Realizing Stochastically Assembled Nanoarrays*

John E. Savage
Department of Computer Science
Brown University
Providence, RI 02912-1910 USA

Abstract

Nanoarrays, crossbars composed of two orthogonal sets of coded parallel nanowires, are expected to play a central role in chips in the future. Given their small size, nanowires will be placed on a chip using chemical self-assembly, a stochastic process that cannot guarantee the identity or location of nanowires. A key challenge facing the use of crossbars will be controlling the uncertainty introduced by self-assembly. In this paper we explore architectural and manufacturing strategies to cope with this uncertainty. We also describe methods of analysis used to compare strategies.

1 Introduction

The end of Moore's Law [22] is in sight unless new technologies are discovered. The 2001 International Technology Roadmap for Semiconductors [1] predicts that within 10-15 years "most known technological capabilities will approach or have reached their limits." Optical lithography is the most expensive part of the chip manufacturing process and the bottleneck in increasing chip density. Unless, it can be replaced, the exponential expansion in capacity and reduction in price that has been the hallmark of the computer industry will come to an end.

Nanotechnology-based self assembly has the potential to replace optical lithographic assembly. This new technology is characterized by materials whose smallest dimensions are nanometers in length and whose assembly into chips is not completely predictable. These materials are so small that they can only be assembled stochastically. Consequently, they are much better suited to the construction of regular structures such as crossbars. (See Figure 1.) The self-assembly of nanometer-sized materials presents new design and analytical challenges.

*This research was funded in part by NSF Grant CCF-0403674. Appears in 2004 Procs. Asilomar Conf. on Signals, Systems and Computers.

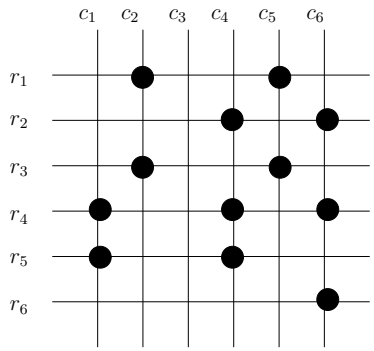

Figure 1. A crossbar in which sites holding 1s are marked by dots.

In this survey paper we outline some of the known issues and progress that has been made in coping with them.

2 Key Nanotechnologies

Key nanotechnologies that are likely to play a central role in the manufacture of chips in the future are semiconducting [4, 23, 15, 17] and metallic nanowires [21] (NWs) and carbon nanotubes [8]. Experiments have shown that they can be integrated into crossbars [26, 16, 18], two orthogonal sets of parallel wires [2, 19, 3]. When a thin molecular layer of the appropriate material is placed between the two sets of wires, at crosspoints defined by intersecting row and column wires, the material in the layer act as switches [3, 9]. Crossbars whose switches have been realized through contacts have also been exhibited [24]. In both cases their conductivity can be made high or low through the application of large positive and negative electrical fields across the crosspoints. Smaller electric fields allow the conductivity to be sensed without changing it. Crossbars can be used as memories or PLAs.

2.1 Nanowires

Nanometer-sized wires (NWs) have been grown using chemical vapor deposition (CVD) [4, 23, 15] and molecular beam epitaxy (MBE) [21, 17]. Growth of crystalline semiconducting NWs by CVD occurs under small gold clusters that are exposed to a gas in an evacuated chamber.

NWs as small as a few nanometers and as long as a few hundred microns have been grown this way. As they grow, it has also been shown that doped regions can be grown along the length of the NW [15, 13]. When these doped regions are exposed to an electric field of the appropriate magnitude, they lose a very significant fraction of their conductivity. That is, they act like field effect transistors (FETs). The doped regions can have lithographic lengths so that they can be controlled by lithographically produced wires (LWs) [5]. This provides a means to differentiate NWs and control them individually, as explained below.

Mod-doped NWs can be assembled into parallel sets of wires using fluidic methods that are akin to logs lining up as they flow downstream [26, 16, 18]. NWs are floated on the surface of a liquid, deposited on a chip and dried.

The growth of NWs by molecular beam epitaxy (MBE) is done by creating a sandwich (superlattice) of alternating layers of GaAs and AlGaAs [21]. The AlGaAs layers are partially etched to create a sawtoothed pattern along one edge. Metal is then deposited in gaps forming very long, closely-spaced NWs. NWs with width of 8 nms, pitch of 16nms, and length in tens of centimeters have been produced. The metal NWs are transferred to a chip by applying an adhesive layer and then pressing the MBE stamp against this layer. Very high quality NWs are produced using this method, known as SNAP.

