NANO LETTERS

2009
Vol. 9, No. 10
3640-3645

Memristor−CMOS Hybrid Integrated Circuits for Reconfigurable Logic

Qiangfei Xia,*,† Warren Robinett,† Michael W. Cumbie,‡ Neel Banerjee,‡ Thomas J. Cardinali,‡ J. Joshua Yang,† Wei Wu,† Xuema Li,† William M. Tong,‡ Dmitri B. Strukov,† Gregory S. Snider,† Gilberto Medeiros-Ribeiro,† and R. Stanley Williams†

Hewlett-Packard Laboratories, 1501 Page Mill Road, Palo Alto, California 94304, and Technology Development Organization (TDO), Hewlett-Packard Company, 1000 NE Circle Blvd., Corvallis, Oregon 97330

Received June 11, 2009; Revised Manuscript Received August 20, 2009

ABSTRACT

Hybrid reconfigurable logic circuits were fabricated by integrating memristor-based crossbars onto a foundry-built CMOS (complementary metal−oxide−semiconductor) platform using nanoimprint lithography, as well as materials and processes that were compatible with the CMOS. Titanium dioxide thin-film memristors served as the configuration bits and switches in a data routing network and were connected to gate-level CMOS components that acted as logic elements, in a manner similar to a field programmable gate array. We analyzed the chips using a purpose-built testing system, and demonstrated the ability to configure individual devices, use them to wire up various logic gates and a flip-flop, and then reconfigure devices.

One possible way to extend Moore's law[1] beyond the limits of transistor scaling is to obtain the equivalent circuit functionality using fewer devices or components, i.e., get more computing per transistor on a chip. One proposal for achieving this end was the hybrid CMOL (CMOS/molecule) architecture of Strukov and Lihkarev,[2] which was modified by two of us to improve its manufacturability and separate the routing and computing functions; this was called FPNI (field-programmable nanowire interconnect).[3] Rather than relentlessly shrinking transistor sizes, FPNI separates the logic elements from the data routing network by lifting the configuration bits, routing switches, and associated components out of the CMOS layer and making them a part of the interconnect. Memristor cross bars[4,5] can be fabricated directly above the CMOS circuits, and serve as the reconfigurable data routing network. A 2D array of vias provides electrical connectivity between the CMOS and the memristor layer. Memristors are ideal for this FPGA-like application because a single device is capable of realizing functions that need several transistors in a CMOS circuit, namely, a configuration-bit flip-flop and associated data-routing multiplexer. A further advantage is that their memory function is nonvolatile, which means they do not require power to refresh their states, even if the power to the chip is turned off completely.[4] Moreover, with appropriate defect-finding and control circuitry, the redundant data paths of the cross bar structure enable alternate routes through the interconnects, resulting in a highly defect-tolerant circuit.[6,7] Numerical simulations showed that this type of architecture can dramatically increase the logic density of an FPGA-like chip without degrading power dissipation or speed even in the presence of large numbers (up to 20%) of defective components.[2,3]

In this paper, we report the successful implementation of the first memristor−CMOS hybrid integrated circuits with demonstrated FPGA-like functionality. The titanium dioxide memristor crossbars were integrated on top of a CMOS substrate using nanoimprint lithography (NIL)[8,9] and processes that did not disrupt the CMOS circuitry in the substrate. To the best of our knowledge, this is the first demonstration of NIL on an active CMOS substrate that was fabricated in a commercial semiconductor fabrication facility. The successful integration shows that memristors and the enabling NIL technology are compatible with a standard logic-type CMOS process.

The concept of our memristor−CMOS hybrid circuits is schematically shown in Figure 1a. The memristor crossbar layers (nanowire layer 1, switching layer, and nanowire layer 2) are fabricated on top of a CMOS substrate. There are two sets of tungsten vias coming up from the CMOS, one for the bottom nanowires of the crossbars (red circles in panels a and b of Figure 1) and the other for the top nanowires (blue circles). To make alignment feasible between the

* Corresponding author, qiangfei.xia@hp.com.
† Hewlett-Packard Laboratories.
‡ Technology Development Organization, Hewlett-Packard Company.

