Emerging Memory Technologies for Reconfigurable Routing in FPGA Architecture

Pierre-Emmanuel Gaillardon, M. Haykel Ben-Jamra, Giovanni Betti Beneventi, Fabien Clermidy, Luca Perniola
CEA, LETI, Minatec, Grenoble, France
pierre-emmanuel.gaillardon@cea.fr

Abstract—Emerging non-volatile resistive memories such as Phase Change Memories (PCMs) are promising candidates to replace Flash and SRAM memories in some applications. This paper introduces a novel memory node for Field-Programmable Gate Arrays (FPGAs). We propose an elementary circuit storing the reconfiguration signals by means of two resistive memories and one programming transistor. The area and write time of the proposed node are investigated and their impact on complex circuits is assessed. We show that the elementary memory node yields an improvement in area and write time of 3.4x and 16x respectively vs. a regular Flash implementation. We use the designed memory node to reconfigure the switchbox in the routing part of FPGAs, and we demonstrate a delay reduction up to 50% on a benchmark of logic circuits that we mapped on an FPGA architecture thanks to the area saving and especially to the low on-resistance of PCMs.

I. INTRODUCTION

Resistive memories are a family of two terminal devices that can store the information in an internal state, which depends on the underlying technology. One of those resistive memories are Phase Change Memories (PCMs) which are considered today to be one of the most promising candidates for next generation of non-volatile memory applications [1]. The interest in PCMs is due to different advantages, including: better scalability (up to few nanometers) [2], faster programming time (in the order of few nanoseconds) [3] and an ameliorated endurance (up to $10^9$ programming cycles) [4]. Some prototypes have been presented recently in order to showcase the viability of high density standalone memories based on PCM technology from the industrial point of view. As a matter of fact, a 60-nm 512-Mb [5] and a 45-nm 1-Gb [4] PCM technology have been recently demonstrated.

The focus in this work is on the utilization of PCMs in reconfigurable logic circuits, such as Field-Programmable Gate Arrays (FPGAs). The reason behind this choice is the fact that in reconfigurable logic, up to 40% of the area is dedicated to the storage of configuration signals, leading to a large cost of reconfiguration in terms of area and routing delay [6]. Traditionally, the configuration is serially loaded in SRAM cells distributed throughout the circuits [6]. As a consequence, the configuration at power-up is a time-consuming operation. Because of SRAM volatility and loading time, non-volatile memories can be used to address this issue. In [7], a complete solution based on Flash memories, able to store the configuration when the power is off, is presented. The drawback of this solution is that it mixes CMOS with Flash technology, resulting in a high cost in terms of technology. On the other hand, PCMs are fabricated in a back-end-of-line process, i.e., like metal layers, allowing us to move all the configuration memory onto the top of the chip, inducing a promising reduction of area in complex reconfigurable circuits.

In this paper, we propose an elementary crosspoint of a FPGA switchbox, formed by one single resistive memory. This node functionally stores the information in its resistive state and can be used to route signals through low-resistive paths, or isolate them by means of high-resistive paths. Both the performances of the node and its impact on FPGA circuits are studied versus the equivalent SRAM and Flash implementations. We show that the PCM switchbox saves the area by a factor of 1.5x and the write time by a factor of 16x compared to Flash. The compact dimensions of the PCM-based device reduce the size of FPGA blocks and routing channels. This yields a delay reduction up to 50% in the best case of complex benchmark circuits.

The organization of the paper is the following. Section II surveys the FPGA architecture. Section III gives an overview of PCM technology. Then, in section IV, we evaluate the single structure with respect to competing technologies and we study the impact of our approach on FPGA design. In section V, we discuss potential opportunities given by the evolution of the technology. Finally, in section VI, we draw the conclusions.

II. ARCHITECTURAL BACKGROUND AND MOTIVATION

A. FPGA architecture

FPGAs are regular circuits formed by several identical reconfigurable logic blocks called Configurable Logic Blocks (CLBs) that are surrounded by reconfigurable interconnect lines [6]. Every CLB is formed by a set of N Basic Logic Elements (BLEs). A BLE is simply a K-input Look-Up-Table (LUT), whose output can be routed to any other LUT input with or without being saved in a flip-flop. Every CLB has I inputs coming from other CLB outputs and from external signals to the CLBs. All design parameters N, K and I can be set by the FPGA architect depending on the targeted system granularity. The overall delay and area strongly depend on those parameters, which have extensively been investigated in [8].