SNAP metallic NWs can be used as masks to produce semiconducting NWs [17]. Silicon (Si) is deposited on silicon oxide (SiO_2) on a Si substrate. After depositing metallic wires, the Si is etched down to the SiO_2. When the metal wires are removed, Si NWs are exposed that sit on SiO_2.

2.2 Nanowire Codes

Two types of NW code have been proposed, namely, **(h, n)-hot codes** [5] and **binary-reflected codes** [10]. In (h, n)-hot codes h regions are doped and $h - n$ regions are not doped. In binary-reflected codes half of n regions are doped in in the pattern $(x_1, x_2 \ldots x_{n/2}, \bar{x}_1\bar{x}_2 \ldots \bar{x}_{n/2})$ where $x_i = 1$ denotes a doped region and $x_i = 0$ denotes and undoped region. Translation from external to internal addresses is much simpler for it than for (h, a)-hot codes.

The **code space** of the (h, n)-hot codes has $C = \binom{n}{h}$ encodings while that of the binary-reflected codes has $C = 2^{n/2}$ encodings.

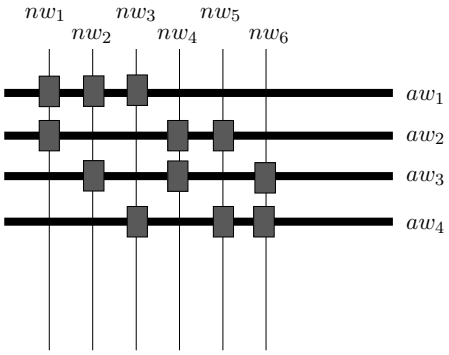

Figure 2. (2,4)-hot addressing of six NWs $\{nw_1, \ldots, nw_6\}$ with four lithographic-level address wires $\{aw_1, \ldots, aw_4\}$. The rectangles identify doped NW regions. If exactly two of the four address wires carry a high voltage, exactly one of the six NWs is conducting.

2.3 Controlling Mod-Doped NWs in a Crossbar

Electric fields are applied by LWs at right angles to the NWs, as suggested in Figure 2. If any doped region has a field near it, it loses its conductivity. Thus, with (h, n)-hot codes, if exactly h LWs carry a field strong enough to immobilize carriers in doped NW regions, exactly one type of NW will have high conductivity. With binary-reflected codes, the same statement applies when half of the LWs carry an immobilizing field. Both types of NW code have the feature that if the codes are repeated along the length of a NW and the NW is shifted relative to the LWs, another code of the same type appears beneath them. This helps to cope with NW misalignment during self-assembly. In fact, random shifting of NWs during assembly greatly reduces the number of coded NWs that need to be manufactured.

2.4 Stochastic Assembly of Mod-Doped NW Crossbars

Mod-doped NWs are assembled into crossbars by first creating a mixture of differently encoded NWs and then depositing a randomly selected sample of NWs on a chip [26, 16, 18]. The NWs will align in parallel, although not perfectly. There is no guarantee of end-to-end registration of NWs, that is, they will shift relative to one another.

3 Controlling MBE NWs with High-K Dielectric Regions

When SNAP is used, all the NWs are exactly the same. Thus, to control individual NWs, some external means of

differentiating them is needed. The idea proposed by Jim Heath [14] is to first dope the NWs so that they are sensitive to external fields and then deposit high-K dielectric regions over NWs to shield them from the influence of fields. If such a region is deposited over the first half of SNAP-generated NWs, a large field provided by a LW at right angles will not affect their conductivity although it will cause the remaining half of the NWs to lose their conductivity. A second region that shields the second half of the NWs from a second LW provides a means to select one half of the NWs to be conducting. This same process can be repeated to divide each half of the NWs into two sets, one conducting and the other not using two more LWs. Repeating this process and using $2\log_2 N$ LWs, N a power of two, provides a means to select one of N NWs to be conducting.

Unfortunately, this method cannot work in practice because it isn't possible to put down high-K dielectric regions whose length and separation is on the order of nms. We revisit this question again later.