10.1021/nl901874j CCC: $40.75
Published on Web 09/01/2009

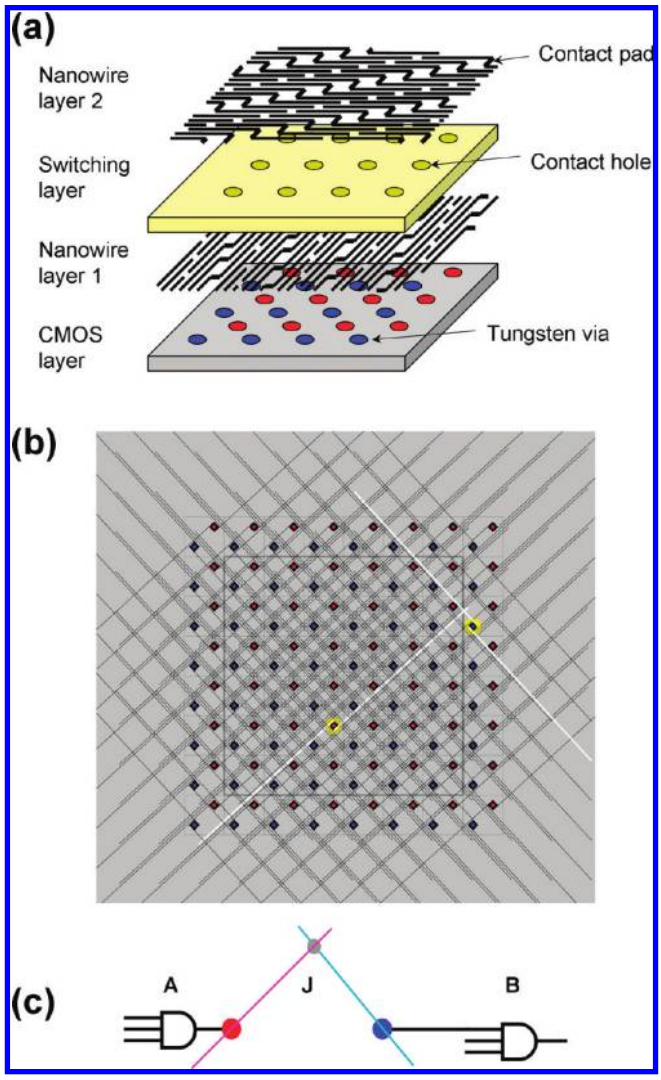


Figure 1. (a) Conceptual illustration of the memristor−CMOS hybrid architecture. Tungsten vias (red and blue) on the CMOS substrate are designed as the area interface between the transistors and the memristors. Two layers of nanowires with a switching material sandwiched between them form the memristor crossbar array. The memristors are connected to the tungsten vias (bottom electrodes to the red vias and top electrodes to the blue ones) using larger connect pads. (b) Schematic top view of the nanowires forming the crossbar. The nanowires attached to red vias run perpendicularly to the nanowires attached to blue vias, as illustrated by the two highlighted nanowires (white). A memristor is formed at each junction where two nanowires cross; if the memristor is turned ON, then an electrical connection is formed between the red via and the blue via: they are "wired" together. The pitch of the nanowires is exaggerated in this illustration, so that the wiring pattern can be seen; in the fabricated circuit, the center-to-center nanowire pitch is 200 nm, whereas the distance between same-color vias is 50 μm. (c) Schematic of a signal path in a logic circuit. A and B denote the CMOS gates and J the memristor. When the memristor is configured, the signal is passed from gate A to gate B through the metal nanowires and memristor.

CMOS and crossbar layers, interconnects between the memristor layer and the CMOS layer are implemented using larger contact pads, or "flags" that connected the nanowires to the tungsten vias in the CMOS substrate. In operation, certain memristors are configured by the CMOS circuitry to be closed, thus connecting components in the CMOS layer to synthesize particular logic circuits in FPGA fashion. Figure 1c illustrates a signal path from one gate output up through a red via, through a configured memristor, down through a blue via to another gate's input. Every signal in the configured logic circuit follows such a path, passing up and down between the CMOS level (where the gates are) and the crossbar (where the connections are defined).

