Nano-Electro-Mechanical Relays for FPGA Routing: Experimental Demonstration and a Design Technique

Chen Chen[1], W. Scott Lee[1], Roozbeh Parsa[1], Soogine Chong[1], J Provine[1], Jeff Watt[3], Roger T. Howe[1], H.-S. Philip Wong[1], Subhasish Mitra[1,2]

[1]Department of Electrical Engineering
Stanford University
Stanford, CA 94305 USA

[2]Department of Computer Science
Stanford University
Stanford, CA 94305 USA

[3]Altera Corporation
101 Innovation Drive
San Jose, CA 95134 USA

***Abstract*— Nano-Electro-Mechanical (NEM) relays are excellent candidates for programmable routing in Field Programmable Gate Arrays (FPGAs). FPGAs that combine CMOS circuits with NEM relays are referred to as CMOS-NEM FPGAs. In this paper, we experimentally demonstrate, for the first time, correct functional operation of NEM relays as programmable routing switches in FPGAs, and their programmability by utilizing hysteresis properties of NEM relays. In addition, we present a technique that utilizes electrical properties of NEM relays and selectively removes or downsizes routing buffers for designing energy-efficient CMOS-NEM FPGAs. Simulation results indicate that such CMOS-NEM FPGAs can achieve 10-fold reduction in leakage power, 2-fold reduction in dynamic power, and 2-fold reduction in area, simultaneously, without application speed penalty when compared to a 22nm CMOS-only FPGA.**

Keywords – NEM relay, FPGA routing, Half-select programming, CMOS-NEM FPGA

1. INTRODUCTION

FPGAs are popular digital design platforms because they enable low design costs and quick turnaround times [Kuon 07]. However, they suffer from several drawbacks compared to ASICs, e.g., larger area, lower performance, and higher power. These drawbacks are mainly due to the overheads associated with on-chip programmable routing, which is widely implemented using NMOS pass transistors controlled by SRAM cells [Kuon 07].

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. Unfortunately, pass transistor threshold voltage (V_t) cannot be further reduced due to leakage power constraints. Transistor gate voltage (*Vdd*) is limited by gate dielectric reliability constraints [Alam 02] and cannot be increased either (referred to as gate boosting [Betz 99]). With existing CMOS technologies, other techniques to address this challenge include the use of triple gate-oxide transistors [Altera, Xilinx] and CMOS transmission gates. These techniques introduce their own set of challenges. In this paper, we explore another alternative: Nano-Electro-Mechanical (*NEM*) relays for FPGA routing.

It has been experimentally demonstrated that NEM relays have zero off-state leakage, steep sub-threshold slope, and low on-state resistance values (R_{on}) compared to silicon CMOS transistors [Gaddi 10, Kam 09, Parsa 10]. Hence, they are promising candidates for designing highly energy-efficient digital systems. However, it is very challenging for such systems to achieve high speed and reliability for the following reasons:

- Large mechanical switching delays (>1ns) [Chen 08, 10a];
- Limited number of cycles that exhibit reliable operation (~ billions of reliable switching cycles) [Kam 09, Parsa 10].

FPGAs are a highly promising on-ramp for NEM relays because they enable unique opportunities by avoiding the above drawbacks of NEM relays while retaining their benefits [Chen 10b]:

- Since FPGA programmable routing switches do not change states after configuration, large mechanical delays of NEM relays do not affect FPGA application performance.
- NEM relays with low on-resistance values improve FPGA application critical path delays.
- Hysteresis in current-voltage (I-V) characteristics of NEM relays can be utilized to create new FPGA programmable routing switches which do not require configuration SRAM cells.
- Reliability associated with NEM relays is less of a concern for FPGAs because FPGA routing switches are generally subjected to a limited number of reconfigurations (~500) [Kuon 07].
- Using *back-end of line (BEOL)*-compatible processes [Chong 11, De Los Santos 04], NEM relays may be encapsulated and placed on top of CMOS circuits [Gaddi 10, Xie 10]. Therefore, substantial chip footprint area reduction may be obtained (Fig. 1).

The major contributions of this paper are:

- We experimentally demonstrate, for the first time, correct functional operation of programmable routing crossbars implemented using NEM relays. Hysteresis properties of NEM relays are effectively utilized to program such crossbars without requiring configuration SRAM cells. Such programming is accomplished by a special half-select programming technique tailored for NEM relays (details in Sec. 2.2 and 2.3).
- By utilizing the unique electrical properties of NEM relays, we present a new design technique for CMOS-NEM FPGAs. This technique significantly improves the energy efficiency of CMOS-NEM FPGAs compared to our earlier results in [Chen 10b]. Simulation results indicate that our new design technique can simultaneously achieve 10-fold leakage power reduction, 2-fold dynamic power reduction and 2-fold area reduction, without incurring any application speed penalty compared to a CMOS-only FPGA at the 22nm technology node (details in Sec. 3).

Section 2 introduces NEM relays, and experimentally demonstrates FPGA programmable routing crossbars using NEM relays. Section 3 presents our design technique for energy-efficient CMOS-NEM FPGAs. Related work is discussed in Sec. 4, and Sec. 5 concludes this paper.

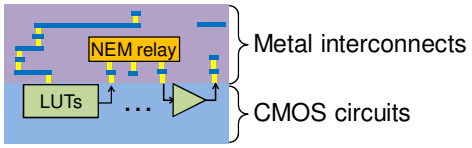


Figure 1: CMOS-NEM FPGA using NEM relays as routing switches (stacked on top of CMOS circuits).

