A Nanowire Array for Reconfigurable Computing

Paul Beckett
School of Electrical and Computer Engineering
RMIT University
Melbourne, Australia 3000
Email: pbeckett@rmit.edu.au

Abstract—A fine-grained reconfigurable array based on complementary, dual-gate, fully depleted, silicon on insulator (DGFD-SOI) nanowire transistors is proposed and analyzed. Both low power and reconfigurable operation may be achieved by altering the switching threshold of the array using the back-gate bias on the complementary double-gate transistors. Simulated performance figures are presented for the array when configured into representative circuits and compared with two similar self-assembled molecular arrays. It is shown that SOI nanowire arrays can achieve dense, low-power reconfigurable operation without the overheads of either level restoration or additional gain blocks that may be required by molecular-based systems.

I. INTRODUCTION

If the various forecasts are to be believed [1]–[5] we are standing at the threshold of an “Aladin’s Cave” of nanoscale devices and technologies. However, the djinn that guard this cave are manufacturability, reliability and power. Unless new, disruptive technologies emerge it will become increasingly difficult to manufacture the complex, heterogeneous logic and interconnection structures that characterize current micro-architectures. This issue, combined with high defect rates and unreliable operation threaten to prevent these nanoscale riches from being fully exploited [6], [7].

Regardless of the manufacturing and reliability issues, it is most likely to be simple power dissipation considerations that will ultimately limit the density and operating frequency of any technology based on manipulating electronic charge (e.g., CMOS), even at the single-electron level [1], [8]. Static power is proportional to the average off current of the transistors \( I_{OFF} = I_D \cdot V_G = 0 \) while active dissipation may be simply expressed as switching events per unit area. Using passive cooling and economical packaging, a maximum power density in the order of 100W/cm\(^2\) holds regardless of the actual density of the switching devices. Thus, as the device density increases, both the static power per device and the percentage of devices active at a given instant must fall.

On the other hand, a more realistic power target for portable devices might be more like 0.01W/cm\(^2\). As a result, interest must eventually switch to some other state variable (e.g. electron spin) that does not rely on the transfer of charge. However, the commitment to an extensive and massively expensive silicon-based infrastructure will ensure that any successful mid-term solution will be based on CMOS-compatible technology. The issue then becomes one of how to exploit to best advantage combinations of the so-called “top-down” (conventional design) or “bottom-up” (self-assembly) fabrication approaches in a way that might be integrated into a standard CMOS line.

One response to this problem has been to look for simplified structures that offer post-manufacture reconfigurability. Such techniques allow circuits to be less sensitive to manufacturing defects, thus enhancing yields and lowering costs while at the same time offering reduced design risk and support for multiple applications via user programmability. The objective is to exploit regular or periodic structures that take optimal advantage of nanoscale fabrication techniques. As an example of this idea, molecular-based nanoscale circuits have been proposed [9]–[11] and built [12], [13] based on regular crossbar array topologies. The ability to fabricate such nanoscale circuitry on top of conventional CMOS [14] allows each level to be separately optimized [15] and the special characteristics of each to be exploited. In hybrid circuits of this type one of the primary functions of the CMOS layer is level restoration. Given that it represents an overhead in this case (i.e., it is only required because of the nanoscale layer), it is not clear at present what the optimum scale reduction (micro to nano) is likely to be (or, indeed, even how to calculate it in the general case). Further, as was shown in [14], nanoscale circuits (a nano-crossbar full adder in that case) may actually turn out to be slower, larger and consume more power than the same design in conventional CMOS once the drivers and sense-amplifiers are taken into account. This is an important result as it shows that mixed nanoscale/CMOS approaches may not necessarily exhibit better performance than conventionally scaled CMOS, unless an appropriate balance is found between the two levels.

This research is motivated by two related questions: what types of simple (regular) CMOS structures can be exploited to create reconfigurable architectures for nanocomputing and how might heterogeneous functionality emerge from an essentially homogeneous array of simple devices? Remaining in the CMOS domain offers a number of advantages, including the availability of three terminal switching devices with intrinsic gain, a stable and well characterized manufacturing base plus compatibility with existing design tools. The disadvantage is that the design is constrained by lithographic patterning and alignment issues. While it is forecast that feature sizes (for logic) will approach 22nm by 2016 or 2018 [16], it is not clear at the moment how this might be achieved. Our premise here is that simplified, regular structures with a minimal number of interconnection layers will have a better...
II. SILLICIDE S/D NANOWIRE DEVICES