We started with CMOS substrates with an active transistor plane and three metal layers fabricated in a commercial foundry using a high-voltage 0.5 μm technology. A 1 μm thick layer of silicon dioxide was deposited onto the completed wafers using tetraethyl orthosilicate (TEOS), and the resulting surface was planarized by chemical mechanical polishing (CMP) to expose the tungsten vias. This step resulted in surfaces with ~50 nm deep depressions above the tungsten vias, which were too deep to allow successful NIL. To achieve surface planarity sufficient for NIL, an extra planarization step was performed by using a thin layer of liquid NIL resist that was flattened with a polished quartz plate and cured with ultraviolet (UV) radiation. Photolithography was used to isolate the areas over the tungsten vias and expose them by etching the cured resist. These vias were extended to the TEOS surface level by metal backfilling (see Methods for fabrication details). The first layer of nanowires (100 nm width, 100 nm spacing; bottom electrodes) and connector pads were defined by NIL with the pads aligned over one set of the tungsten vias on the CMOS substrate. The entire surface of the chip was then coated with a thin layer of titanium dioxide (the switching layer) at a substrate temperature of 270 ± 25 °C, followed by exposing and extending the other set of tungsten vias to the titanium dioxide surface with photolithography, etching, and metal backfilling. The substrate temperature while depositing TiO_2 was much lower than some process temperatures during the CMOS fabrication (for example, TEOS was deposited at 450 °C using PECVD), so that it did not disrupt the underlying CMOS circuitry. The second layer of nanowires (top electrodes) was defined and deposited in a similar fashion to the bottom electrodes, with the pads aligned over the other set of tungsten vias. Finally, photolithography and etching steps were performed to open the input/output (I/O) pads on the chip's peripheral area, which conducted signals in and out of the hybrid circuits for measurements (see Methods for the measurement system). Images of the CMOS substrate and the finished hybrid chip are shown in Figure 2.

To test the electrical behavior of the memristors, an array of "isolated" memristors was fabricated using NIL on a Si wafer with a 100 nm thick thermally grown silicon dioxide. We made this array of devices to acquire the information about internal voltages that are needed to switch the devices using a four-point probe measurement system. This information is an important reference for us to choose proper rail voltages for the CMOS in the hybrid circuits in order not to accidentally change the configured memristor states. Although we have access to each memristor in the hybrid circuits by two-wire measurement, an accurate determination of the electrical characteristics is hindered by series resistance and other parasitic effects. Each memristor in the array had

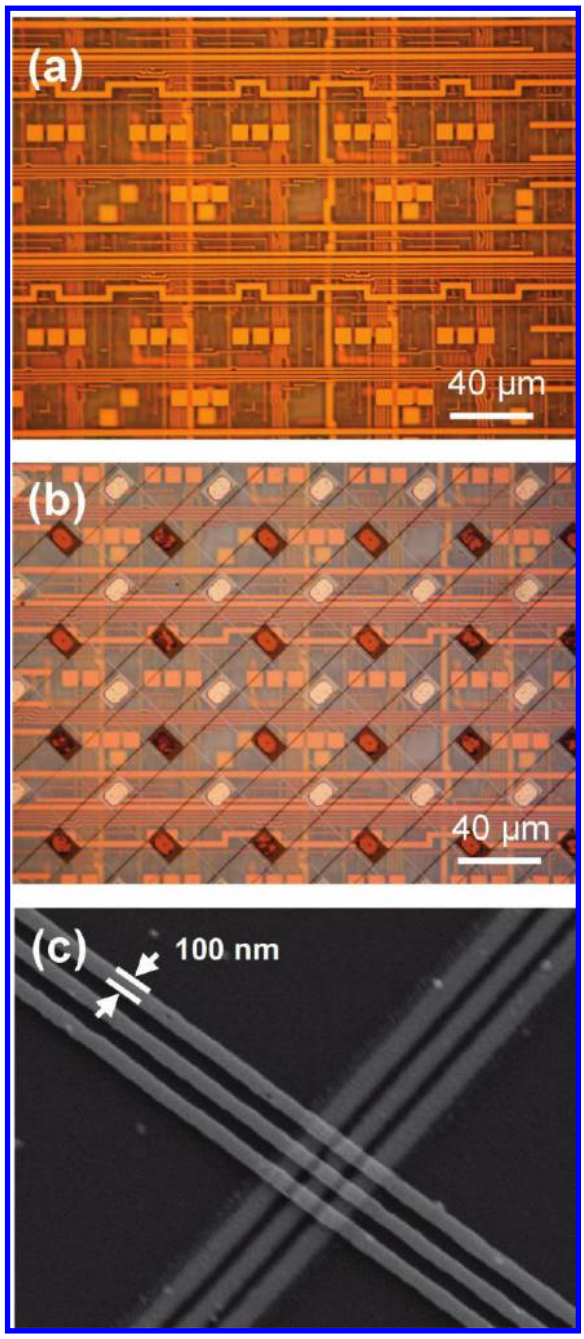


Figure 2. (a) Optical micrograph of the as-received CMOS chip. (b) The hybrid chip with memristor crossbars built on top. (c) Scanning electron microscope image of a fragment of the memristor crossbar array (where three nanowires cross three other nanowires, forming nine memristors) with junction areas of 100 × 100 nm^2.

