Appearing in HotChips 15, Stanford University, August 17–19, 2003

Sub-lithographic Semiconductor Computing Systems

André DeHon
Joint work with: Charles M. Lieber, Patrick Lincoln, and John Savage
contact: <andre@cs.caltech.edu>
Department of Computer Science, 256-80
California Institute of Technology
Pasadena, CA 91125

July 30, 2003

Enabled by advances in our basic scientific understanding at the molecular and atomic scales, we can now engineer designed nanostructures without using lithography. Key features can be a few nanometers wide—a few silicon atoms wide, perhaps the ultimate scale for devices. This allow us to design computing components without the costs or limits of ultra-fine lithography. Design at this scale, however, will not simply be an extension of our familiar VLSI design. We may not be able to directly pattern complex features, but rather must exploit basic physical properties to define feature sizes, self-assembly to create ordered devices, and post-fabrication reconfigurability to define functionality and mask defects. This creates new challenges for design and exposes a different cost structure which motivates different computing architectures than we found efficient in conventional, lithographically patterned silicon. I will review the emerging nanoscale fabrication building blocks, sketch a hybrid fabrication scheme which uses these building blocks along with lithography, and present a plausible architecture for nanoscale electronics based on silicon nanowires. I demonstrate that these nanoscale constructs are sufficient to provide universal logic functionality with all logic and signal restoration operating at the nanoscale.

Lithographic patterning has long been the primary way of defining features in semiconductor processing. Consequently, our ability to build smaller and denser components has been directly tied to our ability to engineer smaller ways to reliably, lithographically define smaller and smaller regions. At the same time, a large part of the exponential cost required to achieve each geometric feature size shrink is directly related to the cost of lithographically defining smaller and smaller feature sizes (*e.g.* [1]).

Scientists are developing a growing repertoire of techniques which allow us to define features (*e.g.* wire widths, spacing, active device areas) without lithography. Many of these are "bottom up" techniques that rely on the fundamental physics of molecules and crystals to self-assemble structures from individual atoms with tight dimensions. The result is wires which are just a few atoms wide and can be packed together at the pitch of just 10's of atoms as well as active devices which may be as small as a single molecule.

This raises an interesting question: *Can we define and build computing systems without lithography?* The benefit to doing so is clear—we can engineer nanoscale systems without being limited by the current lithographic technology and the associated costs of such tight pitch lithography. Another way of asking this question is: *Can we build systems bottom up, from the assembly of individual atoms and molecules, rather than top down, by gross patterning, etching, and doping?* Bottom up construction places a whole new set of constraints on what we can build—we don't have the flexibility to plot anything we'd like to build, but rather, must arrange for regular, crystal-like, structures to self-assemble into useful arrangements. Our bulk assumptions break down as we approach the level of individual atoms for our wires, forcing us to work with law of large numbers ordering affects above the level of wires and devices rather than below it. An immediate consequent is that our computing structures must tolerate inevitable errors and variations in the devices and wires we assemble.

In principal, it now appears that we do have a sufficient set of building blocks to engineer fully nanoscale systems without relying on lithography for the nanoscale features. Lithography can assist at a larger scale to provide an important scale bridge between these molecular-scale devices and the kinds of IO wiring we traditionally use to connect to electronic components.

Sublithographic assembly exploits seeding, molecular self-organization, and time controlled growth and etch to define basic feature sizes. Semiconducting nanowires with diameters down to 3 nm (approximately 6 atoms wide) have been grown using a catalyst seed to control nanowire diameter ([5] [11]). These wires can be composed of different materials or selectively doped along their length using timed growth ([8]). Nanowires can also have a doping pattern which varies radially outward from the nanowire core ([10]). Flow techniques can be used to align a set of nanowires into a single orientation, close pack them, and then transfer them onto a surface ([9] [12]). Switchable molecules can be assembled, placing one or a few molecules under each junction in a crossed array ([2] [3]), providing sublithographic scale, programmable junctions.

With these building blocks, we can build diode junctions by crossing P-doped and N-doped nanowires, use field-effects to control conduction in semiconducting nanowires, and employ a variety of programmable junction effects to implement switchable crosspoints or memory bits (*e.g.* [3], [4]). Using the axial variation of doping or materials, a single nanowire can have regions which are gateable and other regions which are not gateable. These devices are sufficient to build programmable diode-OR gates, programmable memory points, and field-effect based inverting and restoring logic gates. Using flow techniques, tight-pitched, crossed nanowire arrays can be built. By selecting the kinds of nanowires and how they are sandwiched, we can realize a number of useful, array building blocks including programmable OR planes, memory planes, and signal restoration or inversion planes. As a result of the assembly restrictions, much of the challenge here lies in arranging to build the all the necessary functionality into the crossed nanowires—particularly providing a single straight wire which allows bootstrap programming, normal operation, and inter-array communications. Detailed solutions are developed in [6].