In this section, we examine the characteristics of a double-gate complementary pair using silicide source/drain regions and a mid-band metal gate structure. As gate lengths are reduced, performance fluctuations due to random dopant distribution in the channel will become a major problem in CMOS. An alternative approach, employing an undoped channel region and using Schottky barriers at the source and drain has been demonstrated by a number of workers. Metal silicides form natural Schottky barriers to silicon substrates, acting to confine carriers and reducing or eliminating the need for impurities in the channel to prevent current flow in the off condition [17]. Schottky barrier devices were first described more than 30 years ago [18] and have been investigated for many years. Although they generally exhibit poorer performance than conventional devices, this gap may close as devices shrink. For example, there is evidence that quantum confinement effects in nanowires may result in greater mobility values than in bulk silicon [19], due to a reduced probability of scattering leading to ballistic operation [20] (although it appears that MOS transistors already operate surprisingly close to this limit [21]). Even if this is not the case, the fixed subthreshold slope characteristic of MOS (> 60mV/decade) will make it increasingly difficult to achieve low static power as supply voltage reduces due to the need to reduce $V_{TH}$ as well. In these highly resistive Schottky nanowire devices both $I_{ON}$ and $I_{OFF}$ are reduced in the same ratio, thereby offering low power operation, albeit at the expense of performance. Silicide nanowires therefore appear to be a viable replacement for bulk silicon channels, especially for low power systems.

In [22], both p and n-type transistors with gate lengths down to 15nm were fabricated and it was shown that leakage currents (that have been a traditional drawback of Schottky barrier FETs) could be controlled using thin-body SOI techniques. In this work, we have simulated a number of double-gate thin-body Schottky devices (Fig. 1) using a commercial simulator with classical transport models and compared the results with the general characteristics of these and other devices – in [23], for example. The n-type transistors used ErSi$_{1.7}$ source/drain regions ($\phi_{bn}=0.28$eV above Si) while the p-types were formed using PtSi ($\phi_{bp}=0.23$eV). In both cases the gate material was assumed to be Au/Cr with a (mid-band) work function of 4.7eV. The width and length were both 20nm, the body thickness was 5nm and the gate oxide thickness was 1.5nm.

Pairs of p and n-type transistors may be formed on an array of undoped silicon nanowires by depositing (and annealing) erbium and platinum silicide in alternate rows (Fig. 2). The bottom (logic) gates are common to both transistors. Conventional flash memory techniques would be applicable to set the top gate charge. Fig. 3 shows the simulated $I_D/V_G$ performance for these devices with the top gate voltages set between $\pm 2$V. It should be noted that the drift-diffusion models used here tend to underestimate the drain current density at these dimensions. For example, the value of $I_{ON} \approx 0.1\mu$A derived here can be contrasted with the actual devices measured in [22] which exhibited current drives in the order

1 Atlas/Spices, Silvaco Inc.
of 2µA for the 20nm wire in that study. In addition, it has been shown that quantum confinement effects can produce a large threshold voltage shift in double gate devices below 30nm gate lengths [24] - an effect that is not accounted for here. Neither of these effects was important to this part of the study (and in any case, the quantum V_{TH} shift is most pronounced below L_G=10nm). The results shown in Fig. 3 indicate values of I_{ON}/I_{OFF} > 10^6 can be achieved using these devices and that a significant threshold shift can be expected as the voltage on the control gate is modulated (e.g., ΔV_{TH}/ΔV_{FG} = 0.375V in Fig. 3). This is the basis of the operation of the proposed nanowire devices that can be exploited to create reconfigurable LUT structures from simple arrays, as discussed in the following section.

III. RECONFIGURABLE COMPUTING CIRCUITS

A fine-grained reconfigurable array has been described previously [25, 26] that could be said to be “polymorphic” in that its constituent blocks may be arbitrarily configured into state elements, logic, interconnect, or combinations of all three. Although there are many ways to implement such an array, the thin-body, fully depleted, double gate MOSFET devices described in Section II represent an evolutionary path from current CMOS technology. It is found that FDDGSOI devices will be ultimately scalable to gate lengths of about 10nm [27], although achieving the required level of dimensional control will be extremely difficult, as will achieving acceptable performance targets in the face of device parasitics.

In Fig. 4, complementary nanowire devices have been arranged as a 6-input, 6-output NOR-based LUT with each (horizontal) output terminated in a configurable inverter/3-state driver. A NAND-based organization would be equally possible, with a slight rearrangement of the internal connections.