a junction area of 100 × 100 nm^2, with a 36 nm thick titanium dioxide layer sandwiched between a 9 nm thick Pt (with 2 nm Ti adhesion layer) bottom electrode and a 12 nm thick Pt top electrode, the same device geometry as those fabricated on the CMOS substrate. It is necessary that memristors with this geometry be electrically "formed" after fabrication but prior to use as switches. In this forming operation, conductive channels are formed in the switching layer by moving oxygen vacancies toward or away from the Pt/TiO_2 interface under an applied electrical field.[5] These isolated memristors exhibit reversible nonvolatile bipolar

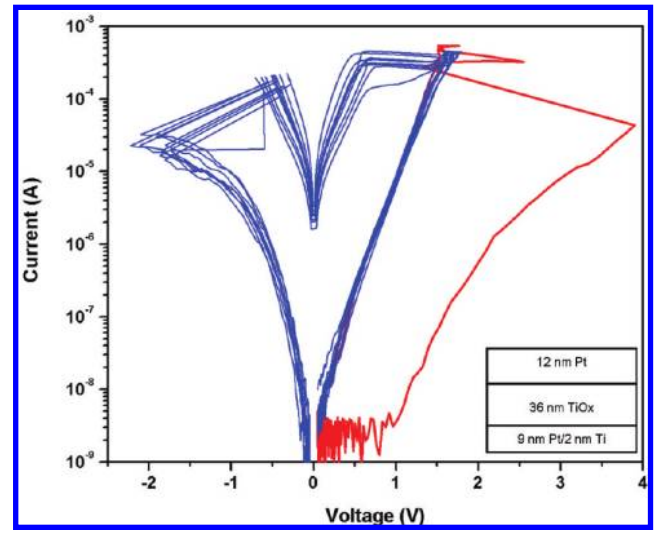


Figure 3. $I-V$ characteristic of memristors during forming and switching. These memristors were made directly on a Si wafer with a 100 nm thick thermal oxide. The cross sectional junction geometry is schematically shown in the inset, with the junction areas of 100 × 100 nm^2. These devices have the same geometry as those used in the hybrid circuits. The forming voltage is about 3.9 V, as shown by the red curve, while the subsequent 10 switching cycles after the junction is electrically formed are shown by the blue curves. The ON/OFF conductance ratio is about 1000 at 0.5 V.

switching behavior with an ON/OFF conductance ratio about 1000 at an applied voltage of 0.5 V as shown in Figure 3.

For the hybrid circuits, certain junctions are configured (turned ON) using the CMOS configuration circuitry to pass electrical signals. After the junctions are configured, output signals from one gate can be propagated to the input of another gate using lower amplitude voltages (0.5−1.7 V) in order to avoid accidentally changing the state of a memristor during logic operations. The CMOS gates internally used standard voltages for this generation, and the logic computations were done at logic voltage rails of 0 and 3.3 V. To make CMOS logic compatible with a wide range of programming and operating conditions of memristive devices, each cell has special circuitry introduced between the logic gate's input/outputs and crossbar wires. More specifically, upshift and downshift circuits were used to convert the voltages between the two sets of rails used in the CMOS and crossbar layers. For the current design, which was intended mostly as a test bed for the memristive devices/circuits development and prototyping, more than half of the area of a cell is taken by the special circuitry (nine transistors with channel widths of 2× to 16× of the minimum feature size). These circuits can be dramatically shrunk or even excluded by optimizing the CMOS design in future generations.

With several selected memristors turned on, gates on the CMOS layer can be "wired up" into digital circuits on the hybrid chip, as shown in Figure 4. For example, configuring the memristor between channel 15 and 16 would result in a NOT gate; with 4 junctions configured, the signals from input channels 98 and 92 can be sent to the transistors on the CMOS layer, and send back through to the output channel 77, completing the logic OR gate. Similarly, other basic logic

