Hybrid RRAM-NEM Relay Design for Programmable Logic Devices

White paper

Akash Levy

March 18, 2019

1 Motivation

Field-programmable gate arrays (FPGAs) are programmable logic devices (PLDs) that enable rapid prototyping of digital circuits and provide an alternative to the expensive application-specific integrated circuit (ASIC) design process today. However, when compared with ASICs performing the same function, SRAM-based FPGAs typically have 10-40 times lower logic density, 3-4 times higher delay, and 5-12 times higher dynamic power dissipation [1], [2], [3]. This is because reconfigurable components incur large overhead—instead of being packed tightly, configurable logic blocks (CLBs) in FPGAs are laid out relatively far apart to make space for routing components, resulting in greater interconnect delay, power dissipation, and chip area. The disparity in the performance of FPGAs and ASICs provides motivation to find ways to reduce reconfigurability overhead, so that ultimately PLDs can become competitive with ASICs at manufacturing scale. Possible solutions lie with PLD designs that trade off fine-grained CLBs with more specialized blocks e.g. coarse-grained reconfigurable arrays (CGRAs) [4], or with emerging nanotechnologies that enable novel design strategies in the space of reconfigurable semiconductor devices (to be discussed later).

Another topic of interest in semiconductor design today is "normally off, instantly on" device operation, which is critical for devices with weak, unreliable, or intermittent power sources. Energy-harvesting devices, wearables, and low-power nodes in the Internet of Things (IoT) will need to be able to operate in this mode to prevent unnecessary power dissipation between periods of action. Today, FPGAs need to load their configuration from off-chip memory into SRAM on startup, which incurs a significant energy cost. This has motivated the development of integrated non-volatile memory (NVM) that can replace (or work well in conjunction with) on-chip SRAMs to enable intermittent computing.

Here, we propose the use of two emerging nanotechnologies in PLD design for bridging the gap between PLDs and ASICs: resistive random access memory (RRAM) and nanoelectromechanical (NEM) relays. We describe a hybrid RRAM-NEMS-based design for implementing PLDs with both "normally off, instantly on" device operation and low-overhead reconfigurable routing components. We exploit the properties of RRAM and NEM relays in a synergistic fashion to develop a novel NV-SRAM-based router that has all of its read circuitry implemented in back-end-of-line (BEOL), with RRAM cells directly controlling gate voltage of hysteretic NEM relays that behave as routers. We compare the design of a CGRA developed using the techniques described here versus designs that use (1) NEM relays only, (2) RRAM only, and (3) standard CMOS technologies (baseline).

2 Emerging nanotechnologies for PLDs

2.1 NEM relays

With technology scaling, it is becoming increasingly difficult to design FPGAs using NMOS pass transistors for programmable routing. An NMOS pass transistor introduces a threshold (V_t) drop when passing high voltage level, after which the voltage level must be restored. Unfortunately, pass transistor threshold voltage (V_t) cannot be further reduced due to leakage power constraints.

NEM relays are switches that are actuated electrostatically at the nanoscale. Their properties have been studied in depth, and they have gained interest for their potential as pass transistor replacements in FPGAs [5] [6]. NEM relays have extremely low static power dissipation and low ON-state resistance—however, they switch much slower than transistors (on the order of nanoseconds) and typically operate at higher voltages. [PRIYANKA: How does the static power and ON resistance compare to CMOS? Would be good to add some representative numbers for a technology.] NEM relays have a sharp I_{DS}-V_G curve (often referred to in literature as an infinite subthreshold slope) and also exhibit hysteresis in switching state based on applied gate voltage (V_G). This has led to NEM relay-based SRAM designs utilizing this property.

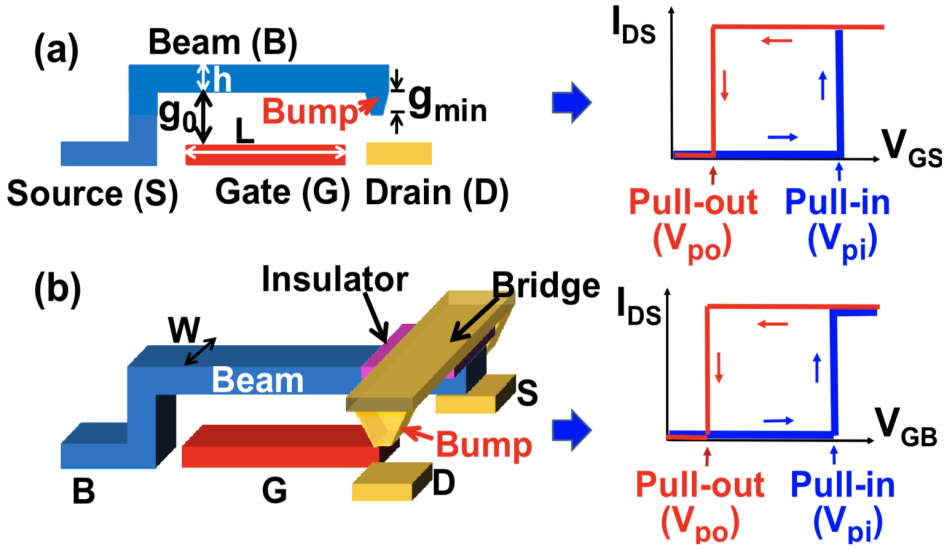


Figure 1: Figure adapted from [6]. (a) A 3T NEM relay and its I_{DS}-V_{GS} characteristics; (b) A 4T NEM relay and its I_{DS}-V_{GB} characteristics. The beam is insulated from the bridge for S/D contacts [10].

There are two kinds of NEM relays we will focus on: 3-terminal (3T) and 4-terminal (4T) devices (see Figure 1). 3T devices have a source, drain, and gate, and the voltage that controls the beam is the voltage between the gate and source. This means that the source voltage affects the beam state, which is undesirable for a device that needs to behave as a pass transistor. The 4T device mitigates this problem by having the beam electrically isolated from the "bridge" that connects the source and drain [7]. Table 2.1 [PRIYANKA: The table number doesn't get generated correctly in the pdf.] (adapted from [8]) describes characteristics of NEM relays found in literature. NEM relays can be fabricated at low-temperatures ($< 400^{o}$C) making them compatible with CMOS back-end-of-line (BEOL), and they make use of materials already available in foundries. Simple lateral designs require 4-7 masks at the top layer of the device [9]. Research into scaling these devices to smaller sizes and integrating them in 3-dimensions is an ongoing effort.

Several studies have examined the performance implications of using NEM relays in PLDs. Architectural analysis with CMOS-NEM routing and logic lookup tables was performed in [26]. Using a 3-D place-and-route scheme with NEM-based routers, a maximum of 41.9% performance gain was achieved over the baseline. However, this study assumed that multiple NEM devices could be viably stacked in 3-D, which has yet to be experimentally demonstrated. Another study claimed a 10-fold reduction in leakage power, 2-fold reduction in dynamic power, and 2-fold reduction in area [5]. A half-select programming scheme was used for the NEM relays which was verified experimentally (see Figure 2), but the use of a 3T device and a small hysteresis window would probably prevent these devices from functioning properly as pass transistor replacements—when a high voltage level is being passed, V_{gs} would fall below V_{po} causing the NEM relay to switch to the off-state unintentionally. The study also used optimistic device parameters/dimensions that have yet to be physically realized.

2.2 RRAM

RRAM is typically based on a metal oxide (such as HfO_2, Al_2O_3, Ta_2O_5) that is already used in CMOS fabrication processes e.g. for gate oxides. Usually, the device programming works via a filamentary process involving soft electric breakdown of the metal oxide to create conductive filaments that can be broken or reformed in subsequent programming (see Figure 3) [27]. RRAM can store multiple bits/cell, however in binary operation, it has a high-resistance state (HRS) which usually corresponds to a '0', and a low-resistance state (LRS) which usually corresponds to a '1'. RRAM has high-speed operation, low operating power, good switch endurance ($> 10^6$ cycles), geometric scalability down to 10 nm x 10 nm, and a relatively simple fabrication process. Many materials have been reported to have resistive hysteresis, and many architectures have been proposed for decoding/sensing the resistance state. Some examples which reported high endurance are given in Table 2.