The routing part of the FPGA is formed by large channels interleaved between the CLBs and having the width W. The channels cross each another and the signals can be routed within the FPGA using switchboxes. A switchbox is a matrix at

---

This work was funded by the French National Research Agency under the program ANR-08-SEGI-012 "NANOGRAIN".
the crossing of channels, which is made of reconfigurable switches.

B. Programmable routing element

In today’s FPGAs, there are two technologies used for reconfigurable switches. Most of the FPGAs are made in a full CMOS process. Then, the switches are transistors or gated buffers which are controlled by a reconfiguration signal stored in a SRAM cell. However, for some specific applications – especially those requiring a higher reliability than the one of SRAM cells – flash technology is integrated into the CMOS process and the switches are made of flash transistors that can be either passing or blocking depending on their configuration. The drawback of the SRAM-based solution is the volatility. While the Flash-based solution addresses the issue of volatility, it is technologically expensive because of the undesirable integration of flash devices into a CMOS circuit.

The main idea in this work is to replace the switching elements that traditionally include MOS transistor by switching elements with a lower resistivity formed by PCMs. Considering that NMOS transistors have a resistance of around 9 kΩ (CMOS 65nm), a phase-change memory element seems to be an attractive solution with a lower on-resistance that has been reported in the range of 50 Ω to 4 kΩ, depending on technology.

III. NON-VOLATILE RESISTIVE MEMORY FOR ROUTING APPLICATION

A. Physics principle

A PCM device relies in the unique properties of chalcogenide alloys as active materials integrated in the memory cell stack. Chalcogenide alloys are semiconducting glasses made by elements of the VI group of the periodic table, such as sulphur, selenium and tellurium, showing reversible phase-change transformation capability. Indeed, by means of a careful control of Joule heating through the cell, it is possible to electrically switch the chalcogenide layer between two stable configurations, i.e., a high-conductive polycrystalline state and a low-conductive amorphous one, as it is highlight in figure 1. A sufficiently high voltage pulses heat the PC layer above the melting temperature of the material (T_m). A rapid quench follows and a part the chalcogenide alloys (depicted as an oval in the PC layer) is stuck in the amorphous phase. The resulting memory cell is in a high resistance state (Figure 1-a). A lower but longer pulse is used to crystallize the amorphous region of the PC layer in order to achieve a low resistance memory cell (Figure 1-b). Phase transformation is fully reversible up to 10⁶ cycles. It is worth noting that even if most of the research work has focused on the GeSbTe alloy (also called GST) [2], other materials, such as GeTe or GeTeCα% (with α representing the percentage of C in GeTeC) have been presented [10]. They have different properties compared to GST. For example, a better high conducting state or better data retention properties have been demonstrated [11].
The required programming voltages and timing pulses to program the PCMs may be applied through the drivers at the inputs and the outputs of the switchboxes (figure 3-a). Figure 4 shows a possible structure of these drivers. They are used for electrical interface between signal channels and the programming unit, which generates the configuration waveforms.

![Figure 3. PCM-based 2x2 switchbox architecture](image)

**D. Programming scheme**

A PCM is programmed by applying a pulsed signal between its two terminals. This conducts the PCMs of the switchbox to be addressed sequentially. A selection of a node is done by selecting the drivers related to its terminals, while other PCM’s terminal are left floating, to avoid parasitic programming. After the selection, the programming unit drives the wanted set or reset pulse to program the resistivity state. An example of sequential programming is shown in figure 4.

![Figure 4. Input (a) and output (b) drivers for PCM-based crossbar](image)

IV. SIMULATION RESULTS

To validate the PCM-based switchbox, we characterized its performances in terms of area and write time. Comparison with memory elements traditionally used in FPGA like MOS SRAM ST cells [6] and flash memories elements [7], are then used to benchmark the structure. Furthermore, since we are interested in the utilization of PCMs in reconfigurable circuits, we evaluated the delay and area of PCM-based FPGAs for a benchmark of widely used logic circuits.