Figure 1 shows how these pieces are assembled into a tile for a universal, programmable computing array. These tiles can be further arrayed to build large-scale devices. The basic tile alternates arrays of programmable OR and (possibly fixed) inversion/NOR planes, realizing a PAL-like structure. All arrays have nanowires which overlap adjacent arrays for signal communications. Lithographic scale wires are used for addressing of stochastically coded wire ends to select and energize a single nanowire in the row and a single in the column of an array; this allows both

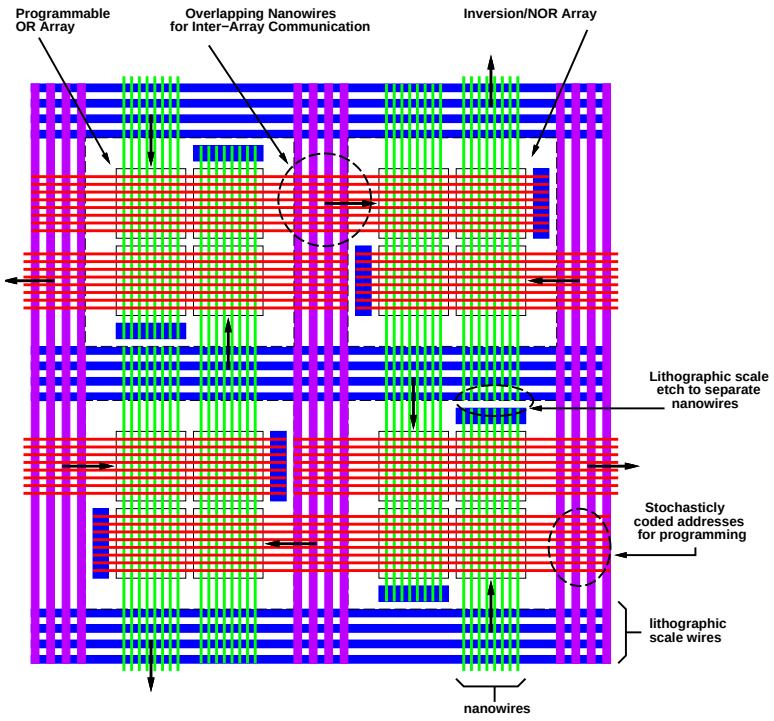


Figure 1: Master Programmable Structure: Not shown to scale. Typical arrays would be a few hundred nanowires on a side to amortize out the cost of the lithographic programming support.

bootstrap device programming and signalling between the lithographic and sublithographic scales [7]. Broken wires in the array are identified during a testing phase. Programming of the nanowire crosspoints serves both to define function and to allow routing around defective wires. The result is a programmable array of large-fanin NOR gates with interconnect. We know NOR gates are universal logic building blocks, so this combination provides a nanoscale programmable array which can implement any logic function. With additional array differentiation, selective blocks can be used as read-write memory (RAM) which can be interfaced with the programmable logic blocks. The array logic structure is developed further in [6].

As this sketch shows, we now appear to have an adequate set of building blocks to construct complete computing systems, both memories and active computing devices, where the critical features, logic operation, and density are all defined at sublithographic scales. Features are defined by seeding and timed growth and self-assembled into regular, tight-pitch arrays. Stochastic coding provides unique addressability for post-fabrication configurability. Configurability allows us to differentiate and customize the regular arrays to perform our designed logic function and allows us to avoid inevitable fabrication defects. Significant work remains at all levels to bring this together and realize a usable computing system; nonetheless, recent advances in technology and organization are offer growing confidence in the feasibility of systems of this scale and capability.

Acknowledgment This research was funded by the DARPA Moletronics program under grant ONR N00014-01-0651.

References

[1] International Technology Roadmap for Semiconductors. <http://public.itrs.net/Files/2001ITRS/>, 2001.

[2] Christopher L. Brown, Ulrich Jonas, Jon A. Preece, Helmut Ringsdorf, Markus Seitz, and J. Fraser Stoddart. Introduction of [2]Catenanes into Langmuir Films and Langmuir-Blodgett Multilayers. A Possible Strategy for Molecular Information Storage Materials. *Langmuir*, 16(4):1924–1930, 2000.