RRAM has been studied for use in FPGAs as an SRAM replacement. An FPGA was fabricated using a 1T2R configuration with two RRAM cells having opposite resistance states to implement a voltage divider controlling a pass

Table 1: NEMS devices reported in literature. Adapted from [8].

Research Group	Type	Area	c-gap	V_{pi}	R_c	Cycles	Material	Circuit
First CNT [11] [12]	3T	-	30 nm	4.5V	1 MΩ	>1	CNT	No
First Top-Down (KAIST) [13]	3T	-	40 nm	4V	-	> 10	TiN	No
Out-of-Plane (UC Berkeley) [7] [14] [15]	4T	-	80 nm	<1V	1kΩ	$>10^8$	W	Yes
	6T	-	80 nm	<1V	1kΩ	$>10^8$	Ru	Yes
	6T	-	250 nm	10V	-	1	TiO2	Yes
In-Plane (Stanford University) [16] [17] [18]	5T	-	-	7.9 V	3 kΩ	10^8	Pt	Yes
Curved (NEMIAC) [19] [20]	3T 4T	15 μm^2	50 nm	0.5 V	5 kΩ	$>10^8$	a-C	Yes
SiC Relays (CWR University) [21] [22]	3T	1 μm^2	100 nm	15 V	10 MΩ	>14,000	SiC	Yes
Sandia National Lab [23]	3T	(10 μm^2)	30 nm	>4 V	-	$2 \cdot 10^6$	Ru	No
Cornell University [24]	3T	(3 μm^2)	200 nm	10 V	10 MΩ	-	-	No
Piezoelectric (CM University) [25]	4T	58 μm^2	-	10 mV	16 kΩ	-	Pt	No

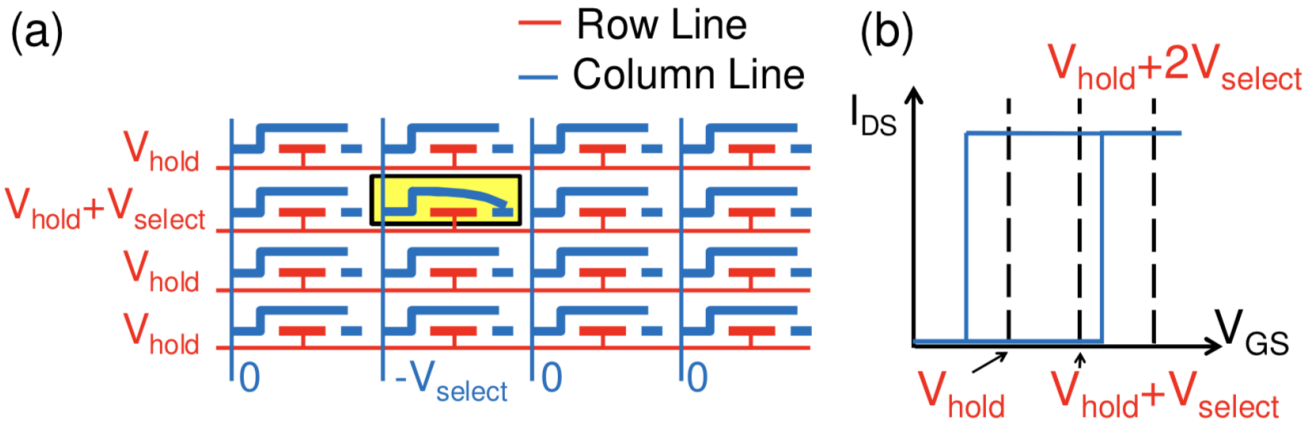


Figure 2: NEM relay half-select programming (adapted from [5]). (a) Array of relays. (b) NEM relay I-V curve with half-select programming voltages.

Table 2: Table of some RRAM device properties reported in literature. Adapted from [28].

Publication & Year	VLSI 2014	VLSI 2014	IEDM 2014	IEDM 2014	VLSI 2015	VLSI 2015
2D/3D Geometry	2D Planar	3D Vertical	2D Planar	3D Vertical	2D Planar	2D Planar
Switch Type	Bi	Bi	Bi	Bi	N/A	Bi
Structure	1T-1R	1R	1T-1R	1R	1T-1R	1R
Top Electrode	Al	TiN	TiN/Ti	Pt	Ir	TiN/TiO2
Material	AlOx/WOx	HfOx	HfO2	HfOx	Ta2O5/TaOx	TiO2/a-Si
Bottom Electrode	W	TiN	TiN	TiN	TaN	TiN
Cell Area (um^2)	0.0324 (.18um)	0.000003	0.01 (100nm)	1 (15nm)	0.0016(40nm)	0.0013
Speed [ns]	~1000	50	40	20	N/A	10
DC Peak Voltage [V]	N/A	<3	<3	<2.5	N/A	6.5
DC Peak Current [uA]	N/A	~500	~60	~100	N/A	5
HRS/LRS Ratio	>180	>10	10	100	10	10
R high [Ω]	$3 \cdot 10^7$	10^6	$2 \cdot 10^5$	$3 \cdot 10^7$	$5 \cdot 10^4$	10^8
R low [Ω]	$3 \cdot 10^4$	10^4	$2 \cdot 10^4$	$3 \cdot 10^4$	$4 \cdot 10^3$	10^7
Endurance	10^9	10^4	10^3	10^8	$> 10^5$	10^6
Retention	10y	20000h	10y	10y	10y	3y
Retention Temp.	85°C	250°C	70°C	RT	85°C	55°C

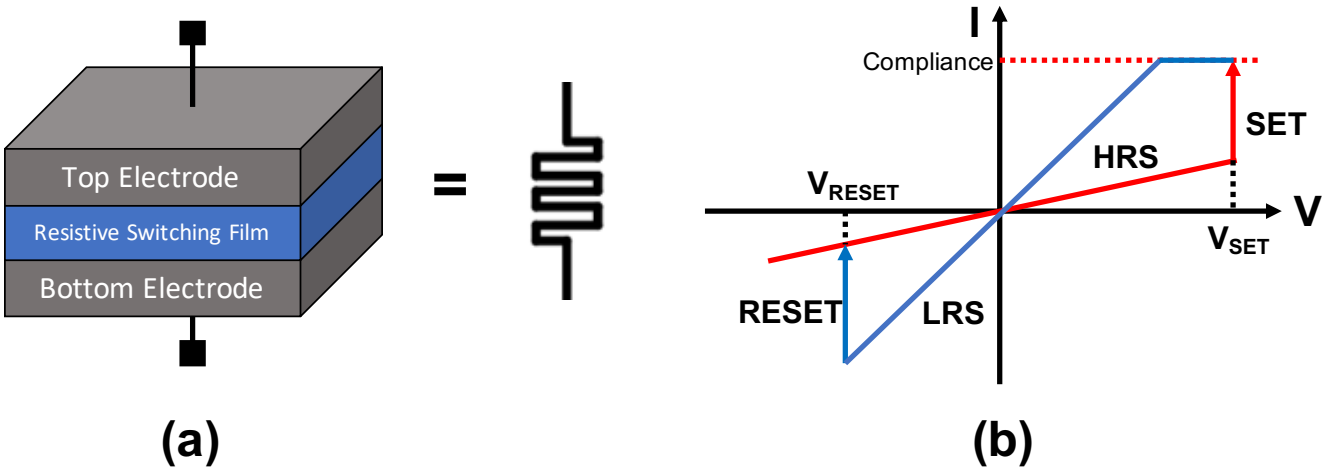


Figure 3: An RRAM device and its physical/electrical characteristics when operating in bipolar mode. (a) Device layers and electrical symbol. (b) I-V characteristic with indications of operating modes (HRS/LRS and SET/RESET).