4 Data Storage in Crossbars

It is expected that data will be stored in nanometer crossbars (nanoarrays) by interposing a thin molecular layer of long molecules between two orthogonal sets of NWs where the molecules at crosspoints form memory cells [2, 19, 3]. Under high positive voltage, they become conducting while under high negative voltage they become non-conducting. The conductivity of a crosspoint is sensed with low voltage that doesn't change the state. Claims made for storage density of $\geq 10^{11}$ bits/cm^2 [21].

5 Designing Crossbars using Mod-Doped NWs

Many challenges present in the design of crossbars made from mod-doped NWs [6]. The number of LWs should be kept small, the chip area should be minimized, and data should be stored efficiently in nanoarrays.

To meet these challenges, good physical models are needed, crossbar architectures must be devised that help mitigate effects of stochastic assembly, and good methods of probabilistic analysis employed to evaluate designs, issues best addressed by physical scientists, electrical engineers, and computer scientists working together.

5.1 Coping with Stochastic Crossbars

The NWs in a crossbar must be connected to ohmic regions at each end so that they will carry current. If it is required that each of the N NWs in each direction in a crossbar have a different encoding with probability at least .99, the size of the code space, C, must be at least $100N^2$ [6]. If $N = 1,000$, a reasonable number, C must be 10^8, an impossible number to manufacture. Multiple ohmic regions dramatically reduce the size of C because one region at a time is activated [6, 10].

Another source of uncertainty in crossbars is the identity of NW encodings. They must be identified after crossbar assembly and a standard mapping assigned to the encodings found in the array. Three methods of producing such a mapping are discussed below [10].

The goal of a crossbar design strategy is to maximize with high probability n_A, the number of unique internal addresses that are accessible externally in each dimension. The designer is free to choose, m, number of ohmic regions, w, number of NWs per ohmic region, C, size of code space, the type of NW encoding (such as (h, n)-hot and binary-reflected code) and the error rate ϵ.

5.2 Evaluation of Mapping Strategies

Three illustrative address mapping strategies are
S_a: All NWs in each ohmic region are unique;
S_b: Most NWs in each ohmic region are unique;
S_c: All C NW encodings occur at least once in each region.

We use the area A of a complete array as a quality measure. It is approximated by the sum of the area of a translation memory, an ohmic region decoder, and the area of the NWs themselves. Figurately, we write $A \approx A_{TranslationMemory} + A_{Decoder} + A_{Nanoarray}$. $A_{TranslationMemory}$ is the product of the number of bits to store a mapping times the area per bit, $A_{Decoder}$ is made proportional to $m \log_2 m$, and $A_{Nanoarray}$ is the square of the NW length devoted to storage and control.

Each architectural strategy has some probability of success. Bounds on such probabilities are obtained by evaluating tails of distributions or using variants on the coupon collector problem. The latter is modeled by cereal boxes each of which is equally likely to have any one of C coupons. We ask "How many boxes need to be purchased so that with probability at least $1 - \epsilon$, all coupons are collected?"

5.2.1 Strategy S_a Versus S_b

The areas used by strategies S_a and S_b when binary-reflected codes are used are shown below. Here C_a and C_b are the sizes of the code space used by them, λ_{nano} and λ_{litho} are the feature sizes of nano and lithographic technologies, and σ is the area of one bit of DRAM storage.

$$\begin{aligned} A_a &\approx 2(n_A \log_2 C_a)\sigma + 2\lambda_{litho}(m \log_2 m) \\ &\quad + (2\lambda_{litho} \log_2 C_a + \lambda_{nano} n_A)^2 \\ A_b &\approx 2(n_b \log_2 C_b)\sigma + 2\lambda_{litho}(m \log_2 m) \\ &\quad + (2\lambda_{litho} \log_2 C_b + \mathbf{2}\lambda_{nano} n_A)^2 \end{aligned}$$

Although the formulas are the same except for a factor of two multiplier, the number of codes C_a needed for strategy S_a is much larger than that needed for strategy S_b. In fact, when $\epsilon = .01$, $w \geq 10$ and $m < 5,000$, all reasonable assumptions, analysis shows that $C_a = 50 n_A w$ and $C_b < 15w$ [10]. Thus, S_b is clearly superior to S_a.

5.2.2 Strategy S_c Versus S_b and DRAM

Strategy S_c, which sets the parameters so that all NW codes appear in every ohmic region with probability at least $1 - \epsilon$, is clearly inferior to S_b. Most wire encodings are repeated multiple times, which is a waste of real estate. However, when binary-reflected codes are used, the translation process is trivial. For this reason, S_c may offer a practical first step to replacing the DRAM with a nanotechnology-based crossbar [10].