[3] C. Collier, G. Mattersteig, E. Wong, Y. Luo, K. Beverly, J. Sampaio, F. Raymo, J. Stoddart, and J. Heath. A [2]Catenane-Based Solid State Reconfigurable Switch. *Science*, 289:1172–1175, 2000.

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

[5] Yi Cui, Lincoln J. Lauhon, Mark S. Gudiksen, Jianfang Wang, and Charles M. Lieber. Diameter-Controlled Synthesis of Single Crystal Silicon Nanowires. *Applied Physics Letters*, 78(15):2214–2216, 2001.

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

[7] André DeHon, Patrick Lincoln, and John Savage. Stochastic Assembly of Sublithographic Nanoscale Interfaces. *IEEE Transactions on Nanotechnology*, 2(?), 2003. *To Appear*.

[8] 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, February 7 2002.

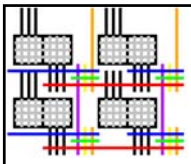
[9] Yu Huang, Xiangfeng Duan, Qingqiao Wei, and Charles M. Lieber. Directed Assembley of One-Dimensional Nanostructures into Functional Networks. *Science*, 291:630–633, January 26 2001.

[10] Mark S. Gudiksend Lincoln J. Lauhon, Deli Wang, and Charles M. Lieber. Epitaxial Core-Shell and Core-Multi-Shell Nanowire Heterostructures. *Nature*, 420:57–61, 2002.

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

[12] Dongmok Whang, Song Jin, and Charles M. Lieber. Nanolithography Using Hierarchically Assembled Nanowire Masks. *Nano Letters*, 3, 2003. to appear shortly.

Web links for this document:
<http://www.cs.caltech.edu/research/ic/abstracts/sublitho_hotchips2003.html>



Sub-lithographic Semiconductor Computing Systems

André DeHon
andre@cs.caltech.edu
In collaboration with
Charles Lieber, Patrick Lincoln,
and John Savage

Approaching the Bottom

- In 1959, Feynman pointed out we had
  - "plenty of room at the bottom"
- Suggested:
  - wires ~ 10-100 atoms diameter
  - circuits ~ few thousands angstroms
    ~ few hundred nm

Approaching the Bottom

- Today we have 90nm Si processes
  - bottom is not so far away
- Si Atom
  - 0.5nm lattice spacing
  - 90nm ~ 180 atoms diameter wire

Exciting Advances in Science

- Beginning to be able to manipulate things at the “bottom” -- atomic scale engineering
  - designer/synthetic molecules
  - carbon nanotubes
  - silicon nanowires
  - self-assembled mono layers
  - designer DNA

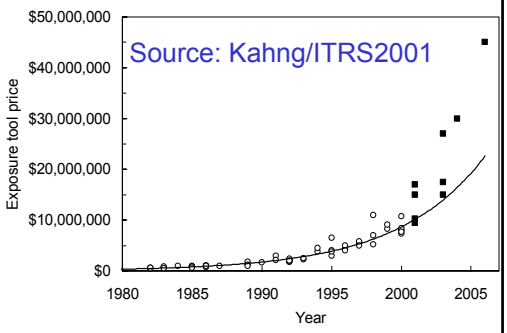
Question

- Can we build interesting computing systems without lithographic patterning?
- **Primary interest:**

  below lithographic limits

Why do we care?

- Lithographic limitations
  - Already stressing PSM
  - …xrays, electron projection…
- Lithographic costs