2. NEM RELAY FOR FPGA ROUTING

We introduce NEM relays in Sec. 2.1, and present an overview of a half-select programming scheme tailored for NEM relays in Sec. 2.2. In Sec. 2.3, we experimentally demonstrate correct functional operation of 2-by-2 NEM relay-based programmable routing crossbars that can be successfully configured using our half-select programming scheme.

978-3-9810801-8-6/DATE12/©2012 EDAA

2.1 *Introduction to NEM Relays*

The structure of a NEM relay is shown in Fig. 2a. The device consists of a movable beam connected to the source electrode (*S*), a drain electrode (*D*), and a gate electrode (*G*). The voltage difference between the gate and source (V_{GS}) controls the position of the beam. When a gate to source voltage is applied, charges on the beam and gate electrodes attract each other, exerting an electrostatic force that pulls the beam toward the gate. For small V_{GS} values, the elastic force of the beam balances the electrostatic force, and the source and drain electrodes are not connected. When V_{GS} is increased to a certain voltage level, defined as the *pull-in voltage* (V_{pi}), the elastic force of the beam can no longer balance the electrostatic force exerted by the gate, and the beam pulls in toward the gate until the beam contacts the drain. Since the beam pulls in through electromechanical instability [Kaajakari 09], the V_{GS} required to release the beam, defined as the *pull-out voltage* (V_{po}), is smaller than V_{pi}, resulting in *hysteresis* in the I-V characteristics of NEM relays (Fig. 2b).

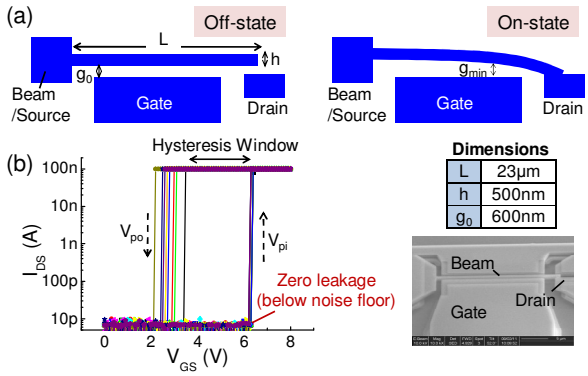


Figure 2: (a) 3-terminal (3T) NEM relay in off and on states. (b) Fabricated 3T NEM relay in our laboratory and measured I-V characteristics for multiple pull-ins and pull-outs (100nA current compliance was applied during testing).

Both V_{pi} and V_{po} are dependent on the device dimensions [Kaajakari 09]. V_{pi} can be calculated as:

$$V_{pi} = \frac{1}{L^2}\sqrt{\frac{16Eh^3 g_0^3}{81\varepsilon}},$$

where E is the Young's modulus of the beam. h and L are the thickness and length of the beam, respectively. ε is the permittivity of the ambient enclosing the relay, and g_0 is the gate-to-beam gap.

Neglecting surface forces, V_{po} can be approximated by:

$$V_{po} = \frac{1}{L^2}\sqrt{\frac{4Eh^3}{3\varepsilon} g_{\min}^2 (g_0 - g_{\min})},$$

where g_{min} is the minimum gap between gate and beam when the beam is pulled in. Actual V_{po} will be less than the estimated value obtained from the above expression because additional elastic force is required to overcome the surface forces (such as *van der Waals force*) present at the beam-drain contact. Figure 2b shows the measured I-V characteristics and the dimensions (h, L and g_0) of a fabricated NEM relay in our laboratory (using a process similar to [Parsa 10]).

Ideally, NEM relays should be operated in controlled testing environments (e.g., in vacuum or nitrogen) to avoid oxygen, moisture, and unexpected contaminants in the air. Recently, BEOL-compatible processes to seal relays in controlled ambient under micro-shells have been demonstrated [Gaddi 10, Xie 10]. Alternatively, oil can be used as a controlled ambient that limits contact corrosion, and reduces switching voltages (V_{pi} and V_{po}) due to larger permittivity (ε) of the oil [Lee 09]. The relays in this work were tested in oil to avoid environmental effects on testing without encapsulation of a controlled ambient. As confirmed by the measured characteristics, our fabricated NEM relays exhibit zero off-state leakage (below the 10pA measurement noise floor). Due to optical lithography limitations, the fabricated NEM relay has relatively large dimensions (Fig. 2b), resulting in high operation voltages (V_{pi}=6.2V, V_{po}=2~3.4V). CMOS-compatible operation voltages (~1V) can be achieved through scaling, as demonstrated both theoretically and experimentally [Akarvardar 09, Chong 11, Kam 09].

2.2 *Half-select Programming using NEM Relays*

SRAM-based CMOS FPGAs (*CMOS-only FPGAs*) use NMOS pass transistors controlled by SRAM cells to implement programmable routing (Fig. 3a). For CMOS-NEM FPGAs, one can simply replace a NMOS pass transistor with a NEM relay and use an SRAM cell to configure the state of the relay. A more beneficial approach, however, is to replace both a routing NMOS pass transistor and its corresponding SRAM cell with a single NEM relay (Fig. 3b).

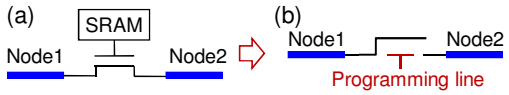


Figure 3: (a) Programmable routing element in CMOS-only FPGAs [Kuon 07]. (b) NEM relay for programmable routing.

