A Design of a Non-Volatile PMC-Based (Programmable Metallization Cell) Register File

Salin Junsangsri,

ECE
ECE Dept, Northeastern University
Boston, MA 02115, USA
(1) 857-225-3755

Junsangsri.s@husky.neu.edu

Jie Han

ECE Dept, University of Alberta
Edmonton, Alberta, Canada T6G
2V4
(1) 780-492-1361

jhan8@ualberta.ca

Fabrizio Lombardi

ECE Dept, Northeastern University
Boston, MA 02115, USA

(1) 617-373-4854

lombardi@ece.neu.edu

ABSTRACT

This paper presents the design of a non-volatile register file using cells made of a SRAM and a Programmable Metallization Cell (PMC). The proposed cell is a symmetric 8T2P (8-transistors, 2PMC) design; it utilizes three control lines to ensure the correctness in its operations (i.e. Write, Read, Store and Restore). Simulation results using HSPICE are provided for the cell as well as the register file array (both one- and two-dimensional schemes). At cell level, it is shown that the off-state resistance has a limited effect on the Read time, because in the proposed circuit the transistor connecting the PMCs to the SRAM is off. While having no significant effect on the Store time, the time of the Restore operation depends on the value of the off-state resistance, i.e. an increase in off-state PMC resistance causes an increase in Restore time. Comparison between non-volatile register files utilizing either PMCs, or Phase Change Memories (PCMs) is provided.The register file using PMCs has a faster Store and Read times than the PCM-based counterpart; this is mostly caused by the difference in resistance values for these two non-volatile technologies. The lower delay involved in these operations confirms that the proposed PMC-based register file offers significant advantages in terms of delay performance.

Keywords
Register File; Non-volatile Operation; Emerging Technology

1. INTRODUCTION
Non-volatile memories (NVMs) have gained considerable attention due to the requirement of portable storage for many computer applications such as consumer electronics [1]. In the past, they have been used as secondary memory for long term persistent storage; however, technology advances have made possible to have NVMs operating higher in the memory hierarchy [2].

A register file is used at the top of the memory hierarchy. It is usually implemented by fast Static Random Access Memories (SRAMs); however, SRAMs are volatile circuits [3] and in many applications, non-volatile storage is required to meet both temporal and space localities in memory accesses [13]. PMC has been recently proposed as a candidate for the next-generation of non-volatile memory due to its simple structure, high resistance ratio, multilevel capability, low power consumption, favorable scalability and high operational speed [3]. This manuscript considers the Programmable Metallization Cell (PMC) (also referred to as the conducting bridge random access memory, CBRAM, or electrochemical metallization ECM) as non-volatile memory element; it deals with the design of a register file whose non-volatile operation is accomplished by using PMCs. The operations related to the SRAM and the non-volatile elements within a cell are analyzed in details for different figures of merit using an 8T2P (8 transistors, 2 PMCs) configuration. The cell operations are then extended to one- and two-dimensional array schemes for register-file implementation. Comparison with the same scheme but employing Phase Change Memories (PCMs) as non-volatile elements is presented.

2. REVIEW
This section presents a brief review of a register file and PMC as relevant to the proposed designs.

Register file: In a microprocessor, the memory architecture is hierarchically organized. On-chip memory is utilized for very fast access time; cache and register files are two of the most commonly used on-chip memories [5]. A register file is embedded in the central processing unit (CPU) and stores both data and mapping information (such as locations); these locations store specific addresses for loading programs, or data to meet spatial and temporal locality requirements [13].The register file is usually designed as a memory array with the fastest access time in the hierarchy. In the register file, two memory technologies are often used. (i) Static random access memory (SRAM): SRAMs are fast, but they are volatile and each cell requires at least 6 transistors. (ii) Dynamic random access memory (DRAM): DRAMs requires refresh and are slower than SRAMs; they require less area due to the reduced circuit complexity. DRAMs are also volatile. A SRAM consists of two crossed inverters with two additional transistors for data control; this configuration is generally known as 6T.