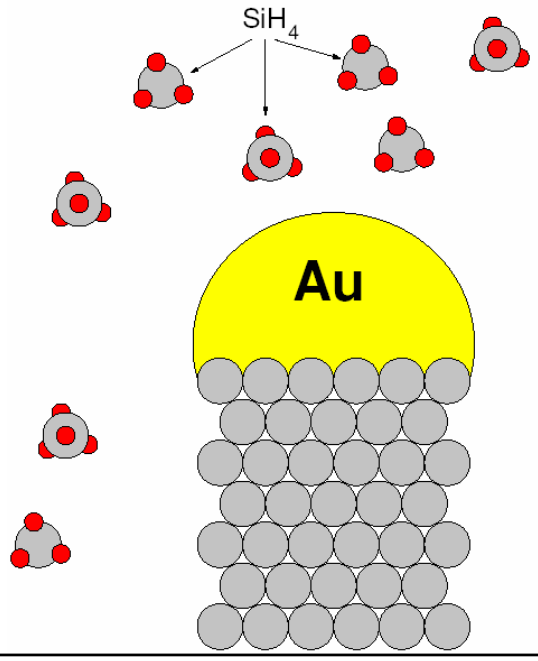
Today's Talk

Bottom up tour: from Si atoms to Computing

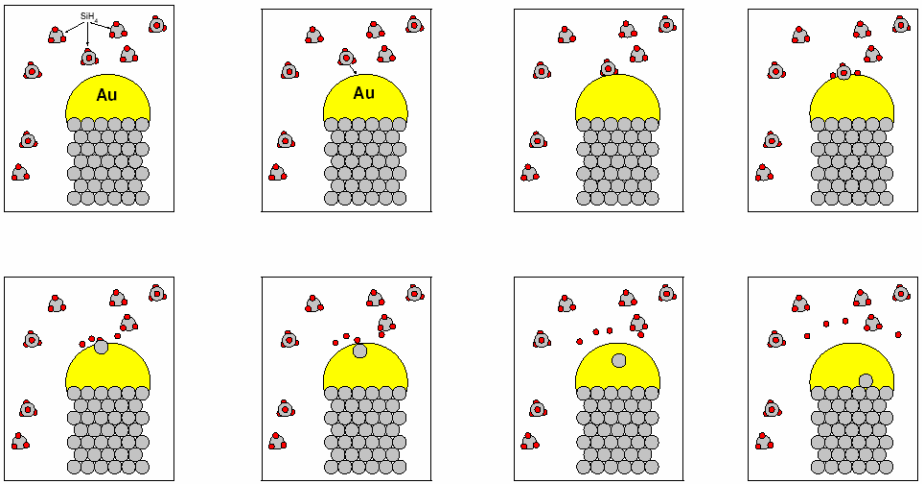
- Nanowire growth
- Nanowire devices
- Nanowire assembly
- Nanowire differentiation
- Nanowire coding
- Nanoscale memories from nanowires
- Nanoscale PLAs
- Defect tolerance
- Universal Computing blocks defined at nanoscale

SiNW Growth

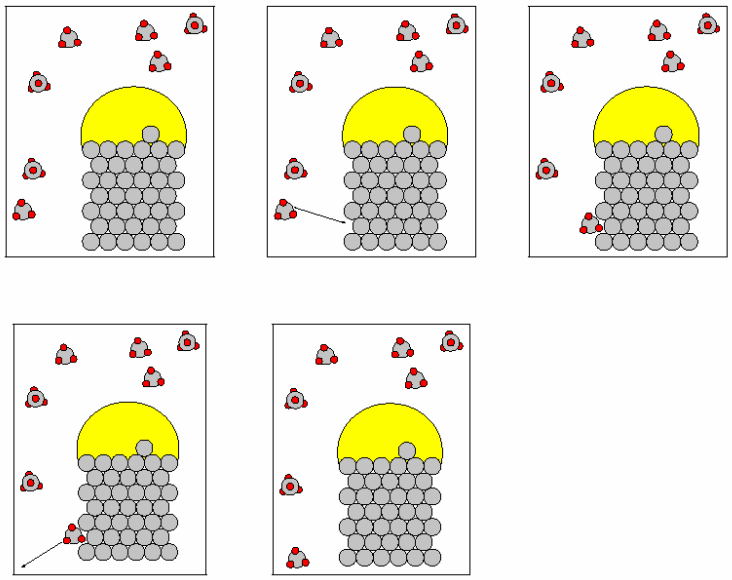
- Atomic structure determines feature size
- Self-same crystal structure constrains growth
- Catalyst defines/constrains structure



SiNW Growth



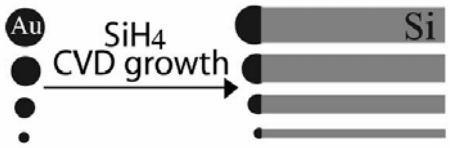
SiNW Growth



Building Blocks

Semiconducting Nanowires

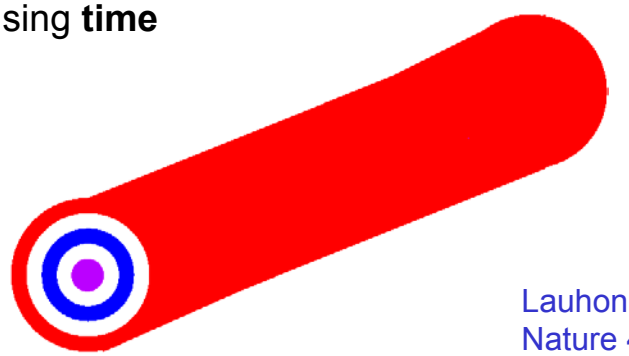
- Few nm's in diameter (*e.g.* 3nm)
  - Diameter controlled by seed catalyst
- Can be microns long
- Control electrical properties via doping
  - Materials in environment during growth
  - Control thresholds for conduction