Based on NEM relay hysteresis properties, we can apply a half-select programming scheme [Olsen 64] that is tailored for NEM relays (details in [Chen 10b]). As shown in Fig. 4, NEM relays are organized in an array with their gates connected to programming row lines and their sources connected to programming column lines. Three voltage levels {*hold voltage* (V_{hold}), *select voltage* ($-V_{select}$), and ($V_{hold}+V_{select}$)} are needed. These three voltage levels are chosen such that the following relationships are satisfied (Fig. 4):

$$V_{po}<V_{hold}<V_{pi}, V_{po}<V_{hold}+V_{select}<V_{pi}, \text{ and } V_{hold}+2V_{select}>V_{pi}.$$

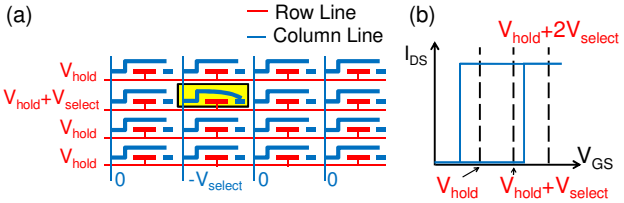


Figure 4: NEM relay half-select programming. (a) Array of relays. (b) NEM relay I-V curve with half-select programming voltages.

Initially, all relays are in pulled-out states, achieved by setting all V_{GS} voltages to 0. The half-select programming scheme is then applied to pull in the desired relays in the array in a row-by-row or column-by-column fashion. For example, to pull in the highlighted relay in Fig. 4, ($V_{hold}+V_{select}$) and $-V_{select}$ are applied to the row and column lines of the highlighted NEM relay, respectively. The remaining row and column lines are biased at V_{hold} and 0 (*GND*), respectively. Hence, only the highlighted relay will be pulled in since its V_{GS} is $V_{hold}+2V_{select}$ ($>V_{pi}$). All other NEM relays will retain their states (either pulled-in or pulled-out) since their V_{GS} values are V_{hold} or ($V_{hold}+V_{select}$), both of which are inside the hysteresis window (i.e., between V_{pi} and V_{po}). After programming, all row lines are biased at V_{hold} to retain the states of the NEM relays.

2.3 *Experimental Demonstration: NEM Relay-Based FPGA Programmable Routing Crossbar*

We demonstrate the correct functional operation of a 2-by-2 NEM relay-based programmable routing crossbar. Figure 5a shows a fabricated 2-by-2 NEM relay-based crossbar on a 4-inch wafer (we could successfully verify correct operation multiple instances of 2-by-2 programmable crossbars). The crossbar is fabricated using four

identical relays that have the same (nominal) dimensions as the relay in Fig. 2b. The crossbar can be configured using the half-select programming scheme of Sec. 2.2 with V_{hold} = 5.2V and V_{select} = 0.8V.

Figures 5b and 5c show the testing waveforms of the crossbar for two different configurations. The waveforms in Figs. 5b and 5c can be divided into three phases: programming, test, and reset. During the programming phase, we used the half-select programming technique of Sec. 2.2 to configure the desired NEM relay(s) in the array. After programming, the configured crossbar can be used as a routing network. The objective of the test phase is to verify correct configuration of the crossbar. Hence, we applied two pulses with 180^0 phase shift to the beams, and monitored the signals on the drain electrodes. After the test phase, the gate voltages were set to 0V to reset the relays to pull-out states (reset phase). By verifying the disappearance of the signals on each drain node, the previously programmed (pulled-in) relays were verified to be reset. After reset, the crossbar was re-programmed using a different configuration.

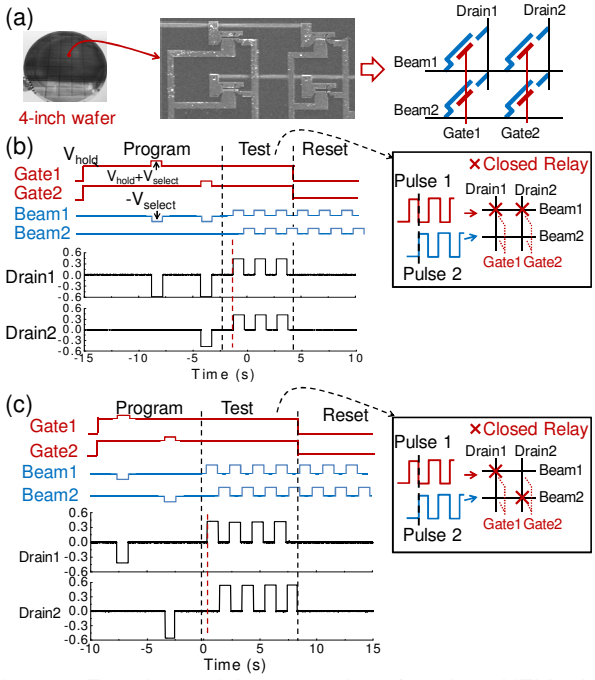


Figure 5: Experimental demonstration of a 2-by-2 NEM relay-based programmable routing crossbar: (a) SEM image of fabricated crossbar on a 4-inch wafer. (b), (c) Example waveforms (all configurations exhaustively verified).

While we could successfully demonstrate correct functional operation of 2-by-2 programmable crossbars, the measured on-resistance (R_{on}) values for the relays in the crossbar are relatively large (~100kΩ as compared to 2kΩ obtained in [Parsa 10] using similar fabrication steps). High R_{on} values are not desirable for FPGA programmable routing [Chen 10b]. Encapsulation in a low pressure, hermetic environment may help avoid surface contaminations and reduce contact resistance [Gaddi 10]. However, more work is needed to obtain low R_{on} (of the order of 2kΩ) consistently at large scale.