Programmable Metallization Cell: The SRAM is a volatile memory circuit; different devices and operations (such as Store and Restore) are commonly used for non-volatile storage. Some of

the non-volatile storage operations are very important when *temporal and space locality considerations* are required in a register file for preserving crucial data at the highest level of the memory hierarchy [13]. The non-volatile property also permits continuous storage of data for improved performance. The Programmable Metallization Cell (PMC) is a non-volatile memory, whose resistance changes by modifying the biasing voltage across the cell. The change in resistance occurs by having metallic ions pass through a solid electrolyte and the subsequent formation and dissolution of a metallic conductive filament (CF) connecting the two electrodes [3].The PMC has two states: *set* (OFF to ON state transition) and the *reset (*ON to OFF state transition). The set state occurs when the positive voltage bias is higher than the positive threshold voltage; this will cause an electrochemical reaction to form a conducting link between the top and bottom electrodes such that ions can tunnel through the cell. Thus, the resistance will change from R_{off} (off state resistance) to R_{on} (on state resistance), as a set state. When the negative voltage is biased over the negative threshold voltage, there is no connection between the electrodes. Therefore, the resistance changes from R_{on} to R_{off} due to the reset state.

3.   CELL DESIGN AND EVALUATION

The proposed cell for a register file is a symmetric circuit; two PMCs (storing data of opposite values) are utilized. The non-volatile elements are connected to a SRAM, such that data can be stored (from SRAM to non-volatile elements), or restored (from non-volatile elements to SRAM). The PMCs are connected to the storage nodes (D and DN) of the SRAM in a balanced (symmetric) scheme, as described in more detail next. This design is referred to as the shared control cell (Figure 1); the two PMCs are controlled by a single signal (given by Ctrl1). As this is an 8T2P symmetric cell, hence the times of the Store and Restore operations are the same for both values ('0' and '1'). The Store and Restore operations of the 8T2P cell are controlled by adjusting the voltages at nodes BL, BLN and WL using 3control signals described in more detail next.

Store Operation: The Store operation is used to transfer the data from the SRAM to the non-volatile elements (i.e. the two PMCs are directly connected to D and DN). In this 8T2P cell, the voltage at node Ctrl2 must be at 0V to allow the supply voltage to the SRAM. The voltage at node WL is high to allow the voltages at BL and BLN to the SRAM (through transistors M5 and M6). Moreover, Ctrl0 is high to allow the voltages at Ctrl1to affect the PMCs. Ctrl1 must be biased to change both PMCs, so the signal on Ctrl1 is (with a full swing from 0V to V_{dd}) reduces the Store time and allows only one PMC to '1' (thus forcing the other node of the SRAM to become '0' in a shorter time). The values of BL and BLN depend on the desired value to be stored.

Restore Operation: The Restore operation occurs when either the supply is restored and/or the data from the non-volatile element(s) is written to the SRAM. Instability in the final step (i.e. the voltage at the storage node D could be close to the voltage at node DN) must be avoided; this is accomplished by having BL, BLN and WL high. Moreover, a PMOS is used to connect or disconnect the supply voltage from the circuit by utilizing Ctrl2; also, Ctrl0 is low to isolate both PMCs from the SRAM. When the voltages at D and DN are close (at a time referred in this paper as the switching restore time, or SRT), Ctrl2 connects the supply voltage to the cell and Ctrl0 allows the voltages at Ctrl1 to the PMCs and restore the data.
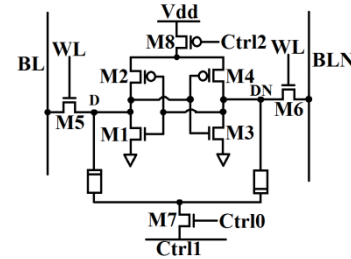

Figure 1. Shared control cell (8T2P)

Next, the proposed cell is evaluated with respect to its operations. For the Read and Write operations, only the SRAM is involved; due to lack of space these operations are not treated furthermore.