From:
Cui…Lieber
APL v78n15p2214

Radial Modulation Doping

- Can also control doping profile radially
  - To atomic precision
  - Using **time**

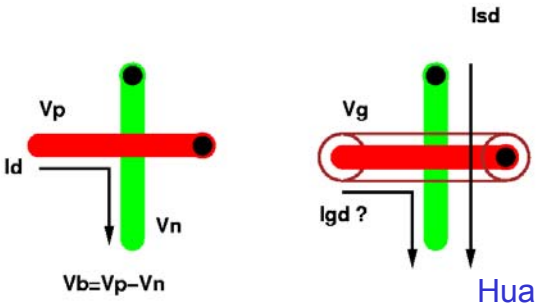


Lauhon et. al.
Nature 420 p57

Devices

Doped nanowires give:

Diode and FET Junctions

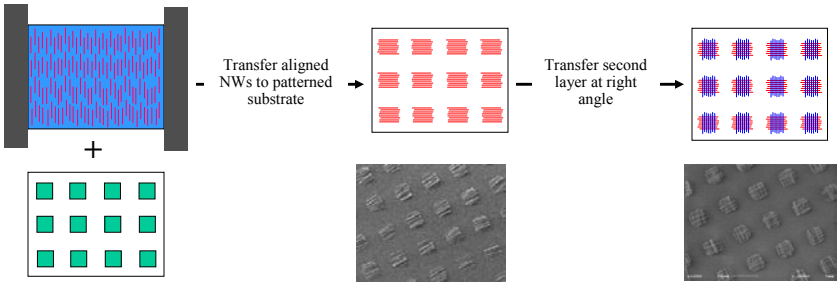


Cui…Lieber
Science 291 p851

Huang…Lieber
Science 294 p1313

Langmuir-Blodgett (LB) transfer

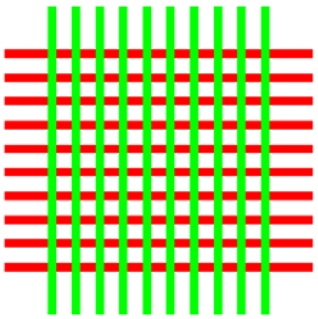
- Can transfer tight-packed, aligned SiNWs onto surface
  - Maybe grow sacrificial outer radius, close pack, and etch away to control spacing



Whang, Nano Letters 2003 (to appear)

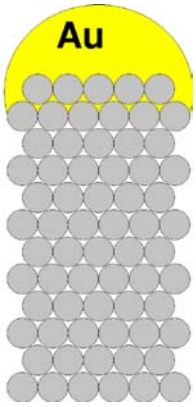
Homogeneous Crossbar

- Gives us homogeneous NW crossbar
  - Undifferentiated wires
  - All do the same thing



Control NW Dopant

- Can define a dopant profile along the length of a wire
  - Control lengths by **timed** growth
  - Change impurities present in the environment as a function of time



Gudiksen et. al.
Nature 415 p617

Björk et. al.
Nanoletters 2 p87

Control NW Dopant

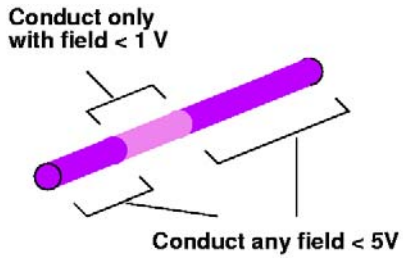
- Can define a dopant profile along the length of a wire
  - Control lengths by **timed** growth
  - Change impurities present in the environment as a function of time
- Get a SiNW banded with differentiated conduction/gate-able regions

Gudskien et. al.
Nature 415 p617

Björk et. al.
Nanoletters 2 p87

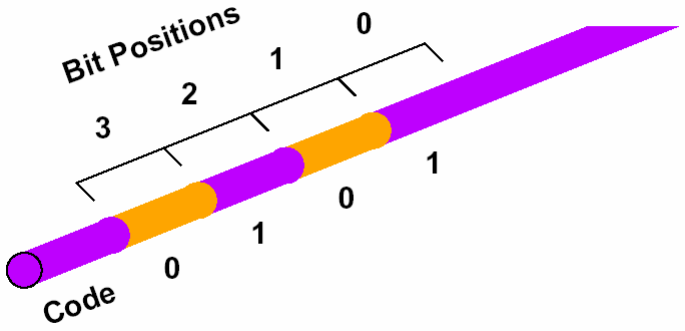
Enables: Differentiated Wires

- Can control which regions of a wire are gate-able
  - Lightly doped regions → gate with low threshold
  - Heavily doped regions → gate with high threshold
- Can engineer so portions of wire oblivious to applied voltage (always conduct) and others controlled