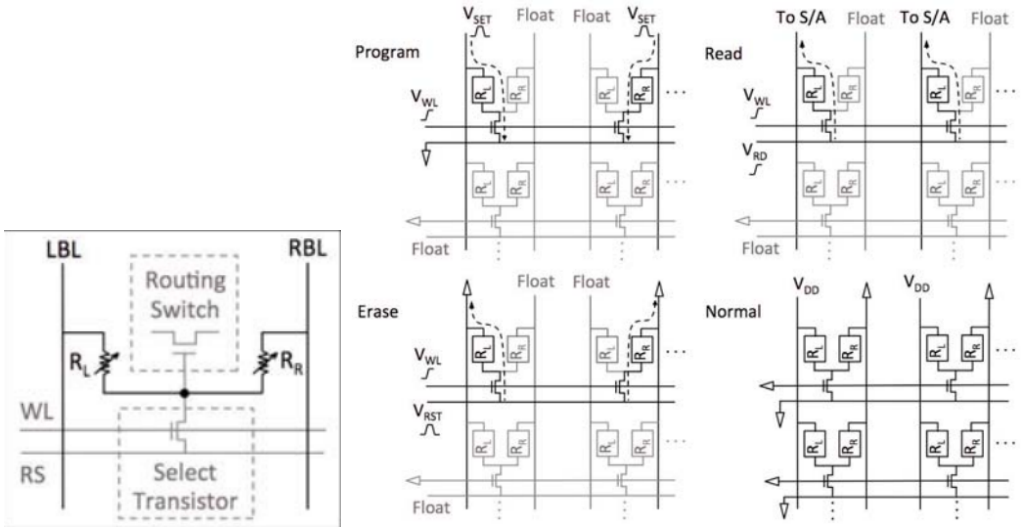


Figure 4: Adapted from [29]. (a) RRAM-based voltage divider cell design. Gray portions refer to circuit elements in the CMOS layer. (b) Schematic of four operation modes of the configuration memory.

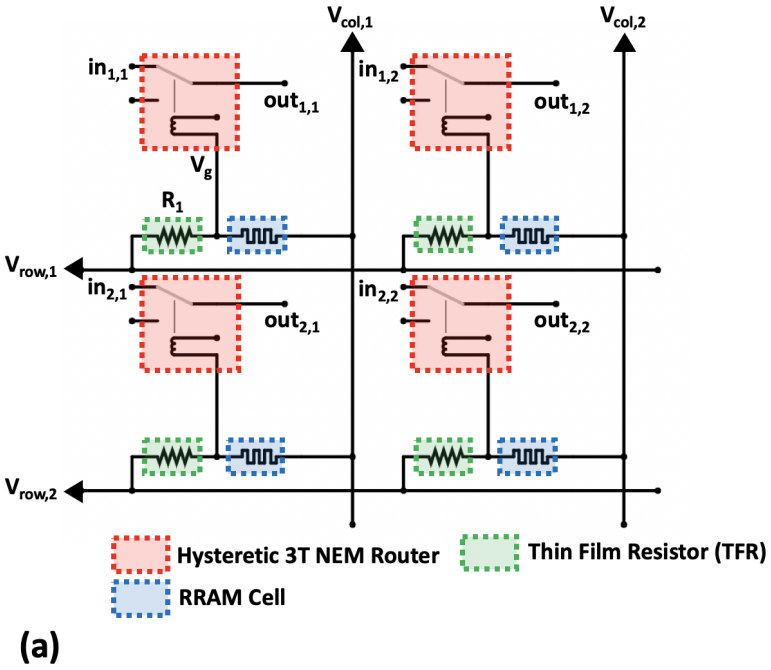
transistor (see Figure 4 for device operation) [29]. The device experimentally achieved a 57% smaller die area and a 28% lower energy-delay-product (EDP) than its SRAM-based counterpart. Additionally, the device had non-volatile configuration memory which enabled quick wake-up.

3 Circuit design

3.1 Strategy

Here, we describe a novel architecture that uses both RRAM and NEM relays in a synergistic fashion. Ultimately, the goal is to have separate physical planes for logic and routing—the front-end-of-line (FEOL) and lower metal layers should implement the reconfigurable logic, while the upper layers in back-end-of-line (BEOL) should implement the reconfigurable routing. This design strategy mimics ASICs - all routing happens in BEOL, while only CLBs are implemented in FEOL. This paradigm should allow separate place and route of routing/logic components, greatly reducing design/verification time, area overhead, and EDP. We make use of the specific device properties of both RRAM and NEM relays to enable this design—using the analog nature of the RRAM, and the sharp I-V characteristic of the NEM relay, we are able to design a 3D standard cell for a router in which the RRAM directly actuates the NEM relay with only BEOL components (no transistors).

As in [5], we use a crossbar array of hysteretic NEM relays as reconfigurable routers—we consider only 4T devices as routers since they have fewer constraints on operating voltage (being unaffected by the signals they pass) and can still be fabricated with only four masks. Instead of setting the NEM relay states in the crossbar fashion described in Figure 2, we have an RRAM cell that stores the NEM relay state even when the power supply is turned off. On startup, the RRAM cell's value is copied to the NEM relay (COPY mode), after which the NEM relay operates in hysteresis mode (HOLD mode). This eliminates the leakage current drawn by the RRAM cells in Figure 4. [PRIYANKA: A little more explanation will be good. Are we referring to the Normal mode operation? R_L and R_H are connected in series between V_{DD} and ground, so there is a static current flowing?] Our design uses a similar RRAM voltage divider approach to that taken in [29]. However, instead of having a select transistor (1T2R) design, we perform SET/RESET operations using the row/column voltages alone (2R design). [PRIYANKA: What was the point of putting the select transistor? What advantages does it have compared to the 2R design?] Our RRAM cells are oriented so that their top electrodes are in electrical contact with each other—hence, when one RRAM cell has a positive voltage across it (i.e. from its top electrode to its bottom electrode), the other cell has a negative voltage across it. [PRIYANKA: I don't understand what the previous sentence means, can we label some more nodes on Figure 5 and explain using those. Also, I am a little confused now - are we using 1 RRAM cell and 1 constant resistor or 2 RRAM cells? Can we make separate sections for the two choices, and do their analysis separately?] We do not need to worry about sneak path currents for reads, since we are only measuring the RRAM voltage locally rather than measuring the current globally. [PRIYANKA: Explain sneak paths.] However, we will need to analyze the



Mode	Param	Value
COPY	V_{row}	V_{copy}
	V_{col}	0
HOLD	V_{row}	V_{hold}
	V_{col}	Z
SET	V_{row}	$\begin{cases} V'_{SET} = \frac{V_{SET}(R_1+R_H)}{R_H}, & \text{sel} \\ Z, & \text{unsel} \end{cases}$
	V_{col}	$\begin{cases} 0, & \text{sel} \\ Z, & \text{unsel} \end{cases}$
RESET	V_{row}	$\begin{cases} V'_{RST} = \frac{V_{RST}(R_1+R_L)}{R_L}, & \text{sel} \\ Z, & \text{unsel} \end{cases}$
	V_{col}	$\begin{cases} 0, & \text{sel} \\ Z, & \text{unsel} \end{cases}$

(b)

Figure 5: (a) 2x2 switch array using RRAM-NEM hybrid design. The difference in voltage between the row and column lines is divided using a constant resistor and the RRAM cell. The result is applied at the gate of the NEM relay which enables programmable connection between *in* and *out*. (b) Row and column voltages applied for different modes of operation. Z indicates floating, sel indicates that the row/col is selected, and unsel indicates that the row/col is unselected.

write half-select problem for SET/RESET operations, since we do not want write operations to disturb unselected cells. [PRIYANKA: Explain write half-select problem.] We control the gate voltage of the NEM relay with a single RRAM cell and a constant resistor (integrated monolithically as a TFR). [PRIYANKA: Where did the section on the TFR go?]