The two operations (Store and Restore) dealing with non-volatile storage are discussed in more detail next (the same process is applicable to both '0' and '1' for the value at nodes D and DN due to the symmetric nature of the proposed cell). Simulation is performed at 32-nm feature size using the Predictive Technology Model (PTM) [9] under the conditions of the parameters shown in Table 1. M7 and M8 allow the correct voltage values to the cell, while control is enforced by Ctrl0, Ctrl1 and Ctrl2 for the correct execution of the Store and Restore operations.

Store Operation: The Store operation is used to transfer the data from the SRAM to the non-volatile elements (i.e. the two PMCs are directly connected to D and DN). The Store operation requires a higher voltage than the other operations, because the resistances of both PMCs are changed by voltage biasing for state switching. The stored values at the PMCs are connected to D and DN, therefore the Store operation starts at the same time on both PMCs. The Store operation is similar to the Write operation for the SRAM; Ctrl0 and Ctrl1 are used as control signals to bias the voltage for storing the desired data in both PMCs. This operation requires an increase of the bias voltage to 3.5V. Moreover, a voltage pulse is needed at Ctrl1 to change the values stored in both PMCs. The timing diagrams of the relevant signals are shown in Figure 2.

The Store operation requires the PMCs to change state in a stable fashion; this process requires more time to change the resistive states from R_{on} to R_{off} than R_{off} to R_{on}. As the resistive states of both PMCs must be changed during this operation, then the proposed cell requires a square pulse signal as input. Figure 3 shows the relationship between the Store time (for the '0' value) and the PMC bias voltage in a semi-logarithm plot scale; the Store time is inversely proportional to the bias voltage, i.e. a smaller bias voltage causes a larger Store time. However, a larger voltage implies larger power dissipation too.

Restore Operation: The Restore operation occurs when either the supply is restored and/or the data from the non-volatile element(s) is written to the SRAM. For the Restore operation, 0.9V (same as the supply voltage at 32nm) is used as voltage value to keep a low power consumption in the circuit. A *switching restore time* (or STR) must be allowed to ensure that both D and DN have correct and stable values (for similar reason as in the Store operation). The values and corresponding timing diagrams of the proposed cell for BL, BLN, WL, Ctrl0, Ctrl1, Ctrl2 are given in Figure 4; the STR is given by 425ps.
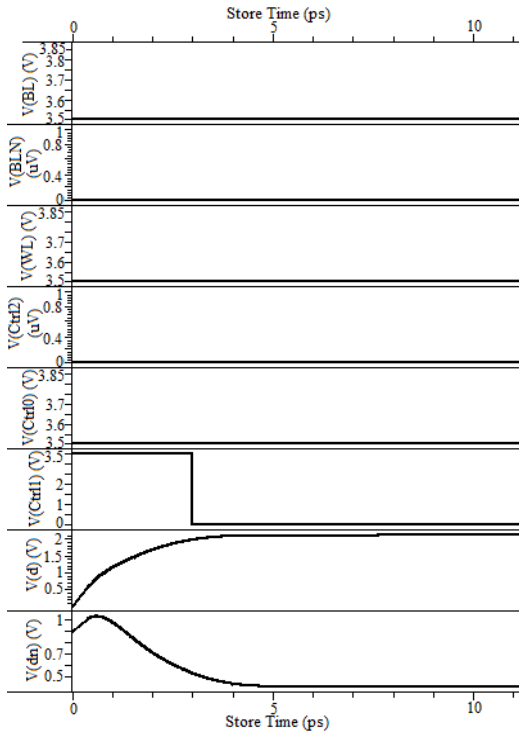
Figure 2. Timing diagram of the Store operation for BL, BLN, WL, Ctrl0, Ctrl1, Ctrl2, V(D) and V(DN).