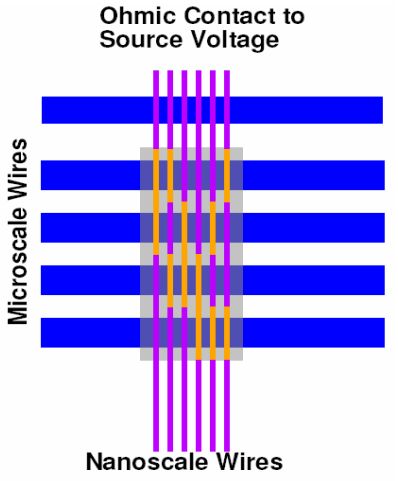
Coded Wires

- By selectively making bit-regions on wires either highly or lightly doped
  - Can give the wire an address



Unique Set of Codes

- **If** we can assemble a set of wires with unique codes
  - We have an address decoder
    - Apply a code
      - k-hot code
    - Unique code selects a single wire

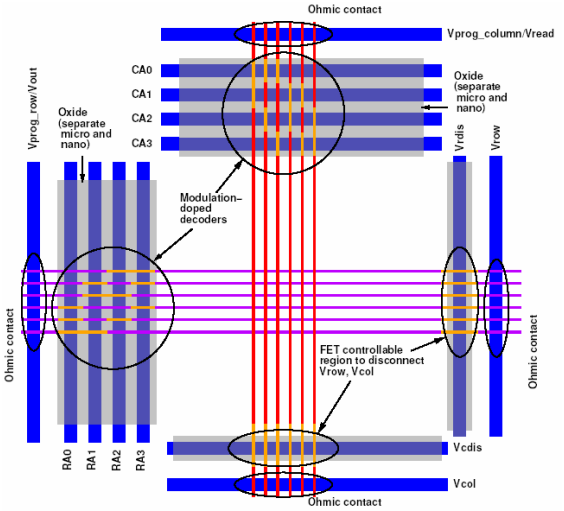


Statistical Coding

- Unique Code set achievable with statistical assembly (random mixing)
- Consider:
  - Large code space (10^6 codes)
  - Large number of wires of each type (10^{12})
  - Small array (10 wires) chosen at random
- Likelihood all 10 unique?
  - Very high! (99.995%)

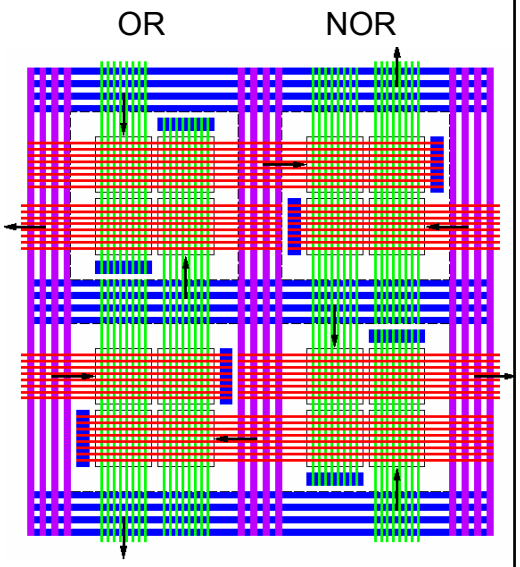
DeHon et. al.
IEEE TNANO to appear

Basis for Sublithographic Memory

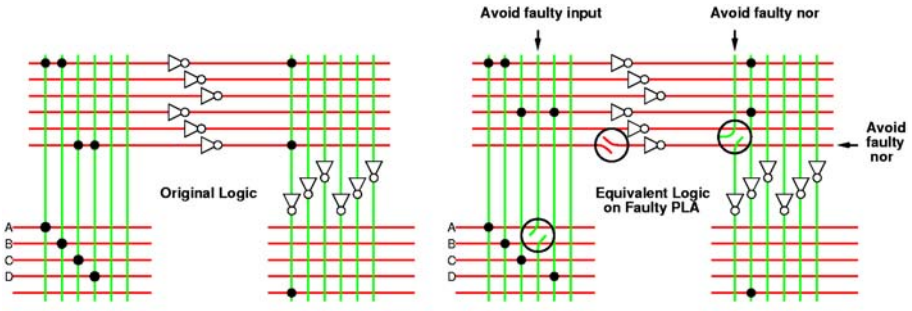


Connected PLAs

- Programmable OR planes like memory
- NW cross arrays for interconnect
- FET planes to restore/invert
- Manhattan routing
- Fully nanoscale computing