Correct half-select programming requires all relays in the array to be configured using the same V_{hold} and V_{select} values. This requires tight control of variations in pull-in voltages and hysteresis windows (V_{pi}-V_{po}) for a large number of NEM relays. To guarantee correct half-select programming, the minimum hysteresis window needs to be larger than the difference between the maximum and minimum pull-in voltages for all relays in the array:

$$Minimum\ \{V_{pi}\text{-}V_{po}\} > V_{pi,\,max} - V_{pi,\,min}.$$

Today's FPGAs typically contain millions of configurable routing switches. As a result, large variations can make it impossible to correctly configure all NEM relays. To examine the feasibility of building larger NEM relay-based programmable routing crossbars, we measured V_{pi} and V_{po} for 100 relays with the same dimensions as those in the crossbar array (fabricated on the same 4-inch wafer). Figure 6 shows the distributions of V_{pi} and V_{po} values. Despite V_{pi} and V_{po} variations, the required half-select programming voltage levels (V_{hold} and V_{select}) to correctly configure all tested NEM relays could still be identified (if they were organized in an array).

As indicated by Fig. 6, the noise margins (i.e., {V_{hold} to $V_{po,max}$, V_{hold}+V_{select} to $V_{pi,min}$, and V_{hold}+2V_{select} to $V_{pi,max}$}) associated with the half-select programming scheme (with the indicated programming voltages) are very small. There is a clear need to minimize variations in V_{pi} and maximize the hysteresis window to increase the yield of NEM relay crossbars. According to the equations for V_{pi} and V_{po} (Sec. 2.1), variations in V_{pi} are mostly due to variations in the dimensions of fabricated relays (such as L, h, and g_0) from our fabrication facilities. Increasing the hysteresis window requires decreasing V_{po} while maintaining V_{pi}, which could be achieved by decreasing the $g_{min}^2 \bullet (g_0 - g_{min})$ term. Furthermore, surface forces that are not accounted for also decrease V_{po}, and increase the hysteresis window.

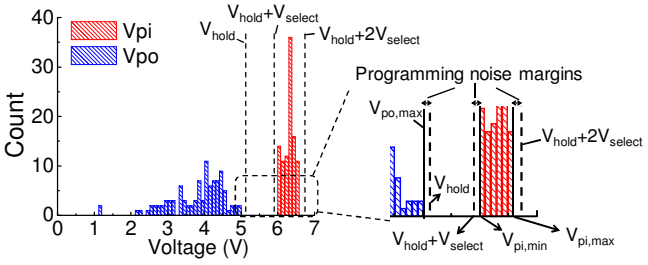


Figure 6: Distributions of V_{pi} and V_{po} for 100 identical relays.

3. CMOS-NEM FPGA DESIGN TECHNIQUE

In this section, we present a new design technique for CMOS-NEM FPGAs that can further improve the benefits predicted in [Chen 10b].

3.1 FPGA Architecture

We focus on the island-style FPGA architecture that is widely used by commercial FPGAs. It consists of an array of Logic Blocks *(LBs)* and programmable routing wires in *routing channels* that connect the LBs (Fig. 7a) [Kuon 07]. *Connection Blocks* (*CBs*) connect routing wires to LB input pins (Fig. 7c). *Switch Boxes* (*SBs*) connect LB output pins to routing wires and one set of routing wires to another set (Fig. 7d) [1].

Each LB (Fig. 7b) contains a cluster of K-input look-up tables (*K-LUTs*), where K is the number of LUT inputs. The LB *cluster size* (N) represents the number of LUTs in each LB. I is the number of LB input pins and the number of LB output pins is also N. A programmable crossbar (Fig. 7b) is used to connect LB input pins to LUT inputs, so that each LB input pin can be connected to any LUT inputs. Each LUT output can also feed back to the LUT inputs through the programmable crossbar. *LB input buffers* are used to drive the local wire interconnects and the capacitive load from the routing crossbars. At each LUT output, a 2-to-1 programmable MUX is used to select either the combinational or registered LUT output. *LB output buffers* are inserted to drive the capacitive loads from the output feedback network and LB output pins.

[1] Note that, unlike today's commercial FPGAs, our FPGA model does not consider non-reconfigurable blocks such as processor cores, signal processing units or high-speed I/O blocks.

Each routing channel consists of W routing wires, where W is defined as *routing channel width*. LB *input* (*output*) *pin flexibility*, F_{cin} (F_{cout}) is defined as the fraction of wires in the routing channel that can be connected to each LB input (output) pin through CB. For example, if F_{cin} is 0.2, each LB input pin can connect to 0.2×W wires in the routing channel. *Switch box flexibility* (F_s) represents the numbers of routing wires each routing wire can be connect to. F_s = 3 means each routing wire can connect to three other routing wires.

An FPGA can be decomposed into repeating *tiles*, where each tile consists of one *LB*, one *SB*, and two *CB*s (Fig. 7a). To achieve smaller area and better performance, routing wires in each routing channel usually span multiple tiles (referred to as *segment wires*). The *wire length* (L) denotes the number of tiles each wire spans [Kuon 07]. For example, an L=4 wire spans four FPGA tiles. Since segment wires are relatively long and have large wire capacitances, (CMOS) buffers, referred to as routing *wire buffers*, are inserted. Table 1 summarizes these architecture parameters and their values we used in our simulations [Chen 10b, Kuon 08]. We refer to LB input buffers, LB output buffers, and wire buffers collectively as *routing buffers*.