We control V_g such that $V_g > V_{pi}$ when the RRAM cell is in the HRS, and $V_g < V_{pi}$ when the RRAM cell is in the LRS. To SET/RESET the RRAM cell, we simply apply the SET/RESET voltage to the corresponding row of that cell and ground the corresponding column. For RRAM devices which require forming, we apply a beam voltage to prevent damage to the NEM relay during the process. See Figure 5 for a visual description of the architecture and the modes of operation. [PRIYANKA: It would be good to replicate Figure 5(a) for the different modes and label the voltages and current flow, like it is done in Figure 4.]

3.2 Device parameters and knobs

Let us start by summarizing the device parameters, accounting for variability:

- $V_{pi,max}$: maximum pull-in voltage
- $V_{pi,min}$: minimum pull-in voltage
- $V_{po,max}$: maximum pull-out voltage
- $V_{po,min}$: minimum pull-out voltage
- $V_{N,max}$: absolute maximum gate voltage for NEM relay, beyond which device failure may result [PRIYANKA: Do all of the above refer to V_{gb}, g = gate, b = beam for the NEM relay?]
- $R_{L,min}$: minimum resistance of RRAM cell in low-resistance state (LRS)
- $R_{L,max}$: maximum resistance of RRAM cell in low-resistance state (LRS)

- $R_{H,min}$: minimum resistance of RRAM cell in high-resistance state (HRS)
- $R_{H,max}$: maximum resistance of RRAM cell in high-resistance state (HRS)
- $V_{SET,min}$: minimum SET voltage of RRAM (positive voltage at which RRAM goes from HRS to LRS)
- $V_{SET,max}$: maximum SET voltage of RRAM (positive voltage at which RRAM goes from HRS to LRS)
- $V_{RST,min}$: RESET voltage of RRAM with minimum magnitude (negative voltage at which RRAM goes from LRS to HRS)
- $V_{RST,max}$: RESET voltage of RRAM with maximum magnitude (negative voltage at which RRAM goes from LRS to HRS) [PRIYANKA: The above four refer to $V_{g,col}$]

Our knobs for mode 'X' are:

- $V_{row,X}$: voltage on a particular row
- $V_{col,X}$: voltage on a particular column
- $V_{beam,X}$: voltage applied to NEM relay beams [PRIYANKA: What are the voltages with respect to? Where is ground?]

We can also define some other useful intermediates (below 'X' will be replaced with the first letter of the mode name):

$$V_{L,min,X} = \frac{(V_{row,X} - V_{col,X})R_{L,min}}{R_{H,max} + R_{L,min}} \tag{1}$$

$$V_{L,max,X} = \frac{(V_{row,X} - V_{col,X})R_{L,max}}{R_{H,min} + R_{L,max}} \tag{2}$$

$$V_{H,min,X} = \frac{(V_{row,X} - V_{col,X})R_{H,min}}{R_{L,max} + R_{H,min}} \tag{3}$$

$$V_{H,max,X} = \frac{(V_{row,X} - V_{col,X})R_{H,max}}{R_{L,min} + R_{H,max}} \tag{4}$$

$$V_{I,min,X} = \frac{(V_{row,X} - V_{col,X})R_{L,min}}{R_{L,max} + R_{L,min}} \tag{5}$$

$$V_{I,max,X} = \frac{(V_{row,X} - V_{col,X})R_{L,max}}{R_{L,max} + R_{L,min}} \tag{6}$$

Above, $R_{1,max}$ and $R_{1,min}$ represent the minimum and maximum counterbalancing resistor values [PRIYANKA: These have been replaced]. $V_{L,min,X}$ and $V_{L,max,X}$ represent the bounds on voltage across the RRAM cell in LRS, and $V_{H,min,X}$ and $V_{H,max,X}$ represent the bounds on the voltage across the RRAM cell in HRS. [PRIYANKA: I get equations 1 through 4 above. What does V_I mean?]

3.3 Device constraints

We need to make sure that (1) our modes function correctly, (2) we do not accidentally SET our RRAM when in COPY mode, (3) we do not destroy our RRAM/NEM relays when doing a SET/RESET, and (4) we do not disturb unselected RRAM cells during a SET/RESET. Figure 6 provides a number line indicating the relative positions of the different voltages and represents the relationships that need to be satisfied. More explicitly, we guarantee that the device parameters satisfy the following order:

$$-V_{N,max} \leq V_{RST,max} \leq V_{RST,min} \leq 0 \leq V_{po,min} \leq V_{po,max} \leq V_{pi,min} \leq V_{pi,max} \leq V_{SET,min} \leq V_{SET,max} \leq V_{N,max} \tag{7}$$

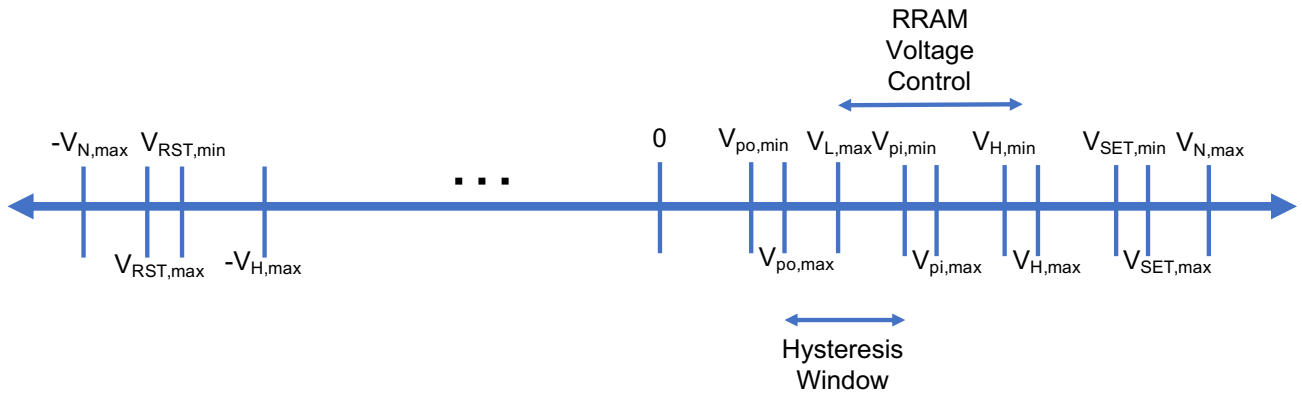


Figure 6: A number line indicating the magnitudes of different device voltage parameters relative to one another.

3.3.1 COPY mode (C) constraints

Constraint 1: In COPY mode, NEM relay must be in ON-state when RRAM is configured as such

$$V_{H,min,C} + V_{col,C} - V_{beam,C} \geq V_{pi,max} \tag{8}$$

Constraint 2: In COPY mode, NEM relay must be in OFF-state when RRAM is configured as such

$$V_{L,max,C} + V_{col,C} - V_{beam,C} \leq V_{pi,min} \tag{9}$$

Constraint 3: In COPY mode, the voltage across the RRAM should be less than the SET voltage:

$$V_{H,max,C} \leq V_{SET,min} \tag{10}$$

3.3.2 SET mode (S) constraints

In SET mode, the convention will be that the device will transition from the OFF state to the ON state. The RRAM cell connected to V_{row} will transition from R_H to R_L, and the RRAM cell connected to V_{col} will transition from R_L to R_H. The transition will happen in 2 steps. Step 1: the first RRAM cell will transition from R_H to R_L (SET) will happen first, since the voltage across the R_H is greater. After this, both RRAM cells will be in LRS. Step 2: the voltage will need to be increased until the second RRAM cell transitions from R_L to R_H (RESET). This will occur at $V_{row} - V_{col}$ equal to twice the RESET voltage.