Table 1. Parameters used in simulation [3]

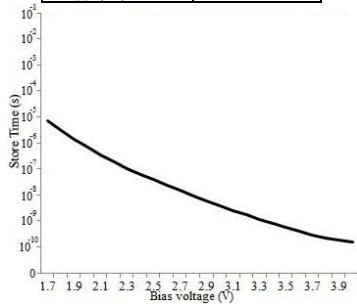| Parameter	Value
R_{on} (Ω)	30.0625k
R_{off} (Ω)	99.9579M
NMOS (nm)	32
PMOS (nm)	32
V_{dd} (V)	0.9

Figure 3. Store time vs.PMC bias voltage

In Figure 4, the voltages of BL, BLN, WL, and Ctrl2 are '1' for 425ps (i.e. the STR); then these voltages have a value of '0' to copy the data from the PMCs to the SRAM. The results (Figure 5) show that the Restore time for the proposed 8T2P cell is 470.73ps.

Consider next the relationship between the cell operations and the PMC resistance; Figure 6 shows the Read/Write times (for both '0' and '1') of the SRAM when varying the value of the off-state resistance of the PMC (and keeping the on-state resistance to its default value as in Table 1).

Figures 6 and 7 show the Store and Restore times (same for '0' and '1') by varying the off-state PMC resistance; while the value of this resistance has no significant effect on the Store time, the time of the Restore operation depends on the value of the off-state resistance. So, an increase in off-state PMC resistance causes an

increase in Restore time. The simulation results show that the off-state resistance has a limited effect on the Read time, because the transistor connecting the PMCs to the SRAM is off.

Static Noise Margin (SNM): The static noise margin (SNM) is defined as the minimum static voltage capable of change the state of a memory cell, the SNM is widely used as a stability criterion [11]. It is well known that the presence of additional hardware (such as a non-volatile element) causes a reduction in the SNM [12]; simulation has found that the SNM for the proposed design is 0.22843V (the 6T (volatile) SRAM has a SNM of 0.2581V). This corresponds to a 15% reduction; this small reduction is due to the utilization of the two control signals Ctr0 and Ctr1 in the operation of the proposed cell.

Power Dissipation: Power dissipation and the power delay product (PDP) are also important metrics in the operations of the cell; the results are shown in Table 2.

Table 2. Power dissipation and PDP of proposed cell

Operation	Power Dissipation	PDP (fWs)
Write	45.405uW	1.106
Read	17.232uW	0.132
Store	46.98mW	185.631
Restore	96.061uW	44.711

As the proposed cell is a symmetric scheme, there is no difference in operation due to the data values; Table 2 shows that only Store incurs in a substantial power dissipation (and corresponding higher PDP); this is caused by the higher voltage level required for writing to the non-volatile elements. However, operationally critical data is preserved in the SRAMs by evoking a Store operation only in occasional cases [14]. Usually high performance computing utilizes very often the non-volatile Restore by which the data stored in the non-volatile elements is loaded into the SRAM, so avoiding data search in the lower memory levels (and incurring in longer latency delays).
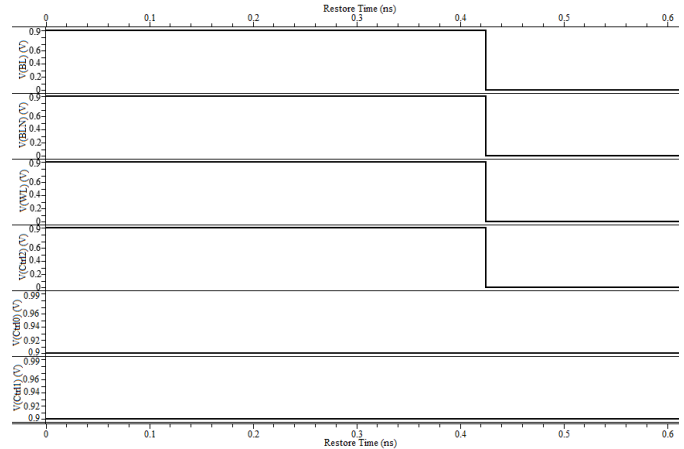