5.2.3 Summary Comparison of Strategies

S_b is superior to the other two mod-doped crossbar design strategies. However, it requires discovery of internal addresses and address translation, unlike S_c, for which the translation process is trivial.

6 SNAP-Based Crossbars

As mentioned above, SNAP can be used to produce very long, straight NWs with very few defects. However, all the NWs are the same. The use of high-K dielectric regions to differentiate them has a problem, regions as small as the pitch of NWs cannot be produced. The proposed solution is to randomly place the smallest regions [14].

Three models for this random placement have been created and analyzed [11]. It has been shown that in all models the number of small regions needed to guarantee that all NWs are controllable is large. This means that this method of controlling NWs is likely to be impracticable.

7 Data Storage in Large Nanoarrays

The goal is to write an array of 1s and 0s into a crossbar in a small number of steps. It will be possible to write subarrays of 1s (stores) or 0s (restores) in one step. We ask "How difficult is it to find optimal storage sequence?" and "Do restores help?".

It has been shown that the first question is **NP**-hard [12]. That is, unless **P** = **NP**, no polynomial time algorithm exists for it. As a consequence, approximation algorithms have been examined. This problem has also been shown **NP**-hard but only when stores are used [12]. It is not known whether the approximation problem becomes easy when restores are also allowed.

As a consequence of these results, other approaches have been explored [12]. They include programming arrays using subarray operations on a bounded number of rows or columns, for which good approximate polynomial-time algorithms exist, and developing programs for particular structured arrays, for which fast algorithms have been found. For example, the diagonal $n \times n$ matrix, which requires n operations when on stores are allowed can be done in $2\lceil \log_2 n \rceil$ steps when restores are also allowed.

8 Other Materials and Systems Issues

New materials are being developed that have the potential to create other realizable nano-electronic components. Core-shell NWs [20] in which shells are superimposed on NW are in this category as is a new method to metallize NWs [27] and greatly reduce their resistance after they have been assembled into arrays. NW resistance can have a large impact on the time to store data in arrays.

Coping with defects and faults [7, 25], minimizing power consumption, and making design decisions to simplify manufacturing are other issues that are under investigation.

9 Conclusions

Nanotechnology offers the promise of continuing the Moore exponential growth of chip density. However, it introduces new problems that need to be addressed. The small size of nanometer-sized objects means that control of their exact placement is replaced by stochastic assembly over which the designer and manufacturer have some control. The principal challenges of nanotechnologies are to cope with stochastic assembly, the defects that arise naturally, and the embedding large problems into arrays of randomly assembled structures. Nanotechnology is an emerging field full of interesting problems.

My collaborators in the work reported here are André DeHon, Benjamin Gojman, Lee-Ad Gottlieb, Charles Lieber, Patrick Lincoln, Eric Rachlin, and Arkady Yerukhimovich.

References

[1] Wolfgang Arden and et al. International technology roadmap for semiconductors, 2001. see http://public.itrs.net.

[2] Yong Chen, Gun-Young Jung, Doublas A. A. Ohlberg, Xuema Li, Duncan R. Stewart, Jon O. Jeppeson, Kent A. Nielson, J. Fraser Stoddart, and R. Stanley Williams. Nanoscale Molecular-Switch Crossbar Circuits. *Nanotechnology*, 14:462–468, 2003.

[3] C. P. Collier, E. W. Wong, M. Belohradský, F. M. Raymo, J. F. Stoddart, P. J. Kuekes, R. S. Williams, and J. R. Heath. Electronically Configurable Molecular-Based Logic Gates. *Science*, 285:391–394, 1999.

[4] Y. Cui, L. Lauhon, M. Gudiksen, J. Wang, and C. M. Lieber. Diameter-Controlled Synthesis of Single Crystal Silicon Nanowires. *Applied Physics Letters*, 78(15):2214–2216, 2001.

[5] André DeHon. Array-Based Architecture for FET-based, Nanoscale Electronics. *IEEE Transactions on Nanotechnology*, 2(1):23–32, March 2003.

[6] André DeHon, Patrick Lincoln, and John E. Savage. Stochastic Assembly of Sublithographic Nanoscale Interfaces. *IEEE Transactions on Nanotechnology*, 2(3):165–174, 2003. see http://www.cs.brown.edu/people/jes/nano.decoders.pdf.