To analyze the operation of this 6x6 LUT, we can start with the equation for saturation drain current in deep sub-micron (DSM) technology developed in [28]:

\[ I_D(sat) \approx \left( \frac{W}{L} \right) \left( \frac{f(R_s)}{T_{OX}} \right) (V_{GS} - V_{TH})^\alpha \]

where L is the physical gate length, T_{OX} the effective gate oxide thickness, including any depletion layer effects, and \( f(R_s) \) is a function of the source/drain resistance for a particular technology. The exponent \( \alpha \) describes the degree of velocity saturation. It is given a value of 1.25 for DSM in [28], and will tend towards 1 in short-channel, ballistic devices [29].

Equating the saturation drain currents in the conventional way [30], we can derive an approximate formula for the switching threshold for each 6NOR:

\[ V_{SW(NOR)} = \frac{V_{THN} + n^{-1}K_r^{1/\alpha}(V_{DD} + V_{THP})}{1 + n^{-1}K_r^{1/\alpha}} \]

where \( V_{THN(P)} \) is the threshold voltage for the N (P) devices, \( n \) is the number of transistors in the stack (6 in this case) and \( K_r = K_E \). Thus \( K_r \) will be a function of the transistor gain ratio—given by \( (W/\sqrt{L})_P/(W/\sqrt{L})_N \)—and the relative mobilities (replaced in (1) by the \( f(R_s)/T_{OX} \) term). It is worth noting that the effective mobility of the Schottky barrier pmos device is greater than that of the nmos due to the lower barrier height of PtSi (0.23V vs. 0.28V for ErSi1-2, see Fig 3).

It can be seen from (2) that for the symmetrical threshold case \( (V_{THN} = -V_{THP}) \), the switching threshold \( (V_{SW}) \) becomes \( V_{DD}/2 \) when \( n^{-1}K_r^{1/\alpha} = 1 \) and thus:

\[ K_r = n^\alpha. \]

It can also be seen from (2) that \( V_{SW} \geq V_{DD} \) when:

\[ V_{THN} + n^{-1}K_r^{1/\alpha}V_{THP} \geq V_{DD} \]

and \( V_{SW} \leq 0V \) when:

\[ V_{THP} + n^{-1}K_r^{1/\alpha}V_{THN} \leq -V_{DD}. \]

Equations (2) – (5) illustrate two important points about this array. Firstly, as the supply reduces with scaling, the range of threshold shifts that will be required to configure the array will also scale down. Secondly, although the optimum value of \( K_r \) (i.e., such that \( V_{SW} = V_{DD}/2 \)) is related to \( \alpha \), in common with all static gates we can adjust \( K_r \) over a fairly wide range with a minimal effect on \( V_{SW} \) (but with an effect on performance [30]).

Table I illustrates this for the devices presented in Fig. 3. In [31], it was shown by simulation that a \( \Delta V_{TH}/\Delta V_{FG} \) of +0.45V could be achieved with ultra-thin body (i.e., \( T_{body} = 5nm \)) double-gate SOI with \( T_{OX} = 1nm \). Thus, gate biases in the range \( \pm 1V \) to \( \pm 2V \) (that are also compatible with oxide reliability [32]) will be sufficient to configure the array.

Table I illustrates this for the devices presented in Fig. 3. Setting the switching thresholds of each complementary pair configures the array into its various modes. Normal operation means that the array is sensitive to that input. Setting large but symmetrical threshold values puts the array in its low standby mode (i.e., while \( V_{SW} \) is still approximately \( V_{DD} \), the propagation delay is severely affected such that the cell would not normally be operated in this mode). Shifting the switching threshold of an individual input to greater than \( V_{DD} \) effectively turns it off, such that it represents a logic “don’t care”.

An example layout of the 6x6 array is illustrated in Fig. 5. This is not necessarily the most compact form, but it does illustrate its structural regularity. In turn, these cells may be

<table>
<thead>
<tr>
<th>( V_{FGN} )</th>
<th>( V_{FGP} )</th>
<th>( V_{THN} )</th>
<th>( V_{THP} )</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>+1</td>
<td>0.45</td>
<td>-0.45</td>
<td>Normal operation</td>
</tr>
<tr>
<td>-1</td>
<td>+2</td>
<td>0.82</td>
<td>-0.82</td>
<td>Low standby power</td>
</tr>
<tr>
<td>-2</td>
<td>+3</td>
<td>1.2</td>
<td>-0.41</td>
<td>logic line disabled</td>
</tr>
<tr>
<td>+1</td>
<td>+2</td>
<td>-0.3</td>
<td>-0.82</td>
<td>unconditional low</td>
</tr>
</tbody>
</table>

TABLE I

GATE BIAS CONDITIONS FOR RECONFIGURABLE ARRAY BASED ON DEVICE CHARACTERISTICS OF FIG. 3 (\( V_{DD} = 1V \)).
organized into a larger array with adjacent connections in the vertical and horizontal directions plus a small number of local feedback connections (Fig. 6). Arranged in this way, each pair of LUT cells contains sufficient resources to develop either a small combinational logic circuit such as a 3-LUT, more complex synchronous state machine elements such as latches and flip-flops or simple asynchronous circuits (Fig. 7), as illustrated in the following sections.