Constraint 4: In SET mode step 1, the voltage across the HRS RRAM should be greater than the SET voltage

$$V_{H,min,S1} \geq V_{SET,max} \tag{11}$$

Constraint 5: In SET mode step 1, the absolute voltage difference between the NEM relay gate and beam before the transition should be less than the maximum gate voltage before NEM device breakdown

$$|V_{L,min,S1} + V_{col,S1} - V_{beam,S1}| \leq V_{N,max} \tag{12}$$

Constraint 6: In SET mode step 1, the absolute voltage difference between the NEM relay gate and beam after the transition should be less than the maximum gate voltage before NEM device breakdown

$$|V_{I,max,S1} + V_{col,S1} - V_{beam,S1}| \leq V_{N,max} \tag{13}$$

Constraint 7: In SET mode step 2, the voltage across the first LRS RRAM (connected to V_{row}) should be greater than the RESET voltage

$$V_{I,min,S2} \geq -V_{RST,max} \tag{14}$$

Constraint 8: In SET mode step 2, the absolute voltage difference between the NEM relay gate and beam before the transition should be less than the maximum gate voltage before NEM device breakdown

$$|V_{I,min,S2} + V_{col,S2} - V_{beam,S2}| \leq V_{N,max} \tag{15}$$

Constraint 9: In SET mode step 2, the absolute voltage difference between the NEM relay gate and beam after the transition should be less than the maximum gate voltage before NEM device breakdown

$$|V_{H,max,S2} + V_{col,S2} - V_{beam,S2}| \leq V_{N,max} \tag{16}$$

Constraint 10: Since we are using half-select programming, we need to make sure that during step 2, our half-select voltages do not cause disturbances to other cells

$$V_{H,max,S2}/2 \leq V_{SET,min} \tag{17}$$

3.3.3 RESET mode (R) constraints

RESET mode will have exactly the same constraints as SET mode by symmetry. The row and column voltages will be swapped.

3.3.4 HOLD mode (H) constraints

Constraint 11: In HOLD mode, the voltage at the gate should be between $V_{po,max}$ and $V_{pi,min}$ (V_{col} is in the high impedance state):

$$V_{po,max} \leq V_{row,H} - V_{beam,H} \leq V_{pi,min} \tag{18}$$

3.4 Tuning the knobs

To tune knobs, we need to convert the inequalities to equalities and solve for the row, column, and beam voltages for each mode.

3.4.1 COPY mode (C)

We can ground the beam in COPY mode, since we have two equations and want two unknowns. We need to focus on solving $V_{row,C}$ and $V_{col,C}$ after converting Constraints 1 and 2 to equalities.

$$V_{H,min,C} + V_{col,C} = \frac{(V_{row,C} - V_{col,C})R_{H,min}}{R_{L,max} + R_{H,min}} + V_{col,C} = V_{pi,max} \tag{19}$$

$$V_{L,max,C} + V_{col,C} = \frac{(V_{row,C} - V_{col,C})R_{L,max}}{R_{L,min} + R_{L,max}} + V_{col,C} = V_{pi,min} \tag{20}$$

Solving for $V_{row,C}$ and $V_{col,C}$ yields:

$$\boxed{V_{row,C} = \frac{R_{H,min}V_{pi,max} - R_{L,max}V_{pi,min}}{R_{H,min} - R_{L,max}}} \tag{21}$$

$$\boxed{V_{col,C} = \frac{R_{H,min}V_{pi,min} - R_{L,max}V_{pi,max}}{R_{H,min} - R_{L,max}}} \tag{22}$$

Usually Constraint 3 is satisfied, given the RRAM/NEM relay device parameters.

3.4.2 HOLD mode (H)

In this mode, we only need to control the row voltage, since the columns will be in high impedance mode. We simply set the row voltage to be anywhere in the hysteresis window, and we can arbitrarily choose it to be the middle of the window:

$$\boxed{V_{row,H} = \frac{V_{po,max} + V_{pi,min}}{2}} \tag{23}$$

3.4.3 SET/RESET mode (S/R) step 1

As before, we can ground the beam, since it does not give us more control. We can assume that with reasonable device parameters that constraints 5 and 6 will be satisfied if we set the column voltage to 0.

$$V_{H,min,S1} = \frac{(V_{row,S1} - \cancel{V_{col,S1}})R_{H,min}}{R_{L,max} + R_{H,min}} = V_{SET,max} \tag{24}$$

$$\boxed{V_{row,S1} = \frac{V_{SET,max}(R_{H,min} + R_{L,max})}{R_{H,min}}} \tag{25}$$

This means that constraints 5 and 6 become:

$$\frac{(R_{H,min} + R_{L,max})R_{L,min}V_{SET,max}}{R_{H,min}(R_{H,max} + R_{L,min})} \leq V_{N,max} \tag{26}$$

$$\frac{R_{L,max}(R_{H,min} + R_{L,max})V_{SET,max}}{R_{H,min}(R_{L,max} + R_{L,min})} \leq V_{N,max} \tag{27}$$

3.4.4 SET/RESET mode (S/R) step 2

We can again ground the beam and the column. We can assume that with reasonable device parameters that constraints 8-10 will be satisfied if we solve for only V_{row} with constraint 7.

$$V_{I,min,S2} = \frac{(V_{row,S2} - \cancel{V_{col,S2}})R_{L,min}}{R_{L,max} + R_{L,min}} = -V_{RST,max} \tag{28}$$

$$\boxed{V_{row,S2} = -\frac{V_{RST,max}(R_{L,min} + R_{L,max})}{R_{L,min}}} \tag{29}$$

This means that constraints 5 and 6 become:

$$-V_{RST,min} \leq V_{N,max} \tag{30}$$

$$-\frac{R_{H,max}(R_{L,max} + R_{L,min})V_{RST,min}}{R_{L,min}(R_{H,max} + R_{L,min})} \leq V_{N,max} \tag{31}$$

$$-\frac{R_{H,max}(R_{L,max} + R_{L,min})V_{RST,min}}{R_{L,min}(R_{H,max} + R_{L,min})} \leq 2V_{SET,min} \tag{32}$$

Equation (32) tells us that our minimum RESET voltage must be close to the SET voltage. We will need to pick RRAM with either a low HRS/LRS ratio or with a RESET voltage that is less than the SET voltage.

3.5 Row/col control simplification

The resulting row/col voltage controllers are given in Figure 7.

4 CGRA Integration

Here we describe the design flow for integration with the AHA CGRA.

4.1 RTL generation using hardware generator

Register-transfer level (RTL) is a design abstraction which models a synchronous digital circuit in terms of the flow of digital signals (data) between hardware registers, and the logical operations performed on those signals. RTL is the lowest abstraction layer and allows relatively little flexibility in terms of parametrization. This motivates the introduction of design tools that operate at a higher level of abstraction. Here, we generate the processing element (PE) tile starting from the Genesis2 code used in the first-generation AHA CGRA[1]. We use the default parameters, which generates a 16-bit-wide PE core with two adders, a counter, simple Boolean operation logic, shift operation logic, a multiplier, absolute value logic, max/min logic, an accumulator, and debugging registers. Genesis2 produces a SystemVerilog (.sv) file that is used in the next step for synthesis.