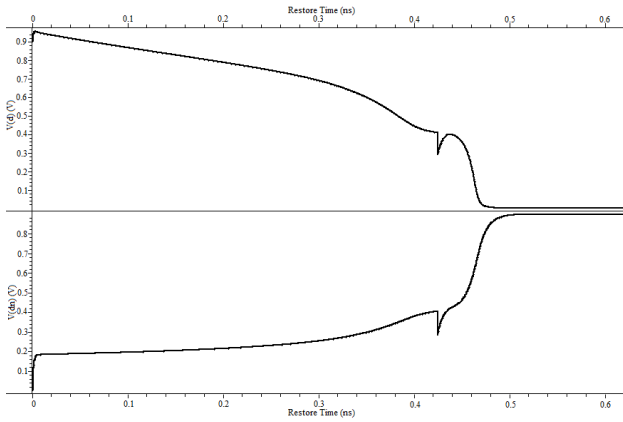**Figure 4. Timing diagram of the Restore operation for BL, BLN, WL, Ctrl0, Ctrl1, and Ctrl2.**

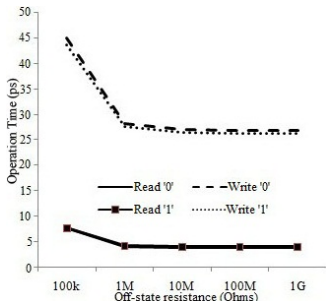**Figure 5. Restore time at nodes D and DN**

Figure 6. Read and write times vs. off-state PMC resistance.

4. REGISTER FILE EVALUATION

In this paper, the register file is designed as consisting of cells arranged in one-dimensional (1-D) and two dimensional (2-D) schemes.

One-Dimensional (1-D) Array: In the 1-D configuration, all register files are connected to common BL and BLN; moreover, simulation considers only one cell of the array to be selected (all other N-1 cells are left unselected).
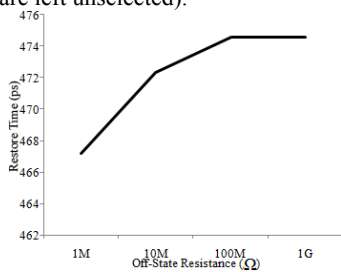

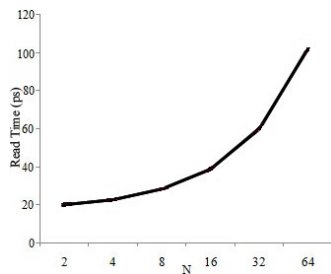**Figure 7. Restore time vs. off-state resistance.**

Figure 8. Read time vs. dimension N

As shown in Figure 8, a larger array size causes more parasitic capacitance resulting in a longer delay for the Read operation.

Simulation has confirmed that N has no significant effect on the Write, Store and Restore times, i.e. nearly constant values. These results confirm that the parasitic capacitance has hardly any effect on the Write, Store and Restore operations due to the bias voltage at BL and BLN, i.e. only the Read operation (Figure 8) shows a dependency on N.
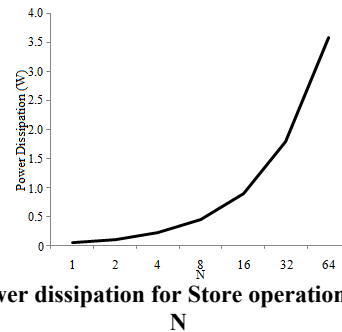

Figure 9. Power dissipation for Store operation vs. dimension N

Power dissipation of the operations of the proposed cell increases at a larger array size (Figures 9 and 10). The Store operation incurs in the highest power dissipation (Figure 9) due to the larger bias voltage. The power dissipation of the Write, Read and Restore operations are shown in Figure 10; all these operations show a dependency of power dissipation on N. The SRAM operations (Read and Write) show the largest increases (the pre-charging process required for the Read operation is the cause of the largest dissipation when N is increased).