[7] André Dehon and Helia Naeimi. A Greedy Algorithm for Tolerating Defective Crosspoints in NanoPLA Design. In *Procs. 2004 Int. Conf. on Field-Programmable Technology*, 2004.

[8] Cees Dekker. Carbon Nanotubes as Molecular Quantum Wires. *Physics Today*, pages 22–28, May 1999.

[9] X. Duan, Y. Huang, and C. M. Lieber. Nonvolatile Memory and Programmable Logic from Molecule-Gated Nanowires. *Nano Letters*, 2(5):487–490, 2002.

[10] Benjamin Gojman, Eric Rachlin, and John E Savage. Decoding of Stochastically Assembled Nanoarrays. In *Procs 2004 Int. Symp. on VLSI*, Lafayette, LA, Feb. 19-20, 2004.

[11] Benjamin Gojman, Eric Rachlin, and John E Savage. Analysis of the Randomized SNAP Decoder, November 6, 2004.

[12] Lee-Ad J. Gottlieb, John E. Savage, and Arkady Yerukhimovich. Efficient Data Storage in Large Nanoarrays. *Theory of Computing Systems*, to appear.

[13] Mark S. Gudiksen, Lincoln J. Lauhon, Jianfang Wang, David C. Smith, and Charles M. Lieber. Growth of Nanowire Superlattice Structures for Nanoscale Photonics and Electronics. *Nature*, 415:617–620, 2002.

[14] Jim Heath. Personal Communication, 2004.

[15] Y. Huang, X. Duan, Y. Cui, L. Lauhon, K. Kim, and C. M. Lieber. Logic Gates and Computation from Assembled Nanowire Building Blocks. *Science*, 294:1313–1317, 2001.

[16] Y. Huang, X. Duan, Q. Wei, and C. M. Lieber. Directed Assembly of One-Dimensional Nanostructures into Functional Networks. *Science*, 291:630–633, 2001.

[17] E. Johnston-Halperin, R. Beckman, Y. Luo, N. Melosh, J. Green, and J.R. Heath. Fabrication of Conducting Silicon Nanowire Arrays. *Applied Physics Letters A*, 2004. to appear.

[18] Franklin Kim, Serena Kwan, Jennifer Akana, and Peidong Yang. Langmuir-Blodgett Nanorod Assembly. *Journal of the American Chemical Society*, 123(18):4360–4361, 2001.

[19] P. J. Kuekes, R. S. Williams, and J. R. Heath. Molecular Wire Crossbar Memory, US Patent Office, US Patent Number 6,128,214. Technical report, October 3, 2000.

[20] Lincoln J. Lauhon, Mark S. Gudiksen, Deli Wang, and Charles M. Lieber. Epitaxial Core-shell and Core-multishell Nanowire Heterostructures. *Nature*, 420:57–61, 2002.

[21] Nicholas A. Melosh, Akram Boukai, Frederic Diana, Brian Gerardot, Antonio Badolato, Pierre M. Petroff, and James R. Heath. Ultrahigh-Density Nanowire Lattices and Circuits. *Science*, 300:112–115, April 4, 2003.

[22] Gordon Moore. Cramming More Components Onto Integrated Circuits. *Electronics*, 38(8):114–117, 1965.

[23] A. M. Morales and C. M. Lieber. A Laser Ablation Method for Synthesis of Crystalline Semiconductor Nanowires. *Science*, 279:208–211, 1998.

[24] Thomas Rueckes, Kyoungha Kim, Ernesto Joselevich, Greg Y. Tseng, Chin-Li Cheung, and C. M. Lieber. Carbon Nanotube-Based Nonvolatile Random Access Memory for Molecular Computing. *Science*, 289:94–97, 2000.

[25] Greg Snider, Philip Kuekes, and R. Stanley Willams. CMOS-like Logic in Defective, Nanoscale Crossbars. *Nanotechnology*, 15:881–891, 2004.

[26] Dongmok Whang, Song Jin, and Charles M. Lieber. Nanolithography Using Hierarchically Assembled Nanowire Masks. *Nano Letters*, 3(7):951–954, 2003.

[27] Yue Wu, Jie Xiang, Chen Yang, Weil Lu, and Charles M. Lieber. Single-Crystal Metallic Nanowires and Metal/Semiconductor Nanowire Heterostructures. *Nature*, 430:61–65, 2004.