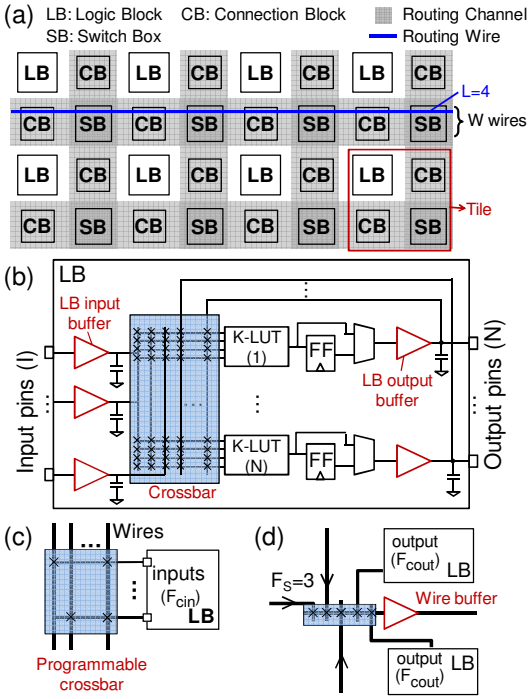


Figure 7: Island-style FPGA architecture. (a) Overall architecture. (b) Logic Block (LB). (c) Connection Block (CB). (d) Switch Box (SB).

3.2 *Selective Buffer Removal / Downsizing for CMOS-NEM FPGAs*

Traditional SRAM-based CMOS FPGAs use NMOS pass transistors controlled by SRAM cells to implement programmable routing. An NMOS pass transistor introduces a V_t drop when it passes high signal level (Fig. 8a). Hence, half latch-based buffers are used for signal restoration and for speeding up the slow rising edge (Fig. 8a). These buffers result in area, performance, and power overheads. NEM relay-based routing switches eliminate the V_t drop problem, which provides unique opportunities to "selectively" remove / downsize the corresponding routing buffers (Fig. 8b).

Figure 9 shows the breakdown of the contributions of various components (i.e., routing buffers, LUTs, etc.) to the overall dynamic and leakage power of a baseline CMOS-only FPGA (simulation details in Sec. 3.3). Routing buffers (LB input/output buffers and wire buffers) consume most of the leakage power and ~30% of the dynamic power. Selective removal / downsizing routing buffers, enabled by NEM relays, create opportunities for significantly improving the energy-efficiency of CMOS-NEM FPGAs.

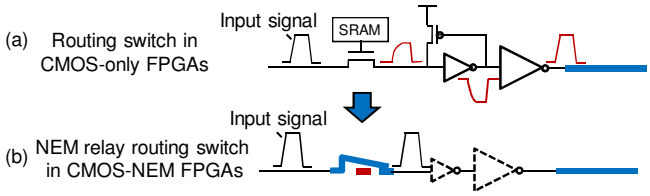


Figure 8: (a) V_t drop with NMOS pass transistor as routing switch in CMOS-only FPGAs. (b) V_t drop is eliminated by NEM relay in CMOS-NEM FPGAs.

As explained earlier (Fig. 7), there are three types of routing buffers that contribute mostly to the routing buffer power: LB input buffers, LB output buffers, and wire buffers. For CMOS-NEM FPGAs, we remove the LB input and output buffers, and downsize wire buffers (i.e., reduce widths of transistors inside wire buffers as determined by our simulations). This is due to the following reasons:

- LB input and output buffers are local buffers. They have fixed capacitive loads from local wire interconnects and the LB routing crossbar, and can be removed due to low R_{on} of NEM relays.
- Wire buffers cannot be entirely removed due to unpredictable loads (e.g., wires may be connected in series without intermediate buffers during mapping of applications onto FPGAs [Kuon 07]).

Table 1. FPGA architecture parameters.

Parameter	Description	Values
N	LUTs per LB	10
K	Inputs per LUT	4
L	Segment wire length	4
F_{cin}	LB input pin flexibility	0.2
F_{cout}	LB output pin flexibility	0.1
F_s	Switch box flexibility	3

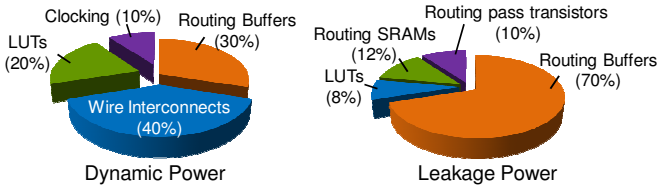


Figure 9: Dynamic and leakage power breakdown of a baseline CMOS-only FPGA.

3.3 *Simulation Methodology*

Our simulation flow is summarized in Fig. 10. With the FPGA architectural parameters in Table 1, we used the VPR tool (an FPGA place and route tool [VPR 5.0]) to estimate the minimum routing channel width (W_{min}) required for all benchmark circuits. The final routing channel width (W=118) is obtained by increasing W_{min} by 20% for "low-stress routing" [Betz 99b].

Using the routing channel width value derived using the approach explained above, we estimated the areas of the baseline CMOS-only and the CMOS-NEM FPGA tiles. In [Chen 10b], actual layouts were drawn for both CMOS-only and CMOS-NEM FPGA tiles using a commercial 90nm CMOS library to estimate layout areas, and to extract interconnect wire lengths. We used the same layout approach. NEM relays were assumed to be stacked between the metal 3 and metal 5 interconnect layers. The obtained area results were later scaled to the 22nm technology node [Chen 10b]. Wire capacitance and resistance values were calculated based on extracted wire lengths using the 22nm PTM interconnect model [Zhao 06].