A. Combinational and Sequential Logic

As an example of combinational and sequential logic synthesis, a simple D-type and a full-adder circuit were simulated using Spice3 level 10 SOI models available at the Nanotechnology Simulation Hub\(^2\). The thin-film double gate transistor models of the D-type were tuned approximately to the characteristics of the devices fabricated in [22] (with a particular focus on the sub-threshold slope) while the full-adder used parameters derived from the simulated devices outlined in section II. Fig. 8 and Fig. 9 show the results for the DFF and full-adder respectively. Table II summarizes the performance of these circuits and, as a comparison, also includes simulation results from [14] and [34] that are discussed further in Section III-C below.

\(^2\)http://nanohub.org
B. Asynchronous Logic

Current programmable systems tend not to support hazard-free logic implementations [35]. Nor do they include special functions such as arbiters and synchronizers. However, most of the building blocks for asynchronous circuits can be described in terms of small asynchronous state machines of a form that is directly supported by the nanowire array. This is illustrated in Fig. 10 for the event-controlled storage element described by Sutherland [33].

C. Power-Delay Performance

Table II summarizes the results for our nanowire array and compares it against two proposals for nanoscale systems that use “bottom-up” (i.e., partial self-assembly) approaches. The crossbar full-adder circuit of [14] is based on rectifying junctions formed from rotaxane molecules such as those described in [36]. The proposed PLA circuits in [34] are based on stochastic self-assembly techniques that form orthogonal silicon nanowire arrays. These would then be randomly connected to a micro-scale address decoder. Although the nanowire and the molecular crossbar circuits are similar in both functional complexity and in overall size, we have to exercise care when making direct comparisons between them as all of the figures listed depend on specific design details that are difficult to quantify in general terms. The thin-body MOSFETs described in [22] exhibit low values of threshold (particularly the n-type devices) due to the gate workfunction chosen. As a result, the static power \( I_{OFF}V_{DD} \) of the nanowire D-type is more than 60 times that of the molecular design. This is certainly too high to support the levels of integration envisaged. For example, if we assuming a low-power portable system with a static power target of 0.01W/cm\(^2\), the density would be power-limited to fewer than 2x10\(^4\) LUT blocks/cm\(^2\), a figure that could be easily exceeded by current low power CMOS techniques. And once the drivers/sense amplifiers are factored into the molecular crossbar case, its static power is almost double that of the high speed CMOS circuits with approximately the same propagation delay. It is only at the maximum figure for passive cooling of 100W/cm\(^2\) that we start to approach the sort of densities promised by the device roadmaps (e.g. O(2 x 10\(^3\)) blocks/cm\(^2\)). Obviously, this does not include the effect of dynamic power, which would apply additional density constraints. While no power figures are presented for the array of [34] (and the proposed manufacturing methodology is purely speculative in any case) it is based on an “nMOS-style” organization that tends to exhibit high static power. The nanowire devices, when configured “on” will be standard transistors constrained by the same \( I_{ON}, I_{OFF} \) and subthreshold slope considerations as conventional devices. As a result, the proposal as presented uses dynamic logic techniques as a way to overall minimize power in this system, although at the expense of introducing additional multi-phase clock distribution overheads.

As a final observation, it is worth noting that at the 22nm technology node all of these proposals appear to result in similar device densities. Although there may be differences in mapping efficiency when each is applied to real circuits, this is very unlikely to account in more than an order of magnitude difference in density. It could be argued that this modest return might not be worth the significant increase in manufacturing complexity required to merge conventional and self-assembled devices.

IV. Conclusion

A fine-grained reconfigurable platform based on complementary, double-gate, fully depleted silicon nanowire transistors has been proposed and analyzed. The transistors are arranged such that the top gates may be programmed to affect the switching threshold of the device and therefore its logic function. Organizing these into simple arrays supports the development of more complex computing functions. Whilst
there are still many technical challenges to be overcome, such
devices offer a number of tangible benefits, not the least of
which is a plausible migration path from conventional (micro-
scale) CMOS.