**A. Performance characterization of single elements**

Table I shows some results of characterization in terms of area and write time for the proposed solution and traditional FPGA memory nodes. We see that the proposed PCM switchbox is the most compact solution, even with the impact of the programming current on access transistor. This advantage is due to the reduction of the memory front-end footprint to only metal lines, compared to 5 transistors for the SRAM cell and 2 transistors for the Flash solution (one pull-up transistor coupled to a floating gate transistor). It is also worth noticing that PCMs offer a significant reduction in writing time versus flash technology. In our context, it is possible to reduce the area by 3.4 and the writing time by 16.6 compared to an equivalent Flash technology.

<table>
<thead>
<tr>
<th>TABLE I. TECHNOLOGY PERFORMANCE EVALUATION (2X2 SWITCHBOX)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cell elements</td>
</tr>
<tr>
<td>--------------------------------------------</td>
</tr>
<tr>
<td>SRAM</td>
</tr>
<tr>
<td>Flash cell</td>
</tr>
<tr>
<td>PCM cell</td>
</tr>
<tr>
<td>Flash vs. PCM</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

**B. FPGA Evaluation flow**

1) Methodology

We used a set of logic circuits taken from the MCNC benchmark [14], which we first synthesized using the ABC tool [15]. We then performed the technology mapping with a library of 4-input LUTs (K=4) using ABC as well. Subsequently, we performed the logic packing of the mapped circuit into CLBs with N=10 BLEs per CLB and I=22 external inputs using TPACK [16]. Finally, the placement and routing were carried out using VPR [16]. We synthesized the considered benchmark twice. First we designed the SRAM-based LUTs and MUXs in the CMOS 65 nm process, using a pass-gate design. Then, we replaced the SRAM-based switchboxes by their PCM counterparts.

2) Simulation of large circuits

We mapped the benchmark in PCM- and SRAM-based FPGAs, and we simulated the FPGA delay. The delay estimation is shown in figure 6. The benchmarks in figure 6 show a delay reduction ranging from 38% to 51%, with 44% on average. The main benefits of using PCMs instead of SRAM cells are the compact area of the cell and the lower internal resistance of data paths. As a matter of fact, we extracted from our design kit the internal resistance of a pass-gate cell which is in the order of the on-resistance of an n-type transistor (9.1 kΩ); while the experimental results show that PCMs have a lower on-resistance that has been reported close
to $4 \, k\Omega$ [10]. This makes the PCM-based switchboxes potentially faster than the SRAM-based counterparts.

![Figure 6. Delay estimation for FPGAs synthesized with GST-PCM- and SRAM-based Switchboxes](image)

3) Impact of PCM technologies

We demonstrated with the previous simulations that the combination of the switchbox design and the GST technology results in a significant delay improvement. In the following, we showcase the impact of the PCM technology type on the delay improvement. We remind the reader that other materials exist, which may replace the GST as a phase-change material. These include GeTe and GeTeCα%. Besides the difference in RESET current and time, the on-resistance depends on the chosen material and it has an impact on the routing path resistance. We simulated the delay of the FPGA benchmark with various resistances corresponding to different materials. We notice that the delay improvement is linear with the decrease of the on-resistance. However, the delay sensitivity is low: a decrease of the on-resistance by 2 orders of magnitude from about 4000 $\Omega$ to about 50 $\Omega$ causes a delay reduction of only 5%.

![Figure 7. Variation of the FPGA critical path delay with the PCM on-resistance (delay averaged over the whole benchmark set)](image)

V. DISCUSSION AND OPPORTUNITIES

The presented architecture can be employed to store configuration data in programmable circuits. Even if the PCM technology seems to be already mature for mass production, research and development on materials and cell design activities are still maintaining a key role for non-volatile resistive memories. The proposed circuit could be generalized to any other resistive memories. It is then interesting to consider technologies or cell design with better characteristics such as higher programming speed or lower programming current. For example, programming current plays an influence on the size of selection devices and the write energy. New cell design could reduce programming energy, as shown in [17], where a sublithographic heater is used to increase current density.

VI. CONCLUSIONS

This paper introduced a novel switchbox circuit, designed to replace traditional configuration circuitry in reconfigurable logic circuits. The switchbox is only based on PCM resistive memories, to build the full interconnect points. Extra buffers are used to mix programming and data paths. We showed that the proposed solution can reduce the size of the elementary node by a factor of 3.4x compared to traditional flash memories. Finally, we evaluated the impact in FPGA design and we showed that the critical path delay can be reduced by a factor up to 51% due to the speed of PCM-based switchboxes.

REFERENCES