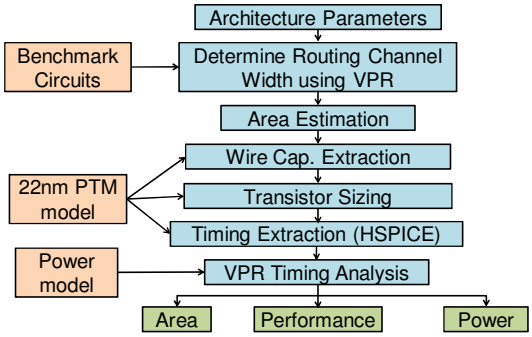


Figure 10: Simulation flow.

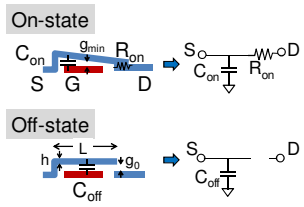
VPR requires various parameters, e.g., LUT input to output delays, LB input pin to LUT input delays, for timing analysis of each benchmark circuit mapped on the FPGA. To obtain these parameters, we created circuit netlists that represent various signal paths (e.g., LB input pin to LUT input, LUT input to output) in the target FPGA model, and then used HPSICE to simulate the netlists together with wire loads extracted from layout. For NEM relays, we used the equivalent circuit models (Fig. 11) in their on- and off-states [Chen 10b] (NEM relays will be either in on- or off-states after FPGA reconfiguration. They will not change states during normal FPGA operation). Based on experimental measurements of our fabricated devices (which have larger dimensions), we scaled the NEM relay device parameters to the 22nm technology node through simulations [Akarvardar 09, COMSOL]. The device parameters and scaled dimensions are shown in Fig. 11. For FPGA power analysis, we used an approach similar to [Jamieson 09]. This technique uses leakage power values for each circuit block and dynamic power values for each circuit node (obtained using HSPICE simulations based on the 22nm PTM transistor and wire interconnect models) and incorporates appropriate switching activities of various circuit nodes.

3.4 Simulation Results

In Fig. 12, we present results obtained for four large benchmark circuits (with > 10K equivalent 4-input LUTs) [Pistorius 07] and geometric means of the 20 largest MCNC benchmarks circuits [Yang 91]. Each circuit was mapped onto CMOS-NEM and CMOS-only FPGA models using VPR to obtain application critical path delays, and leakage and dynamic power characteristics.

Starting from a baseline CMOS-only FPGA (Sec. 3.3), we replaced NMOS routing switches and routing SRAMs with NEM relay-based programmable crossbars stacked on top of CMOS. For each segmented wire, we designed an inverter chain (with minimum-sized inverter as its first stage) to drive the capacitive load of the wire (extracted from layout). We swept the fanout of each stage (and, hence, size) of the chain to obtain the delay-optimal implementation [Weste 10]. Since segmented wire lengths are similar, all inverter chains driving segmented wires have the same size. Next, we "reduced" the size of each chain by redesigning it using the above approach while pretending that it drives a smaller capacitive load (up to 8-times smaller than the segmented wire load). This provides multiple implementations of "smaller" inverter chains with trade-offs between delay vs. power.

As shown in Fig. 12, for application critical path delays, an optimized CMOS-NEM FPGA consumes 2-fold lower dynamic power and 10-fold lower leakage power compared to the baseline CMOS-only FPGA. The footprint area of the CMOS-NEM FPGA is simultaneously reduced by 2-fold (by stacking NEM relays on top of CMOS). To quantify the benefits of our selective buffer removal / downsizing technique, we also analyzed a CMOS-NEM FPGA design which does not use our technique. For similar application critical path delays, a CMOS-NEM FPGA which does not selectively remove / downsize routing buffers achieves only 1.8-fold area reduction, 1.3-fold dynamic power reduction, and 2-fold leakage power reduction compared to the baseline CMOS-only FPGA. These benefits come from area reduction by stacking NEM relays on top of CMOS, low R_{on} values, and zero leakage of NEM relays.



Device Dimension	Value
L	275nm
h	11nm
g_0	11nm
g_{min}	3.6nm

Parameter	Value
R_{on}	2kΩ (Experimental data)
C_{on}	20aF (Simulation)
C_{off}	6.7aF (Simulation)

Figure 11: Equivalent circuits and device parameters for NEM relays in on- and off-states.

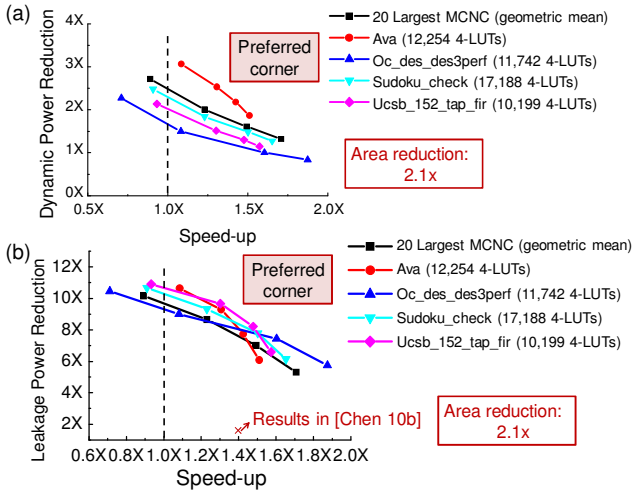


Figure 12: Power-speed trade-offs comparing CMOS-NEM FPGAs to a CMOS-only FPGA. (a) Dynamic power reduction vs. speed-up; (b) Leakage power reduction vs. speed-up.