Two-Dimensional (2-D) Array: A register file is usually implemented as a two-dimensional (2-D) array of dimension N; it is assumed that only one cell (the so-called center cell) is selected, all other cells are unselected. By varying the dimension of the 2-D array, simulation shows that the Write, Store and Restore times are only marginally affected because the bias voltages on BL and BLN are sufficiently high to overcome the parasitic capacitance; BL and BLN in the Read operation are not biased, so the effect of the parasitic capacitance makes the Read time longer at a larger array size (Figure 11).
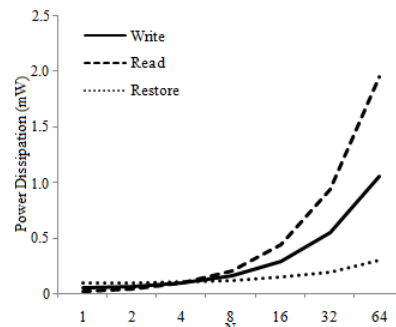

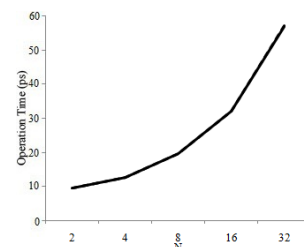**Figure 10. Power dissipation for Write, Read, and Restore operations vs. dimension N**

Figure 11. Read time vs. N for 2-D array

The Restore operation for a larger array size requires a longer time in the first step (due to the larger parasitic capacitances at the bit and word lines); a value of 500ps is found by simulation and used as STR, this corresponds to an increase of 75ps from the single cell case (i.e. nearly 15%).
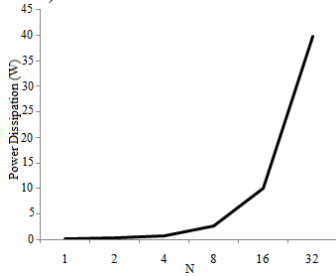

Figure 12. Power dissipation of Store operation vs. N

When considering power dissipation, a larger memory size causes a power dissipation higher than for example the 1-D array. The result for the Store operation of a 2-D array is shown in Figure 12; its trend is similar to the Store operation of a 1-D array. The power dissipation for the three other operations is shown in Figure 13.

5. COMPARATIVE EVALUATION

In this Section, the technology of the resistive non-volatile elements is changed and the corresponding register files are assessed. A PCM (phase change memory) element is also considered [6]. This comparison is performed using the parameters of Table 1; in particular the supply voltage is 0.9V (for the Read, Write and Restore operations) and a Store operation requires 3.5 V as bias voltage.
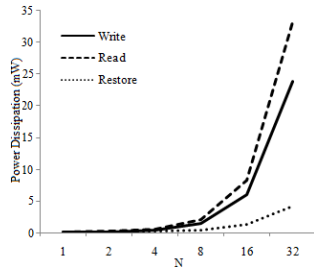

Figure 13. Power dissipation of Write, Read, and Restore operations vs. N

A MOSFET has a standard P/N width ratio of 2.5 (as used throughout the previous sections), so the W/L ratio must be increased to at least 9 to ensure that a PCM can perform a Store operation. The simulation results for the register file cells made of these two technologies are shown in Table 3 (bold entries denote the best values); the delay is the same for the '0' and '1' data values due to symmetry in the cell.

Table 3. Operation delays of PMC and PCM cells when W=9L, P/N ratio = 2.5

Operation	PMC Cell	PCM Cell
Write	26.16ps	**25.84ps**
Read	**8.07ps**	9.276ps
Store	**3.951ps**	287.10ns
Restore	470.73ps	**296.84ps**

Note that a 6T (non-volatile) SRAM cell incurs in delays of 24.583ps, and 6.8618ps for the Read and Write operations, respectively; so a penalty is incurred for the presence of non-volatile elements in a register-file cell. Table 3 shows that a PCM-based cell has worse performance in two out of four operations (so

except for the Write and Restore). A state change in a PCM incurs in a more substantial delay than a PMC, especially for a Store operation (nearly two orders of dimensions); this occurs because the PCM requires a rather long time [6] to change the material structure of the GST compound (and hence its resistivity). The increase in W/L is needed for achieving the required bias voltage for the Store operation in a PCM-based cell.