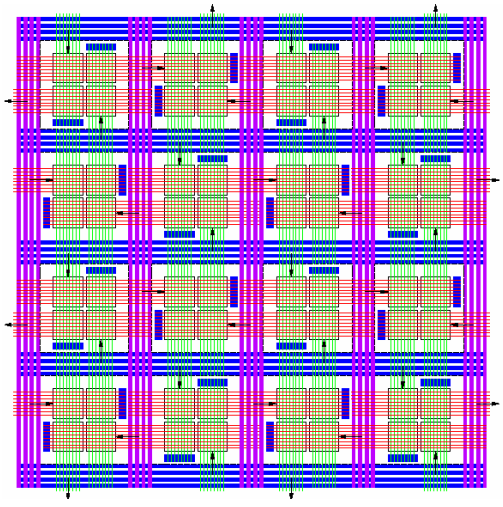
Defect Tolerant



All components (PLA, routing, memory) interchangeable;
Allows local programming around faults

Universal Computing Device

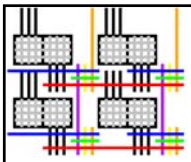
- Tile Array Block
- Programmable Array
- NOR universal
- Implement any computation



DeHon
IEEE TNANO v2n1

Construction Review

- Seeding control NW diameter
- Timed growth controls doping profile along NW
- LB flow to assemble into arrays
- Timed etches to separate/expose features
- Assemble on lithographic scaffolding
- Stochastic construction of address coding allow micro→nanoscale addressing
- Differentiate at nanoscale via post-fabrication programming
- All compatible with conventional semiconductor processing
  - Key feature is decorated nanowires

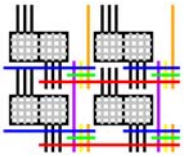


Summary

- Can engineer designer structures at atomic scale
- Must build regular structure
  - Amenable to self-assembly
- Can differentiate
  - Stochastically
  - Post-fabrication programming
- Sufficient for Memories and Universal, Programmable Architecture
- Sufficient building blocks to define computing systems without lithography

Additional Information

- <http://www.cs.caltech.edu/research/ic/>
- <http://www.cmliris.harvard.edu/>

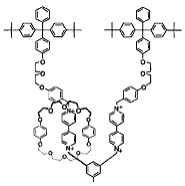


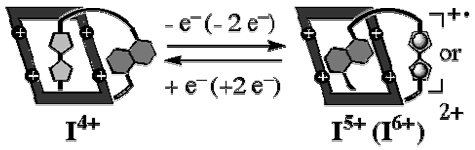
Additional Slides

- Memory Elements
- Logic
- Code Size
- Array Size

Switches / Memories

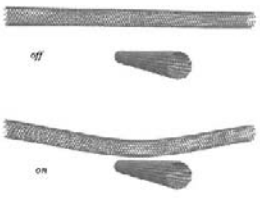
Molecular Switches





Collier et. al.
Science 289 p1172

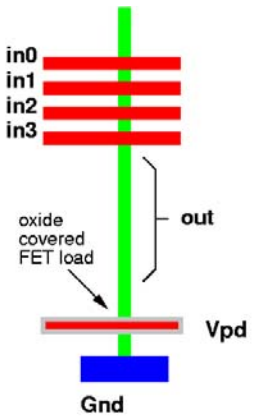
Electrostatic Switches

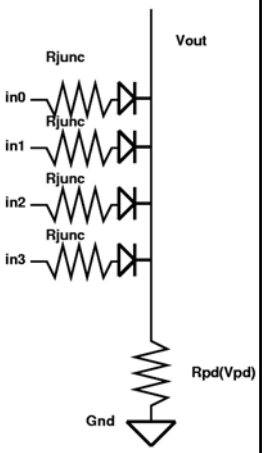


Ruekes et. al.
Science 289 p04

Diode Logic

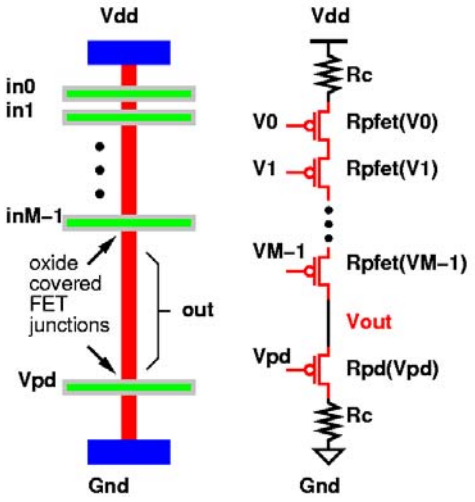
- Arise directly from touching NW/NTs
- Passive logic
- Non-restoring
- Non-volatile Programmable crosspoints