Extending CMOS into the nanoscale domain will require a
careful balance between power and performance. SOI
nanowire arrays offer low-power, complementary operation,
without the overheads of either level restoration or additional
gain blocks (as required, for example, by molecular logic
circuits). In addition, silicid source-drain techniques provide
one means by which device processing might be simplified
and thus made more tractable at nanoscale dimensions.
However, the (static) leakage power demonstrated in the devices
fabricated to date would have to be lowered by between two
and four orders of magnitude – with a resultant increase in
propagation delay – before these could provide the sort of
device riches promised in the nanoscale era.

It is becoming increasingly clear that moving to non-CMOS
technologies will not guarantee, per se, higher densities or
improved performance. Many of the gains to be made in
the nanoscale era will be achieved by exploiting innovative
architectures to compensate for the effects of poor device
performance, reliability and, in particular, power.

ACKNOWLEDGMENT

The author would like to thank Andrea Lobo of the nano-
scale PLA project group who provided the layout for Fig. 5.

REFERENCES

International Symposium on Low Power Electronics and Design,
ISLPED’02, Monterey, California, USA, 2002, pp. 172–177.
and P. Avouris, “Carbon nanotube devices for future nanoelectronics,”
in Third IEEE Conference on Nanotechnology, IEEE-NANO, 2003,
pp. 236–239.
barriers and potential solutions,” in IEEE/SEMI Advanced Semiconduc-
molecular electronics,” in Proc. 28th International Symposium on Com-
[10] A. DeHon, “Array-based architecture for PET-based, nanoscale electron-
ics,” IEEE Transactions on Nanotechnology, vol. 2, no. 1, pp. 23–32,
2003.
P. D. Franzon, and D. P. Nackashi, “Nanocell logic gates for molecular
computing,” IEEE Transactions on Nanotechnology, vol. 1, pp. 100–109,
2002.
[12] Y. Huang, X. Duan, Y. Cui, L. J. Lauhon, K.-H. Kim, and C. M.
Lieber, “Logic gates and computation from assembled nanowire building
a circuits perspective,” in International Symposium on Circuits and
[15] ——, “A case for CMOS/nano co-design,” in Proc. International Sym-
[16] SIA, International Technology Roadmap for Semiconductors, 2004 up-
date.
in Advanced Workshop on Frontiers on Electronics, WOFE ’97, 1997,
pp. 97–100.
transistor using Schottky barrier contacts for source and drain,” in
assembled using silicon nanowire building blocks,” Science, vol. 291,
nanotransistor: A simulation study.” [Online]. Available
[21] F. Assad, Z. Ren, P. Bendix, and M. S. Lundstrom, “Performance
limits of silicon MOSFET’s,” in International Electron Devices Meeting,
[22] J. Ledrizzers, P. Xuan, E. H. Anderson, J. Bokor, T.-J. King, and C. Hu,
“Complementary silicid source/drain thin-body MOSFETs for the 20
nm gate length regime,” in International Electronic Devices Meeting,
barrier MOSFETs on SOI by a self-assembly CoSi2-patterning method,”
approach describe the source-drain tunnelling in decanano double-gate
MOSFETs?” Journal of Computational Electronics, vol. 1, pp. 289–293,
2002.
gate transistors,” in Proc. IEEE International Conference on Field-
and high-speed applications,” in IEICE Transactions on Electronics, vol.
CMOS speed with gate oxide and voltage scaling and interconnect
loading effects,” IEEE Transactions on Electron Devices, vol. 44, no. 11,
[29] T. Sakurai and A. Newton, “Alpha-power law MOSFET model and its
applications to CMOS inverter delay and other formulas,” IEEE Journal
[32] D. J. Frank, R. H. Dennard, E. Nowak, P. M. Solomon, Y. Taur, and H.-
S. P. Wong, “Device scaling limits of Si MOSFETs and their application
[34] A. DeHon and M. J. Wilson, “Nanowire-based subbitigraphic pro-
grammable logic arrays,” in ACM/SIGDA 12th International Symposium on
implementing asynchronous circuits,” in IEEE Design and Test of
[36] C. P. Collier, E. W. Wong, M. Belohradsk, F. M. Raymo, J. F. Stoddart,
P. J. Kuekes, R. S. Williams, and J. R. Heath, “Electronically config-
1999.
and threshold voltage control of double-gate MOSFETs,” in
International Electron Devices Meeting, San Francisco, CA, USA, 2000,