Next, consider the 2-D level comparison for the register file; results are obtained for square register files of dimension 16 and 32. The results in Figure 14 show the Write and Read times for the PMC and PCM-based register files. These results show that the Write time is not significantly affected by the resistive element type and N; the Read time however, increases at a higher value of N and is substantially larger in a PCM-based register file.

Figures 15 and 16 show the Store and Restore times of the PMC- and PCM-based register files when N is 16 and 32. The increase in array size does not significantly affect the Store or Restore operations. Figure 15 shows the benefit of storing data into a PMC (rather than a PCM) due to the order of magnitude time difference between these non-volatile technologies. Figure 16 shows that the reverse (i.e. for a Restore operation), the register file made of PCM-based cells is significantly better than the PMC counterpart. This occurs due to the resistances at the on-/off- states for these technologies, a smaller resistance causes a faster Restore time due to the larger voltage across the cell.
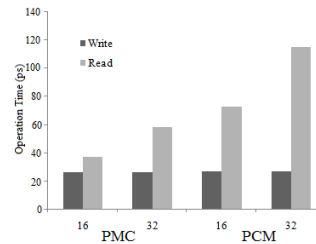

Figure 14. Write and Read times of PMC and PCM-based 2-D register files vs. N
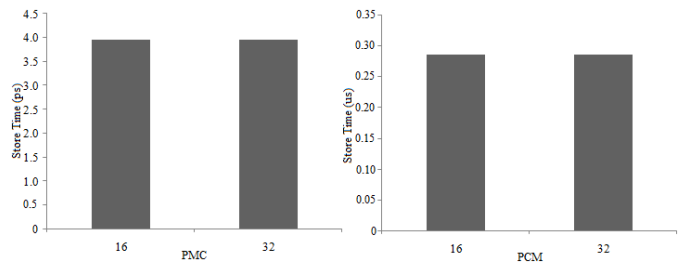

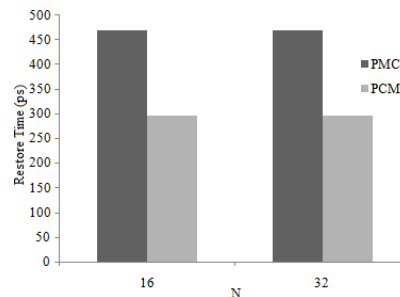**Figure 15. Store time of PMC and PCM-based 2-D register files vs. N**

Figure 16. Restore time of PMC and PCM-based 2-D register files vs. N

6. CONCLUSION

This paper has presented the design and evaluation of a non-volatile register file; the cell in the register file consists of a SRAM and a Programmable Metallization Cell (PMC) in a symmetric 8T2P (8-transistors, 2PMC) design. Three control lines are utilized in the proposed cell to ensure the correctness in its four operations (i.e. Write, Read, Store and Restore. Simulation results using HSPICE have been provided for the cell as well as the register file array (both one- and two-dimensional schemes). Comparison in operation times between non-volatile register files utilizing either PMCs, or Phase Change Memories (PCMs) have also been assessed. Table 4 shows the ranking of these register files. The following features are evident. (i) The register file using PMCs has a faster Store and Read times than the PCM-based counterpart; this is mostly caused by the difference in resistance values for these two non-volatile technologies and the delay involved in these operations [6]. (ii) The Write and Restore times are better for a PCM-based register file; as these operations are not as frequent as the others at the highest level of a (non-volatile) memory hierarchy [2], then the proposed PMC-based register file offers significant advantages in terms of delay performance.