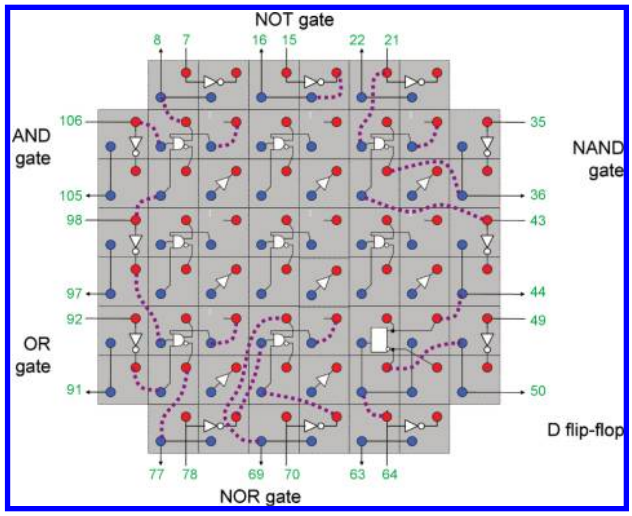
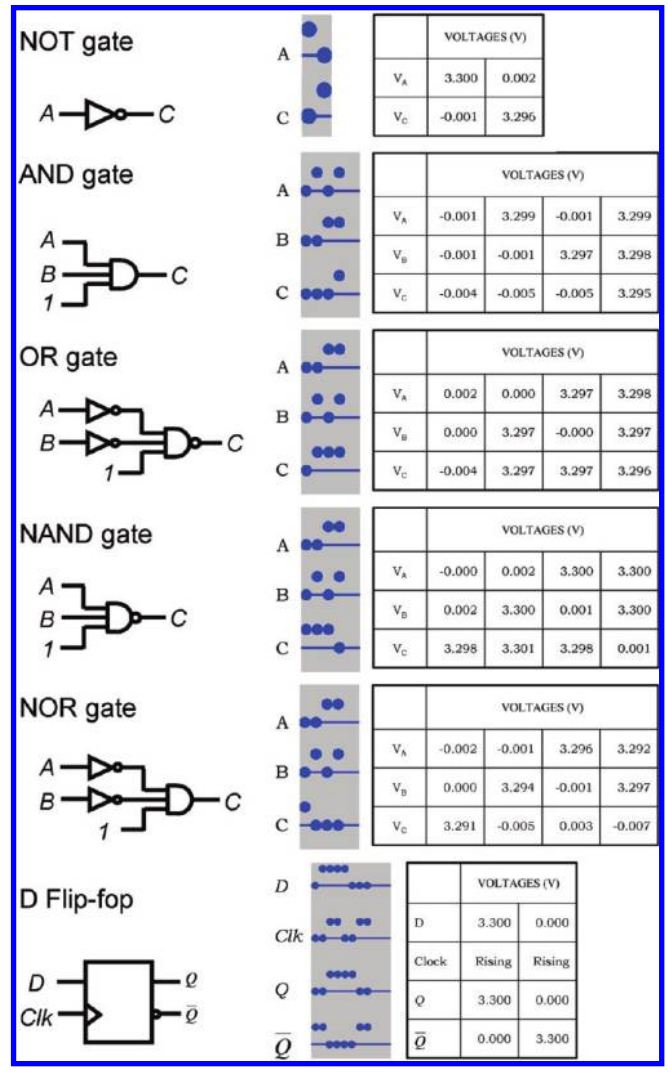


Figure 4. CMOS layer fabric on a die and possible wiring for digital circuits using memristors. The compute fabric provides three types of computing elements: logic gates (3-input AND gate with both true and complement outputs), flip-flops and repeaters. Repeaters are used to send signals long distance because the nanowire length is limited. This figure shows one of the many possible signal routings for NOT, AND, OR, NAND, and NOR gates and a D-type flip-flop. The purple dashed lines represent configured memristors in the data routing network. Red and blue dots are the tungsten vias in the CMOS layer. The green numbers are channel numbers used in the testing system. The size of each cell in this 8 × 8 cell fabric is 50 μm by 50 μm.

gates such as AND, NAND, NOR and a positive-edge triggered D flip-flop can be configured in this hybrid circuit. Since these underlying gates were designed to have three-input terminals, one of the inputs (the "spare" one) was intentionally connected to logic "1".

Although each memristor can be individually accessed using the configuration circuitry in the CMOS layer with the control software, it is possible the neighboring memristors that share the same nanowire electrode might be configured at the same time. In order to avoid simultaneously configuring multiple junctions, all other memristors are biased at half of the write voltage, which does not change the states of any memristors. The voltage drop across the designated memristor was high enough to change its state while all its neighbors in a crossbar remained at their previous states.

After the selected memristors were configured, logic functions were thereby implemented in the hybrid circuits. For all the logic gates depicted in Figure 4, satisfactory operation was achieved and the results are shown in Figure 5. In each case, an equivalent circuit and the visualization of the digital results from the test system are displayed along with the measured truth table. The demonstration of the logic operations showed that (1) the memristors were configured correctly by the CMOS circuitry, (2) the transistors in the CMOS layer were successfully connected using the configured memristors and could communicate to perform higher-level functions, and (3) most importantly, the fabrication processes for building the memristor crossbars did not disturb the underlying CMOS, which is an absolute requirement for a hybrid technology. This illustrates the possibility of implementing more complex compute circuits, which were