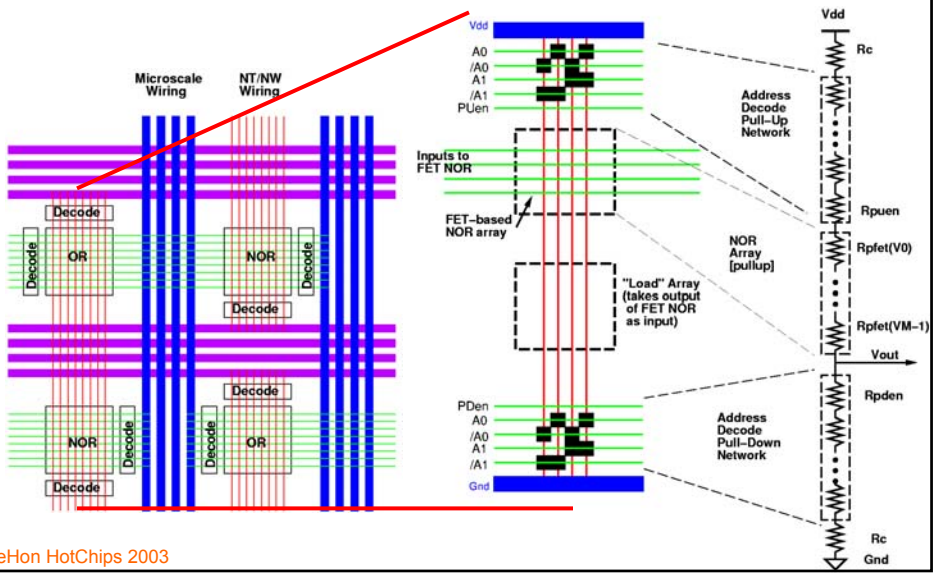


PMOS-like Restoring FET Logic

- Use FET connections to build restoring gates
- Static load
  - Like NMOS (PMOS)
- Maybe precharge

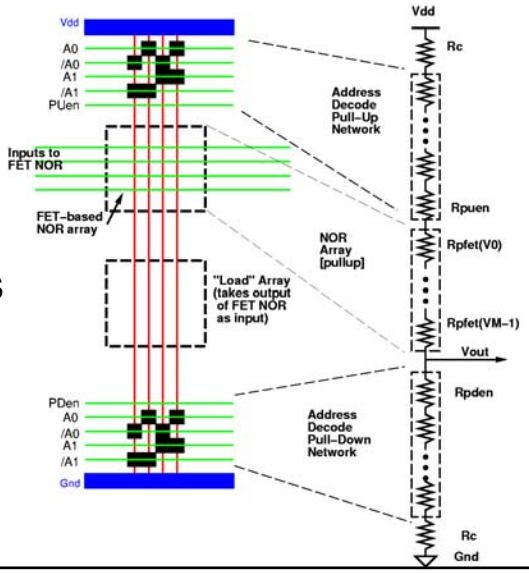


Recall PLA



Operating Array

- Decoders allow program array
  - OR, NOR
- Isolatable
- Dual role of loads during operation
- Output used directly by consumer



Codespace: How Large?

- How large does code space really need to be?
  - Addressing N wires
  - With code space $100N^2$
  - Has over 99% probability of **all** wires being unique
  - For logarithmic decoder:
    - Need a little over 2× bits of sparse code

Array Size

- Larger crossbar
  - Amortize out microscale addressing overhead
- Smaller crossbars
  - Shorter wires
    - Less capacitance → faster, less energy
    - Less likely to fail
  - More efficient for logic

Array Size Summary

- Based on
  - Relative size of structures
    - Micro vs. nano
  - Overhead of current model
  - Current defect rate estimates
- Modest arrays appropriate
  - 512 NT/NW per side
  - A(512)=30
  - A_{side} =30*90nm + (512+11)*10nm
  - 45-65% yield ?
  - 400-800 nm^2/crosspoint

90nm DRAM
49,000 nm^2

22nm DRAM
3400 nm^2?