4. RELATED WORK

Existing publications related to this paper belong to following topics: NEM relayed-based FPGAs, FPGAs using emerging devices different from NEM relays (e.g., [Cong 11, Dong 09, Paul 11]), and digital logic design using NEM relays (e.g., [Chen 08, Chen 10a, Choi 07, Chong 09, Dadgour 07, Fujita 07]). For space constraints, we focus on the first topic since it is directly related to this paper.

In our earlier work [Chen 10b], we introduced CMOS-NEM FPGAs that use NEM relays as routing switches without requiring configuration SRAM cells. This paper experimentally demonstrates correct functional operation of NEM relays (rather than relying on simulations only), and further enhances the CMOS-NEM FPGAs using our routing buffer removal / downsizing technique.

In [Zhou 07], the authors introduced a hybrid CMOS-NEM approach where carbon nanotube-based (*CNT*-based) NEM relays were used as SRAM cells inside LUTs. The authors in [Wang 10] discussed a similar idea of using *CNT*-based NEM relays as configuration memories for NMOS pass transistors. Our work differs from [Zhou 07, Wang10] in that each NEM relay in our work has the function of both a SRAM cell and a pass transistor. [Wang 11] introduced a NEM-based FPGA which can operate at high temperatures (>500^{0}C). Unlike our CMOS-NEM FPGA, NEM relays were used as logic elements. We do not use NEM relays for LUTs to avoid FPGA performance degradation due to large mechanical delays of NEM relays. A 3D CMOS-NEM FPGA was discussed in [Dong 11], where two layers of CMOS-NEM FPGAs (that are similar to our

work in [Chen 10b]) were stacked using face-to-face bonding process for further reduction in FPGA power and area. After [Chen 10b], [Sirigir 10] also introduced a similar idea of using NEM relays as routing switches in FPGAs. This paper differs from [Sirigir 10] because we experimentally demonstrate the use of hysteresis properties of NEM relays to configure the state of each FPGA routing switch. Moreover, our technique of selective removal / downsizing of FPGA routing buffers creates new opportunities for improving FPGA energy-efficiency. [Liu 08] introduced mechanical suspended-gate FETs (*SG-FETs*) that have hysteresis properties similar to NEM relays. However, unlike NEM relays, an on-state SG-FET behaves similar to an NMOS pass transistor, which still suffers from the V_t drop problem. Moreover, it may be challenging to stack such SG-FETs on top of CMOS circuits.

5. CONCLUSION

In this paper, correct functional operation of a 2-by-2 NEM relay-based FPGA programmable routing crossbar has been successfully demonstrated experimentally using hardware prototypes. We also demonstrated that the routing crossbar can be configured by utilizing hysteresis properties of NEM relays and without requiring configuration SRAM cells.

NEM relay-based FPGA routing switches do not introduce any V_t drop when passing logic signals (unlike traditional pass transistor-based FPGA routing switches). This paper utilizes this fact to remove or downsize routing buffers in CMOS-NEM FPGAs. The resulting CMOS-NEM FPGAs exhibit 2-fold lower dynamic power, 10-fold lower leakage power, and 2-fold smaller footprint without any impact on application critical path delays compared to a baseline CMOS-only FPGA at the 22nm technology node (obtained through simulations).

Future research directions include:

- Experimental demonstration of CMOS-NEM FPGAs (including integration of NEM relays on top of CMOS using CMOS back-end of line processes beyond [Gaddi 10 and Chong 11]).
- Experimental demonstrations of NEM relays with consistently small R_{on} values (<2kΩ) and small V_{pi} variations.
- Exploration of new FPGA architectures that utilize unique properties of NEM relays.

6. ACKNOWLEDGEMENT

This work was sponsored by DARPA (NBCH 1090002). The authors would like to thank DARPA program managers Dr. Akintunde I. Akinwande and Dr. Amit Lal for their support.

7. REFERENCES

[Alam 02] Alam, M., *et al.*, "A future of function or failure*?" IEEE Circuits and Device Magazine,* Vol. 18, Issue 2, pp. 42-48, 2002.

[Altera] http://www.altera.com.

[Akarvardar 09] Akarvardar, K., *et al.*, "Nanoelectromechanical Logic and Memory Devices," *ECS trans.*, Vol. 19, No. 1, pp. 49-59, 2009.

[Betz 99] Betz, V., *et al.*, "Architecture and CAD for Deep-Submicron FPGAs," *Kluwer Academic Publishers*, 1999.

[Chen 08] Chen, F., *et al*., "Integrated Circuit Design with NEM Relays," *Proc. Intl. Conf. CAD*, pp. 750-757, Nov. 2008.

[Chen 10a] Chen, F., *et al.*, "Demonstration of Integrated Micro-Electro-Mechanical Switch Circuits for VLSI Applications," *ISSCC*, pp. 150-151, 2010.

[Chen 10b] Chen, C., *et al.*, "Efficient FPGAs using Nanoelectromechanical Relays," *Intl. Symp. FPGA,* pp. 273-282, 2010.

[Choi 07] Choi, W.Y., *et al.,* "Compact Nano-Electro-Mechanical Non-Volatile Memory (NEMory) for 3D Integration," *Proc. Intl. Electron Dev. Meeting*, pp. 603-606, 2007.

[Chong 09] Chong, S., *et al.*, "Nanoelectromechanical (NEM) Relay Integrated with CMOS SRAM for Improved Stability and Low Leakage," *Proc. Intl. Conf. CAD*, pp. 478-484, 2009.

[Chong 11] Chong, S., *et al*., "Integration of Nanoelectromechanical (NEM) Relays with Silicon CMOS with Functional CMOS-NEM Circuit," *Proc. Intl. Electron Dev. Meeting*, pp. 1-4, 2011.