	VOLTAGES (V)	
V_A	3.300	0.002
V_C	-0.001	3.296

	VOLTAGES (V)			
V_A	-0.001	3.299	-0.001	3.299
V_B	-0.001	-0.001	3.297	3.298
V_C	-0.004	-0.005	-0.005	3.295

	VOLTAGES (V)			
V_A	0.002	0.000	3.297	3.298
V_B	0.000	3.297	-0.000	3.297
V_C	-0.004	3.297	3.297	3.296

	VOLTAGES (V)			
V_A	-0.000	0.002	3.300	3.300
V_B	0.002	3.300	0.001	3.300
V_C	3.298	3.301	3.298	0.001

	VOLTAGES (V)			
V_A	-0.002	-0.001	3.296	3.292
V_B	0.000	3.294	-0.001	3.297
V_C	3.291	-0.005	0.003	-0.007

	VOLTAGES (V)	
D	3.300	0.000
Clock	Rising	Rising
Q	3.300	0.000
$\bar{Q}$	0.000	3.300

Figure 5. Equivalent computing circuits, visualized digital results, and the measured truth tables for the logic gates in the hybrid circuits. In the visualization results, the lower blue dots are logic 0 and the upper dots are logic 1. The clock frequency for the flip-flop is 50 Hz.

limited by the number of available gates on the demonstration chip.

Once configured, the "wiring" between the logic gates on the hybrid chip is stored by a nonvolatile mechanism and retains its state when power is removed. We tested the NOT gate over a period of 5 months after it was first configured and found that each time it functioned properly with no observable degradation in the voltage levels delivered. We expect the retention time is much longer. This demonstrated that the memristor connections in the circuit are nonvolatile with respect to both long periods of power-down in the system and repeated cycling of the logic operation. This nonvolatility of the memristors provides a potential power advantage. In the FPGA-like circuit, the memristive data routing circuit (which is small and consumes no power to retain its configuration state) replaces the configuration bits and data-routing multiplexers of a traditional FPGA design (which is relatively larger and does consume power). In a memory application, the power advantage for a memristive

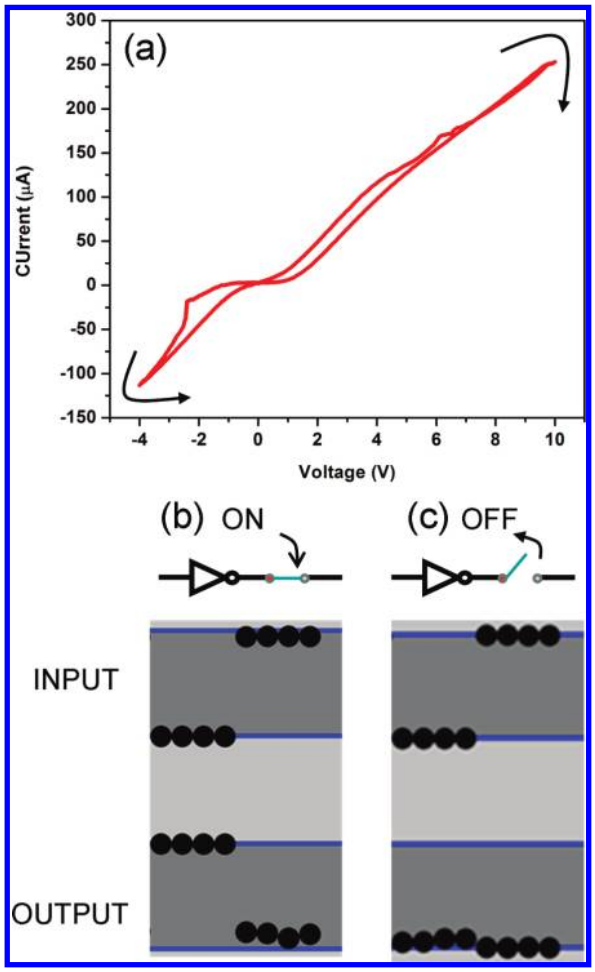


Figure 6. (a) A memristor in the hybrid circuit initially in the OFF state was turned ON using a negative voltage and then turned back OFF using a positive voltage. Note that the voltage scale is large because this is a two-wire measurement and most of the voltage dropped on the metal nanowire electrodes. (b) and (c) Computation results for a NOT gate when the connecting memristor is at different states. A string of 00001111 was inverted by the NOT gate when the memristor is ON (b), while the signal did not pass through when it is OFF (c). The reading voltage is 0.5 V.

memory cell is even clearer, since any memory cells not being accessed could be unpowered for long times.