Table 4. Ranking of PMC- and PCM-based register files

Operation	PMC	PCM
Write time	2	1
Read time	1	2
Store time	1	2
Restore time	2	1

7. REFERENCES

[1] R. K. Gupta, S. Krishnamoorthy, D. Y. Kusuma, P. S. Lee, M. P. Srinivasan "Enhancing charge-storage capacity of non-volatile memory devices using template-directed assembly of gold nanoparticles" Nanoscale Issue 7 pp. 2296-2300, Jan. 2012

[2] H. Li, Y. Chen, L. Benini, G. De Micheli, B. Al-hashimi, W. Mueller "An Overview of Non-Volatile Memory Technology and the Implication for Tools and Architectures" Design, Automation and Test in Europe pp.731-736, 2009

[3] D. Liu, Nannan Wang, Guang Wang, ZhengZheng Shao, Xuan Zhu, ChaoyangZhang"Programmable metallization cell based on amorphous $La_{0.79}Sr_{0.21}MnO_3$ thin films for memory applications" Journal of Alloys and Compounds, vol. 580, pp. 354-357, Dec. 2013. doi: 10.1016/j.jallcom.2013.06.095

[4] P. Junsangsri, F, Lombardi and J. Han "HSPICE Macromodel of a Programmable Metallization Cell (PMC) and its application to memory Design" IEEE/ACM international Symposium on NANOARCH, pp. 45-50, July 2014, doi: 10.1109/NANOARCH.2014.6880477

[5] Hao YAN, Yan LIU, Dong-hui WANG, Chao-huan HOU "A Low-power 8-Read 4-Write Register File Design" IEEE Conference Publications. On Microelectronics and Electronics, pp. 178-181, Sept. 2010. doi: 10.1109/PRIMEASIA.2010.5604933

[6] P. Junsangsri and F. Lombardi "A New Comprehensive Model of a Phase Change Memory (PCM) Cell" IEEE Trans. On Nanotechnology, vol. 13, no. 6, pp. 1213–1225, Nov. 2014.

[7] M. H. Kaffashian, R. Lotfi, K. Mafinezhad, H. Mahmoodi "Impact of NBTI on performance of domino logic circuits in nano-scale CMOS " Microelectronics Journal, vol. 42, no. 12, pp. 1327-1334, Dec. 2011

[8] K. Osada (2011). Fundamental of SRAM Memory Cell. In K. Osada, K. Ishibashi (Eds.), Low Power and Reliable SRAM Memory Cell and Array Design (pp. 6) Springer Series in Advanced Microelectronics, vol. 31

[9] PTM Model, 32-nm MOSFETs. Retrieved from http://www.ptm.asu.edu

[10] P.E. Dodd, F.W. Sexton "Critical charge concepts for CMOS SRAMs" IEEE Trans. On Nuclear Science, vol. 42, no. 6, pp. 1764 – 1771, Dec. 1995.

[11] E. Seevinck, F. J List, J. Lohstroh "Static-Noise Margin Analysis of MOS SRAM Cells" IEEE Journal of Solid-state circuits, vol. 22, no. 5, pp. 748-754, Oct. 1987

[12] W. Wei, K. Namba, J. Han and F. Lombardi, "Design of a Non-Volatile 7T1R SRAM Cell for Instant-on Operation." IEEE Transactions on Nanotechnology, vol. 13, no. 5, pp. 905-916, 2014.

[13] Q. Zhu, K. Vaidyanathan, O. Shacham, M. Horowitz, L. Pileggi, F. Franchetti "Design Automation Framework for Application-Specific Logic in Memory Block" IEEE ASAP 23rd pp.125-132, July 2012.

[14] W. Wei, K. Namba, and F. Lombardi, "Design and Analysis of Non-Volatile Memory Cells for SEU Tolerance," Proc. 17th IEEE Symposium on Defect and Fault Tolerance in VLSI and Nanotechnology Systems, pp. 69-74, Amsterdam, October 2014