizes a comparison of power dissipation of the motion-vector prediction function; an SRAM-based FPGA, a CMOS-based NVFPGA, the proposed NVFPGA without write skipping, and the proposed NVFPGA with write skipping. In the CMOS-based NVFPGA, standby power dissipation is

reduced by using power gating while standby power is still large since each NVSRAM cell has a leakage current path. In contrast, standby power dissipation of the proposed NVFPGA is greatly reduced since NVSRAM cells in the NVLUT circuits are replaced by MTJ devices and the number of leakage current paths is greatly reduced, which results in a standby power reduction of 90% compared to that of the SRAM-based FPGA. In Fig. 9, we can also see that the MTJ write power dissipation accounts for 41% of the total power dissipation in the proposed NVFPGA without write skipping. Meanwhile, it is reported that image frames generally consist of smooth backgrounds or large objects where the intensity of pixels is spatially correlated and about 75% of pixels are the same [13]. Therefore, we assume that the probability of $D \neq M$ is 25%, which results in an MTJ write power reduction of 77% with write skipping. The area penalty of the 4-input NVLE due to the additional circuit for the write skipping is about 10%, which becomes smaller as the number of inputs since the area of the NVLUT circuit is dominant in a multi-input NVLE (for example, area penalty becomes 5% when the number of inputs is six). As a result, the total power of the proposed NVFPGA is reduced by 70% compared to that of the SRAM-based FPGA.

4 Conclusion

An NVFPGA chip has been presented using 90 nm CMOS/70 nm MTJ technologies. Since tile-level power gating is utilized in the proposed NVFPGA, standby power is completely eliminated. Moreover, the write power for storing temporal data in NVFFs before power-off is also eliminated by skipping wasted MTJ write operations. As a demonstration of the proposed FPGA, a motion-vector prediction function is implemented and the total power reduction of 70% compared to that of an SRAM-based FPGA and the write power reduction of 77% compared to an NVFPGA without write skipping. As a next target, it is important to develop a CAD environment which makes it possible to perform logic mapping and physical place-and-route for minimizing the total power dissipation with optimum tile-level power gating.

Acknowledgments

This research is granted by the Japan Society for the Promotion of Science (JSPS) through the “Funding Program for World-Leading Innovative R&D on Science and Technology (FIRST Program)”, initiated by the Council for Science and Technology Policy (CSTP). A part of this work was conducted at TIA-SCR. This research is also supported by VLSI Design and Education Center (VDEC), the University of Tokyo in collaboration with Cadence, Inc. The authors thank N. Sakimura and R. Nebashi, Y. Takako, and A. Tamakoshi.