More importantly, the memristor's states (ON or OFF) can be changed by applying the proper voltages, making the hybrid circuits reconfigurable. For example, a NOT gate was configured by turning the appropriate connecting memristor into the ON state by applying a negative write voltage on the top electrode (Figure 6). After several months of periodically testing this gate, a positive erase voltage turned the memristor back to the OFF state, so that no signal could passed through, thus decommissioning the NOT gate and making the resources available for another purpose. This demonstrated reconfigurability provides significant flexibility in data routing, and it can also reduce power consumption by disconnecting components that are not needed for a particular task.

The redundant data paths and regular structure of the crossbar architecture make it feasible to route around defective transistors, memristors, or wires, and therefore makes the hybrid circuits defect tolerant. The configuration circuitry can be used to access and test memristors, nanowires, and gates, and thus detect defective resources that can then be avoided when compiling an application onto the hardware.[6] During our first experiments to configure an OR gate, channel 91 was chosen as the output (Figure 4). However, the memristor connecting the red via of the logic gate output and the blue via of channel 91 was not working properly. Instead, we chose the blue via of channel 77 and configured the memristor that connects this blue via and the same red via of the logic gate output. As shown in Figure 5, choosing the alternative memristor resulted in successful OR-gate functionality, demonstrating the defect-tolerant capability of the current architecture.

From a circuit-fabrication point of view, our current work incorporated technologies from different generations to implement a novel architecture. The CMOS substrates we used were fabricated using a 0.5 μm technology platform, which was the prevailing technology in the early 1990s. However, the 100 nm crossbars and devices were built on top of the CMOS using NIL, a candidate for the next generation lithography (NGL) with sub-5-nm resolution, and deposition and etching procedures designed to be compatible with and not disruptive to the underlying CMOS. The gap between the two generations of technologies in terms of both feature size and implementation technology was substantial. However, our demonstration illustrates an important point: it may be possible and even necessary in the future to freeze CMOS technology at a particular feature size because of either physical or economic constraints but then continue to build on top of that platform scaled crossbars that can continue to add functionality to the hybrid system as the density of memristors increases.

The successful integration of memristors with CMOS opens up opportunities in other areas. One application is for nonvolatile random access memory (NV-RAM) integrated with logic using standard CMOS processing technology, which is compatible with the simple structures and the materials used for memristors. Another interesting application of the hybrid technology is for non-Boolean neuromorphic computing. With memristors as electronic synapses and silicon transistors as "neurons", hybrid circuits may be able to implement self-organization and learning.[10]

In summary, we have demonstrated the first hybrid integrated circuits combining foundry-produced CMOS with memristors, which were fabricated by NIL, deposition and etching procedures, and materials that were compatible with the underlying circuitry. These demonstration chips provided an FPGA-like functionality, with configured memristors defining data paths to wire logic gates into digital circuits. The memristors were reconfigurably programmed by the CMOS circuitry, thus leading to a new type of hybrid circuit architecture in which transistors and memristors are mutually complementary.

Methods. ***Device Integration.*** CMOS substrates from an HP fabrication facility were used. The CMOS layer has transistors with three metal layers that were fabricated using 0.5 μm technology. A tetraethyl orthosilicate (TEOS) oxide

layer was deposited on top of the chips upon finishing the CMOS fabrication, followed by a chemical mechanical polishing (CMP) to expose the tungsten vias. There are two sets of tungsten vias, and they are 800 nm in diameter and 50 μm apart within one set (hence a 35 μm distance between a blue and a red via because these are interstitially distributed evenly). Because of the dishing effect (the polishing rate of tungsten and TEOS are different), the exposed tungsten pins were at a lower level ($\sim$50 nm or lower) below the polished TEOS surface. To overcome this problem, an extra planarization process was usually carried out. Previously, pressing thermoplastic polymers such as PMMA above their glass transition temperatures have been used to planarize nonflat surfaces for NIL.[11,12] In our case, we used a layer of ultraviolet (UV) curable liquid material ($\sim$63 nm thick) and pressed it with a blank quartz plate to make a surface flat enough for NIL (peak to valley roughness <10 nm).[13] Tungsten vias in the CMOS were then exposed using photolithography and reactive ion etching (RIE) of the planarized layer. A double metal layer (90 nm thick Pt with 10 nm thick Ti adhesion layer) was deposited into the holes to the level of the planarized surface, followed by a liftoff process in acetone. In this manner, the tungsten vias for connecting to the bottom electrodes were extended by the double layer metal plugs.