[1] `https://github.com/StanfordAHA/CGRAGenerator`

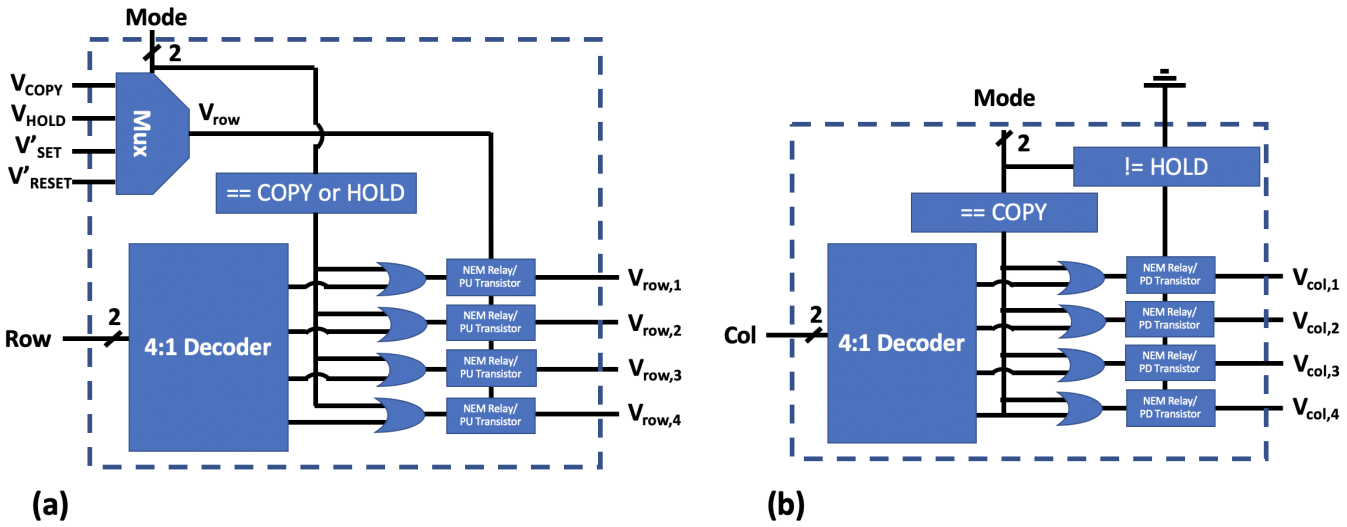


Figure 7: (a) Row voltage controller. In COPY/HOLD mode, all rows receive the same voltage. During a SET/RESET, a single row receives a voltage and the rest are floating. (b) Column voltage controller. In COPY mode, all columns are connected to ground. In HOLD mode, all columns are floating.

4.2 Synthesis

With the RTL generated using Genesis2, we perform synthesis using the Synopsys 32/28nm Generic Library (SAED32). We configure the constraints to minimize area, and leave timing/power unconstrained. This way, we can produce an estimate of how much area we should expect to have in the worst case for our routing components. The results of synthesis are below (all areas are reported in μm^2).

Number of ports:	900
Number of nets:	2621
Number of cells:	1713
Number of combinational cells:	1568
Number of sequential cells:	124
Number of macros/black boxes:	0
Number of buf/inv:	237
Number of references:	32
Combinational area:	3966.171250
Buf/Inv area:	308.022529
Noncombinational area:	1023.183769
Macro/Black Box area:	0.000000
Net Interconnect area:	1466.856481
Total cell area:	4989.355019
Total area:	6456.211500

The design checks indicate that there are some unconnected ports and shorted/constant outputs, but we ignore these warnings. The power report indicates a total power of 102 uW on low-effort analysis mode. The timing report indicates a critical path length of 8.1393 ns, which means a 100 MHz clock rate is within reach assuming the interconnect does not introduce a large amount of delay. Finally, we take our synthesized RTL as input to the PNR tool we instrument.

4.3 Place-and-Route (PNR)

For place-and-route, we use Synopsys IC Compiler. We begin by loading our design and the SAED32 libraries, including the standard cells, memory cells, and hardware models. We then specify a floorplan, which contains the

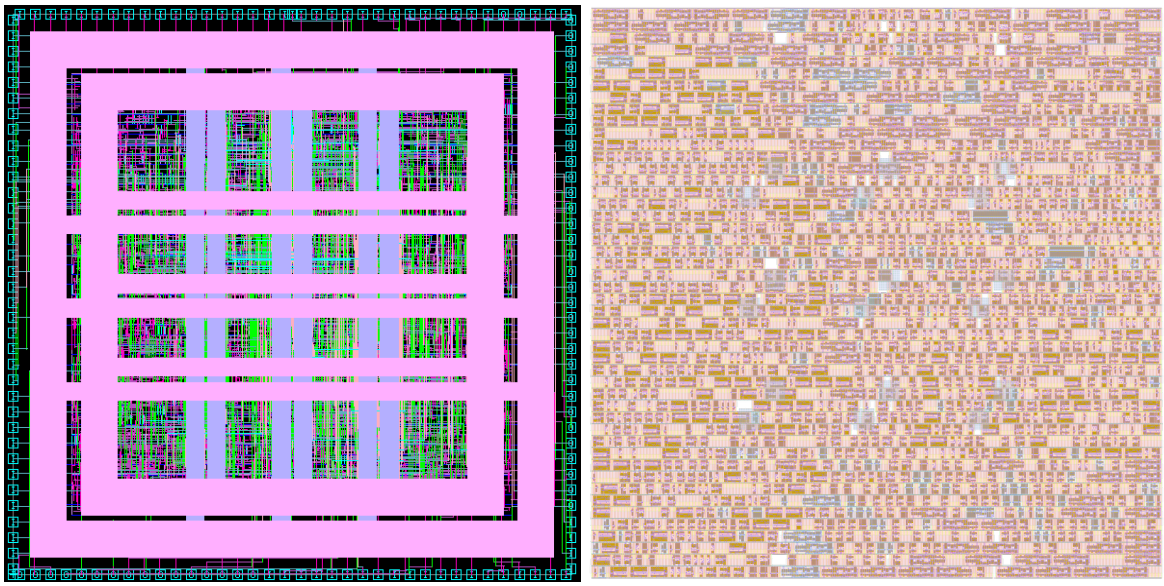


Figure 8: Rendered layouts of PE tile. (Left) View of mostly routing wires in Synopsys IC compiler. (Right) View of mostly standard cells in Virtuoso.

dimensions of the PE. We start with very loose dimensions, and then tighten them as far as possible. We create our floorplan using an aspect ratio and core utilization—also, we leave I/O margins to prevent congestion when routing to I/O pins. We start with a core utilization of 0.5 and an I/O margin of 50 um. In our second pass, we reduce the I/O margin to 10 um after unsuccessfully attempting a reduction to 5 um. In our third pass, we increase the core utilization to 0.75, and in our 4th and final pass, we further increase the core utilization to 0.9. The results of PNR are reported below (all areas are reported in μm^2).

Number of ports:	147
Number of nets:	521
Number of cells:	279
Number of combinational cells:	236
Number of sequential cells:	32
Number of macros/black boxes:	0
Number of buf/inv:	67
Number of references:	52
Combinational area:	4271.906485
Buf/Inv area:	534.718982
Noncombinational area:	1033.857812
Macro/Black Box area:	0.000000
Net Interconnect area:	1520.911512
Total cell area:	5305.764297
Total area:	6826.675809

In Figure 8 are two images of the synthesized PE tile, one from Synopsis IC Compiler, and one from Virtuoso using the GDSII streamout.

4.4 Custom technology file for RRAM/NEM relays

We modify the Synopsys techfile to incorporate the new layers we will need for the RRAM and NEM relays. Between M8 and M9, we add a layer called **RRAM**, which will include the bottom electrode (BE), resistive switching film, and the top electrode (TE). We do not need to explicitly separate layers for each of these, since the patterns for each of these layers is the same and can be processed with a single mask. Next, we introduce 5 layers for the NEM relay:

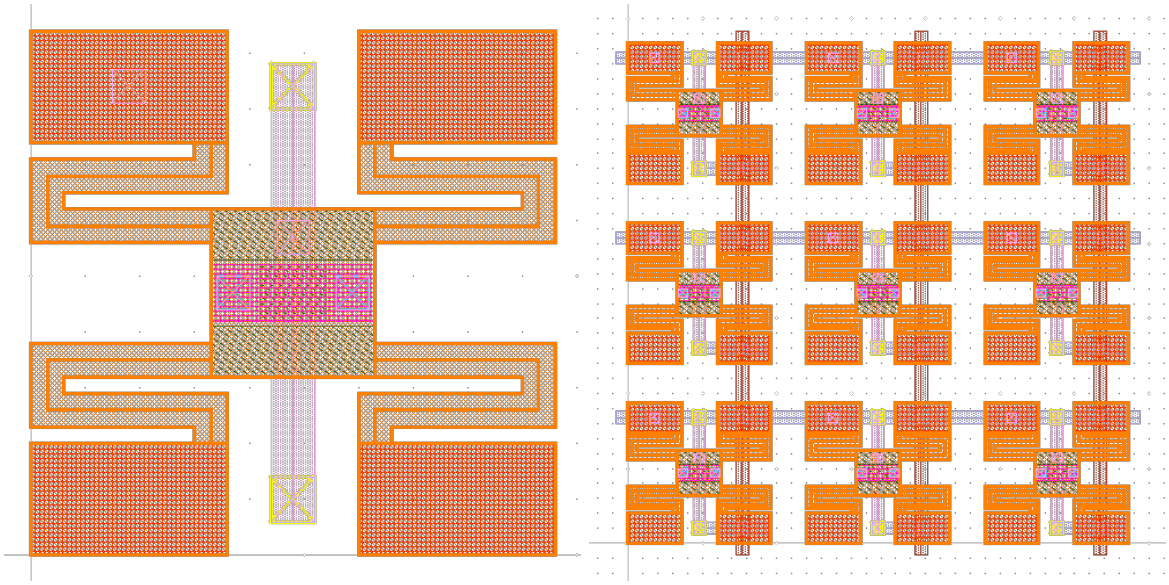


Figure 9: Rendered layouts of a single 1-bit hybrid RRAM-NEMS standard cell, as well as a crossbar array of them.