[COMSOL] http://www.comsol.com/

[Cong 11] Cong, J., *et al.*, "mrFPGA: A Novel FPGA Architecture with Memristor-Based Reconfiguration," *Symp. Nanoscale Architectures*, pp 1-8, 2011.

[Dadgour 07] Dadgour, H.F., *et al.*, "Design and analysis of hybrid NEMS-CMOS circuit for ultra low-power applications," *Design Automation Conference*, pp. 306-311, 2007.

[De Los Santos 04] De Los Santos, H. J., *et al.*, "RF MEMS for Ubiquitous Wireless Connectivity," *Microwave Magazine*, pp. 36-49, 2004.

[Dong 09] Dong, C., *et al.*, "FPCNA: Field Programmable Carbon Nanotube Array,"*Intl. Symp. FPGA,* pp. 161-170, 2009.

[Dong 11] Dong, C., *et al.*, "Architecture and Performance Evaluation of 3D CMOS-NEM FPGA," *System Level Interconnect Prediction,* 2011.

[Fujita 07] Fujita, S., *et al.*, "3-D Nanoarchitectures with Carbon Nanotube Mechanical Switches for Future On-Chip Network Beyond CMOS Architecture," *IEEE trans. Circuits and Systems I*, pp. 2472-2479, 2007.

[Gaddi 10] Gaddi, R., *et al.*, "MEMS technology integrated in the CMOS back end," *Microelectronics Reliability,* Vol. 50, pp. 1593-1598, 2010.

[Jamieson 09] Jamieson, P., *et al.*, "An Energy and Power Consumption Analysis of FPGA Routing Architectures", *Intl. Conf. on Field-Programmable Tech.*, pp. 324-327, 2009.

[Kaajakari 09] V. Kaajakari, *Practical MEMS, Small Gear Publishing,* 2009.

[Kam 09] Kam, H., *et al.*, "Design and Reliability of a Micro-Relay Technology for Zero-Standby Power Digital Logic Applications", *Proc. Intl. Electron Dev. Meeting,* 2009.

[Kuon 07] Kuon, I., *et al.*, "FPGA Architecture: Survey and Challenges," *Foundations and Trends in Electronic Design Automation*, Vol. 2, No. 2, pp. 135-253, 2007.

[Kuon 08] Kuon, I. and J. Rose, "Area and Delay Trade-offs in the Circuit and Architecture Design of FPGAs," *Intl. Symp. FPGA*, pp. 149-158, 2008.

[Lee 09] Lee, J.O., *et al.*, "3-terminal nanoelectromechanical switching device in insulating liquid media for low voltage operation and reliability improvement," *Proc. Intl. Electron Dev. Meeting*, pp. 1-4, 2009.

[Liu 08] Liu, M., "CMOS-Nano FPGA Utilizing Mechanical Switches," *Intl. Conf. Microelectronics*, pp 288-291, 2008.

[Olsen 64] Olsen, K.H., *et al.*, "Magnetic Core Memory," *U.S. Patent 3161861*, 1964.

[Parsa 10] R. Parsa, *et al.*, "Composite polysilicon-platinum lateral nanoelectromechanical relays," *Solid-State Sensors, Actuators, and Microsystems Workshop, Hilton Head*, pp. 7 – 10, 2010.

[Paul 11] Paul, S., *et al*., "A Circuit and Architecture Codesign Approach for a Hybrid CMOS–STTRAM Nonvolatile FPGA," *IEEE Trans. Nanotechnology,* Vol. 10, Issue 3, pp. 385-394, 2011.

[Pistorius 07] Pistorius, J., *et al.*, "Benchmarking Method and Designs Targeting Logic Synthesis for FPGAs," *Proc. Intl. Workshop Logic and Synthesis,* pp. 230-237, 2007

[Sirigir 10] Sirigir, V.K., *et al.*, "Ultra-Low-Power Ultra-Fast Hybrid CNEMS-CMOS FPGA*," Intl. Conf. Field Programmable Logic and Applications*, pp. 368-373, 2010.

[VPR 5.0] http://www.eecg.utoronto.ca/vpr/.

[Wang 10] Wang, W., *et al.*, "cFPGA: CNT emerging memory-based FPGA," *Proc. Intl. Symp. Circuits and Systems (ISCAS)*, pp. 1444-1447, 2010.

[Wang 11] Wang, X., *et al.*, "High-temperature (>500^{0}C) reconfigurable computing using silicon carbide NEMS switches," *Design Automation and Test in Europe (DATE)*, pp. 1-6, 2011.

[Weste 10] Weste, N.H.E. and D. Harris, "Principles of CMOS VLSI Design: A Circuit & System Perspectives," *Addison Wesley*, 2010.

[Xie 10] Xie, J., *et al*., "Wafer-level Vacuum Sealing and Encapsulation for Fabrication of CMOS MEMS Thermoelectric Power Generators", *Proc. Intl Conf. Micro Electro Mechanical Systems*, pp. 1175-1178, 2010.

[Xilinx] http://www.xilinx.com.

[Yang 91] Yang, S., "Logic synthesis and optimization benchmarks, version 3.0," *Technical Report MCNC*, 1991.

[Zhao 06] Zhao, W., and Y. Cao, "New generation of Predictive Technology Model for sub-45nm early design exploration," *Proc. Intl. Symp. Quality Electronic Design,* pp. 585-590, 2006.

[Zhou 07] Zhou, Y., *et al.*, "Low Power FPGA Design Using Hybrid CMOS-NEMS Approach," *Proc. Intl. Symp. Low Power Electronics and Design*, pp. 14-19, 2007.