NIL was then carried out on the planarized CMOS substrate. The master molds for NIL with nanoscale features (100 nm wide, 100 nm spacing lines of 210 μm long fanning out from grid contact pads) were first fabricated by electron beam lithography (EBL) and reactive ion etching (RIE) on a Si substrate with 50 nm thick thermal SiO_2. The size of the contact pads was 10 μm by 15 μm, and the pads were composed of grids of 400 nm pitch to assist the flow of the UV curable resist to obtain a uniform residual layer during imprinting. Daughter molds were duplicated onto optically flat quartz substrates using NIL and RIE. Quartz substrates were chosen because they are transparent to the UV light used in NIL. After the quartz molds were treated with an antisticking layer, NIL was carried out to pattern the bottom electrode on the planarized CMOS substrates with a double layer of resists (transfer layer and UV-curable resist layer on top). After the residual UV resist and the transfer layer were removed by RIE, 9 nm Pt/2 nm Ti bottom electrodes were deposited in an electron beam evaporator at ambient temperature, followed by a liftoff process in acetone. A 36 nm thick titanium dioxide layer was then sputter coated as the switching material at a substrate temperature of 270 $\pm$ 25 °C. The other set of tungsten vias for the top electrodes was exposed and extended in the same fashion, with 120 nm Pt/15 nm Ti layers deposited to extend the tungsten vias to the titanium dioxide surface level using photolithography, RIE, electron beam deposition, and liftoff. Similarly, the top electrodes (12 nm thick Pt) were fabricated using the same processes as for the bottom electrodes. A final photolithography and RIE process were carried out to open the input/output (I/O) pads in the peripheral area, which provide electrical access to the hybrid circuits.

Circuit Configuration and Measurements. The I/O pads of the hybrid chip were contacted with a custom-built probe card in a Karl-Suss PM6 probe station. These I/O pads of the hybrid chip were each connected electrically to 112 independent Source-Measurement Units (SMUs) in an electrical testing system, the PLT 1000. The voltages on these 112 channels were controlled and monitored by custom developed software, written in Java and C++ languages. This testing system supplied power and ground voltages to the chip, configuration voltages for the nanoswitches, and digital input signals to the configured compute circuits. Using its measurement capabilities, the testing system also read the digital outputs of the logic circuits.

Acknowledgment. We thank Philip J. Kuekes for helpful discussions. This work is supported in part by the US Government's Nano-Enabled Technology Initiative.

References

(1) Moore, G. E. Cramming more components onto integrated circuits. *Electron. Mag.* **1965**, *38*, 114–117.
(2) Strukov, D. B.; Likharev, K. K. CMOL FPGA: a reconfigurable architecture for hybrid digital circuits with two-terminal nanodevices. *Nanotechnology* **2005**, *16*, 888–900.
(3) Snider, G. S.; Williams, R. S. Nano/CMOS architectures using a field-programmable nanowire interconnect. *Nanotechnology* **2007**, *18*, 035204.
(4) Strukov, D. B.; Snider, G. S.; Stewart, D. R.; Williams, R. S. The missing memristor found. *Nature* **2008**, *453*, 80–83.
(5) Yang, J. J.; Pickett, M. D.; Li, X.; et al. Memristive switching mechanism for metal/oxide/metal nanodevices. *Nat. Nanotechnol.* **2008**, *3*, 429–433.
(6) Heath, J. R.; Kuekes, P. J.; Snider, G. S.; Williams, R. S. A defect-tolerant computer architecture: opportunities for nanotechnology. *Science* **1998**, *280*, 1716–1721.
(7) DeHon, A; Naeimi, H. Seven strategies for tolerating highly defective fabrication. *IEEE Design Test Comput.* **2005**, *22*, 306–315.
(8) Chou, S. Y.; Krauss, P. R.; Renstrom, P. J. Imprint of sub-25 nm vias and treches in polymers. *Appl. Phys. Lett.* **1995**, *67*, 3114–3116.
(9) Chou, S. Y.; Krauss, P. R.; Renstrom, P. J. Imprint lithography with 25-nanometer resolution. *Science* **1996**, *272*, 85–87.
(10) Snider, G. S. Self-organized computation with unreliable, memristive nanodevices. *Nanotechnology* **2007**, *18*, 365202.
(11) Sun, X.; Zhuang, L.; Zhang, W.; Chou, S. Y. Multilayer resist methods for nanoimprint lithography on nonflat surfaces. *J. Vac. Sci. Technol., B* **1998**, *16*, 3922–3925.
(12) Li, N.,Application of nanoimprint lithography in nano-electronic devices. *Ph.D. thesis*, Princeton University, 2008.
(13) Xia, Q. F.; Tong, W. M.; Wu, W.; et al. On the integration of memristors with CMOS using nanoimprint lithography. *Proc. SPIE* **2009**, *7271*, 727106.

NL901874J