the NEM relay contact layer (**NEMCONT**), the NEM relay dimple layer (**NEMDIMP**), the NEM relay anchor layer (**NEMANC**), the NEM relay conductive channel layer (**NEMCHAN**), and finally the NEM relay beam layer (**NEMBEAM**). We correspondingly add entries for the device parameters corresponding to those reported in literature. We also update the display resource file to incorporate the new layers. Finally, we begin standard cell design of the hybrid RRAM-NEMS routers.

4.5 Standard cell design of routers

In Figure 9 are the layouts of (1) a single hybrid RRAM-NEMS standard cell, and (2) a 3x3 crossbar of standard cells. We use GDS3D to render the 3D layouts shown in Figure 10, which enables us to obtain a side/bottom view and a 3D top-view.

5 Device specification

Based on the above constraints, the target application, and the device characteristics observed in literature, we develop a list of target specifications for the RRAM and NEM relays we would need.

5.1 RRAM specification

5.1.1 Plan

Material: TiN/Ti/HfO_2/TiN (available in foundry already)
Cell area: 0.01 μm^2 (100 nm side length, should be compact enough)
Speed: $<$ 1000 ns (read speed is not that important, since only happens on startup)
SET voltage: 0.8 V (almost CMOS-compatible)
RESET voltage: -0.8 V (almost CMOS-compatible)
HRS resistance: $2 \cdot 10^5$ Ω (able to match to TFR)
LRS resistance: $2 \cdot 10^4$ Ω (able to match to TFR)
Retention: $>$ 5y at RT (approximate lifetime of PLD)
Endurance: $> 10^4$ (assuming a max of 1000 writes in lifetime, should be fine)

5.1.2 Backup 1

Material: Ir/Ta_2O_5/TaO_x/TiN
Cell area: 0.01 μm^2 (100 nm side length, should be compact enough)

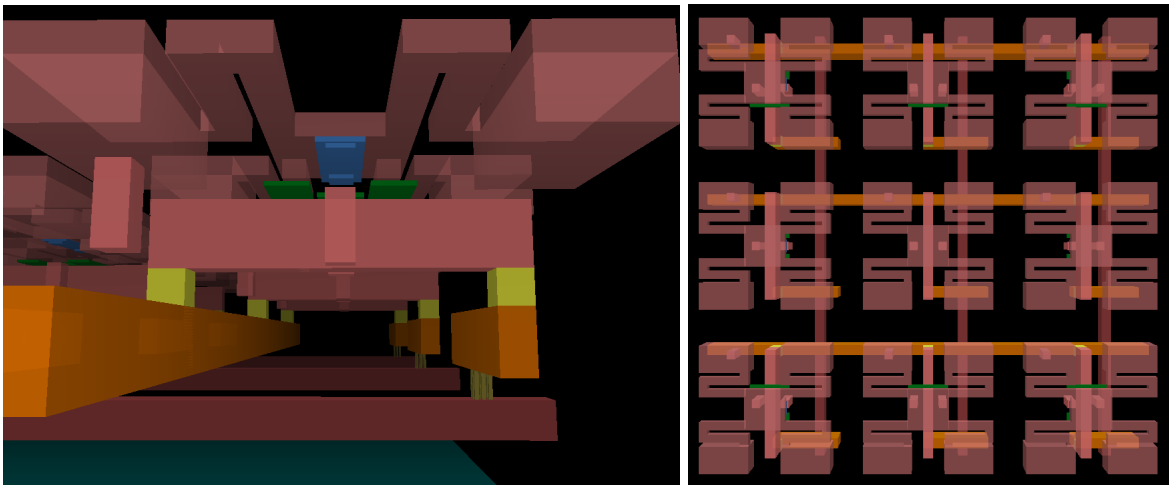


Figure 10: Rendered GDS3D layouts of a side/bottom view and a top view.

Speed: $<$ 1000 ns (read speed is not that important, since only happens on startup)
SET voltage: 6 V (NEM relay-compatible)
RESET voltage: -6 V (NEM relay-compatible)
HRS resistance: $2 \cdot 10^5$ Ω (able to match to TFR)
LRS resistance: $2 \cdot 10^4$ Ω (able to match to TFR)
Retention: $>$ 5y at RT (approximate lifetime of PLD)
Endurance: $> 10^4$ (assuming a max of 1000 writes in lifetime, should be fine)

5.2 NEM relay specification

5.2.1 Plan

Device type: 4T (can operate at CMOS levels)
Material: poly-Si or Al, Pt contact (available in foundry already)
Area: 0.1 μm^2 (1 μm x 100 nm, should be compact enough)
V_{pi}**:** 0.6 V (CMOS-compatible)
V_{po}**:** 0.1 V (CMOS-compatible)
$V_{hys,max}$**:** 0.5 V (CMOS-compatible, large enough hysteresis window)
$V_{hys,min}$**:** 0.2 V (CMOS-compatible, large enough hysteresis window)
$V_{N,max}$**:** 1.1 V (CMOS-compatible, above RRAM SET voltage)
Endurance: $> 10^7$ (assuming a max of 1 million switch-ons)
R_{ds}**:** $<$ 10 kΩ (should be much lower than transistor to get improvement)

5.2.2 Backup 1

Device type: 3T
Material: poly-Si or Al with Pt or a-C contact (BEOL compatible)
Area: 1 μm^2
V_{pi}**:** 8 V (large but still unlikely to cause problems)
V_{po}**:** 4 V (large but still unlikely to cause problems)
$V_{hys,max}$**:** 7.5 V (large but still large enough hysteresis window)
$V_{hys,min}$**:** 5 V (large but still large enough hysteresis window)
$V_{N,max}$**:** 10 V (CMOS-compatible, above RRAM SET voltage)
Endurance: $> 10^7$ (assuming a max of 1 million switch-ons)
R_{ds}**:** $<$ 30 kΩ (should be much lower than transistor to get improvement)

References

[1] I. Kuon, R. Tessier, J. Rose, *et al.*, "FPGA architecture: Survey and challenges," *Foundations and Trends in Electronic Design Automation*, vol. 2, no. 2, pp. 135–253, 2008.

[2] S. S. Wong and A. El Gamal, "The prospect of 3D-IC," in *Custom Integrated Circuits Conference, 2009. CICC'09. IEEE*, pp. 445–448, IEEE, 2009.

[3] Z. Zhang, Y. Y. Liauw, C. Chen, and S. S. Wong, "Monolithic 3-D FPGAs," *Proceedings of the IEEE*, vol. 103, no. 7, pp. 1197–1210, 2015.

[4] M. Wijtvliet, L. Waeijen, and H. Corporaal, "Coarse grained reconfigurable architectures in the past 25 years: Overview and classification," in *Embedded Computer Systems: Architectures, Modeling and Simulation (SAMOS), 2016 International Conference on*, pp. 235–244, IEEE, 2016.

[5] C. Chen, W. S. Lee, R. Parsa, S. Chong, J. Provine, J. Watt, R. T. Howe, H.-S. P. Wong, and S. Mitra, "Nano-electro-mechanical relays for FPGA routing: Experimental demonstration and a design technique," in *Design, Automation & Test in Europe Conference & Exhibition (DATE), 2012*, pp. 1361–1366, IEEE, 2012.

[6] C. Chen, R. Parsa, N. Patil, S. Chong, K. Akarvardar, J. Provine, D. Lewis, J. Watt, R. T. Howe, H.-S. P. Wong, *et al.*, "Efficient FPGAs using nanoelectromechanical relays," in *Proceedings of the 18th annual ACM/SIGDA international symposium on Field programmable gate arrays*, pp. 273–282, ACM, 2010.

[7] R. Nathanael, V. Pott, H. Kam, J. Jeon, and T.-J. K. Liu, "4-terminal relay technology for complementary logic," in *Electron Devices Meeting (IEDM), 2009 IEEE International*, pp. 1–4, IEEE, 2009.

[8] A. Peschot, C. Qian, and T.-J. Liu, "Nanoelectromechanical switches for low-power digital computing," *Micromachines*, vol. 6, no. 8, pp. 1046–1065, 2015.

[9] K. L. Harrison, W. A. Clary, J. Provine, and R. T. Howe, "Back-end-of-line compatible poly-SiGe lateral nanoelectromechanical relays with multi-level interconnect," *Microsystem Technologies*, vol. 23, no. 6, pp. 2125–2130, 2017.

[10] F. Chen, H. Kam, D. Markovic, T.-J. K. Liu, V. Stojanovic, and E. Alon, "Integrated circuit design with NEM relays," in *Proceedings of the 2008 IEEE/ACM International Conference on Computer-Aided Design*, pp. 750–757, IEEE Press, 2008.

[11] J. Jang, S. Cha, Y. Choi, G. A. Amaratunga, D. Kang, D. Hasko, J. Jung, and J. Kim, "Nanoelectromechanical switches with vertically aligned carbon nanotubes," *Applied Physics Letters*, vol. 87, no. 16, p. 163114, 2005.

[12] J. Jang, S. Cha, Y. Choi, T. Butler, D. Kang, D. Hasko, J. Jung, Y. Jin, J. Kim, and G. Amaratunga, "Nanoelectromechanical switch with low voltage drive," *Applied Physics Letters*, vol. 93, no. 11, p. 113105, 2008.

[13] W. W. Jang, J. O. Lee, J.-B. Yoon, M.-S. Kim, J.-M. Lee, S.-M. Kim, K.-H. Cho, D.-W. Kim, D. Park, and W.-S. Lee, "Fabrication and characterization of a nanoelectromechanical switch with 15-nm-thick suspension air gap," *Applied Physics Letters*, vol. 92, no. 10, p. 103110, 2008.

[14] H. Kam, V. Pott, R. Nathanael, J. Jeon, E. Alon, and T.-J. K. Liu, "Design and reliability of a micro-relay technology for zero-standby-power digital logic applications," in *2009 IEEE International Electron Devices Meeting (IEDM)*, pp. 1–4, IEEE, 2009.

[15] I.-R. Chen, L. Hutin, C. Park, R. Lee, R. Nathanael, J. Yaung, J. Jeon, and T.-J. K. Liu, "Scaled micro-relay structure with low strain gradient for reduced operating voltage," *ECS Transactions*, vol. 45, no. 6, pp. 101–106, 2012.

[16] S. Chong, K. Akarvardar, R. Parsa, J.-B. Yoon, R. T. Howe, S. Mitra, and H.-S. P. Wong, "Nanoelectromechanical (NEM) relays integrated with CMOS SRAM for improved stability and low leakage," in *Proceedings of the 2009 International Conference on Computer-Aided Design*, pp. 478–484, ACM, 2009.

[17] R. Parsa, M. Shavezipur, W. Lee, S. Chong, D. Lee, H.-S. Wong, R. Maboudian, and R. Howe, "Nanoelectromechanical relays with decoupled electrode and suspension," in *Micro electro mechanical systems (MEMS), 2011 IEEE 24th international conference on*, pp. 1361–1364, IEEE, 2011.

[18] R. Parsa, W. S. Lee, M. Shavezipur, J. Provine, R. Maboudian, S. Mitra, H.-S. P. Wong, and R. T. Howe, "Laterally actuated platinum-coated polysilicon NEM relays," *Journal of Microelectromechanical Systems*, vol. 22, no. 3, pp. 768–778, 2013.

[19] D. Grogg, C. L. Ayala, U. Drechsler, A. Sebastian, W. W. Koelmans, S. J. Bleiker, M. Fernandez-Bolanos, C. Hagleitner, M. Despont, and U. T. Duerig, "Amorphous carbon active contact layer for reliable nanoelectromechanical switches," in *Micro Electro Mechanical Systems (MEMS), 2014 IEEE 27th International Conference on*, pp. 143–146, IEEE, 2014.

[20] C. L. Ayala, D. Grogg, A. Bazigos, M. F.-B. Badia, U. T. Duerig, M. Despont, and C. Hagleitner, "A 6.7 MHz nanoelectromechanical ring oscillator using curved cantilever switches coated with amorphous carbon," in *2014 44th European Solid State Device Research Conference (ESSDERC)*, pp. 66–69, IEEE, 2014.

[21] X. Feng, M. Matheny, C. A. Zorman, M. Mehregany, and M. Roukes, "Low voltage nanoelectromechanical switches based on silicon carbide nanowires," *Nano letters*, vol. 10, no. 8, pp. 2891–2896, 2010.

[22] X. Wang, S. Narasimhan, A. Krishna, F. G. Wolff, S. Rajgopal, T.-H. Lee, M. Mehregany, and S. Bhunia, "High-temperature (¿ 500 c) reconfigurable computing using silicon carbide NEMS switches," in *Design, Automation & Test in Europe Conference & Exhibition (DATE), 2011*, pp. 1–6, IEEE, 2011.

[23] D. A. Czaplewski, G. A. Patrizi, G. M. Kraus, J. R. Wendt, C. D. Nordquist, S. L. Wolfley, M. S. Baker, and M. P. De Boer, "A nanomechanical switch for integration with CMOS logic," *Journal of Micromechanics and Microengineering*, vol. 19, no. 8, p. 085003, 2009.

[24] J. Rubin, R. Sundararaman, M. Kim, and S. Tiwari, "A single lithography vertical NEMS switch," in *Micro Electro Mechanical Systems (MEMS), 2011 IEEE 24th International Conference on*, pp. 95–98, IEEE, 2011.

[25] U. Zaghloul and G. Piazza, "10–25 nm piezoelectric nano-actuators and NEMS switches for millivolt computational logic," in *Micro Electro Mechanical Systems (MEMS), 2013 IEEE 26th International Conference on*, pp. 233–236, IEEE, 2013.

[26] C. Dong, C. Chen, S. Mitra, and D. Chen, "Architecture and performance evaluation of 3d cmos-nem fpga," in *Proceedings of the System Level Interconnect Prediction Workshop*, p. 2, IEEE Press, 2011.

[27] H.-S. P. Wong, H.-Y. Lee, S. Yu, Y.-S. Chen, Y. Wu, P.-S. Chen, B. Lee, F. T. Chen, and M.-J. Tsai, "Metal–oxide RRAM," *Proceedings of the IEEE*, vol. 100, no. 6, pp. 1951–1970, 2012.

[28] H. Wong, C. Ahn, J. Cao, H. Chen, S. Fong, Z. Jiang, C. Neumann, S. Qin, J. Sohn, Y. Wu, *et al.*, "Stanford memory trends," 2016.

[29] Y. Y. Liauw, Z. Zhang, W. Kim, A. El Gamal, and S. S. Wong, "Nonvolatile 3D-FPGA with monolithically stacked RRAM-based configuration memory," in *Solid-State Circuits Conference Digest of Technical Papers (ISSCC), 2012 IEEE International*, pp. 406–408, IEEE, 2012.