

Comparison of Analog and Digital Nanosystems: Issues for the Nano-Architect

Pritish Narayanan, Teng Wang, Michael Leuchtenburg, Csaba Andras Moritz
Dept. of Electrical and Computer Engineering
University of Massachusetts
Amherst MA USA
{pnarayan, andras}@ecs.umass.edu

Abstract Improvements in manufacturing and assembly of novel devices such as the carbon nanotube (CNT) and semiconductor nanowires has led to the exploration of new nano-architectures utilizing such devices for applications such as general purpose computing, image processing etc. This paper discusses different possible nanoscale implementations of Cellular Neural Networks (CNN). It is seen that while an analog nanoscale implementation of the CNN may be difficult with self-assembly based approaches given the requirements for customization of devices and arbitrary routing, a digital equivalent may be realizable in the near term. One such digital CNN design based on the NASIC nanoscale fabric is shown. A specialized architecture for CNN with Resonant Tunneling Diodes (RTD) is also discussed.

I. INTRODUCTION

Hybrid nanoscale architectures are a promising alternative to end of the line CMOS. Various architectures have been proposed targeting digital and analog applications. For instance, NASICs use semiconductor nanowire grids with supporting CMOS control to implement digital microprocessors [1]; CMOL is a 3D fabric with molecular devices, CMOS and nanowire layers that may be used for reconfigurable digital logic [12] as well as analog neural network implementations [17]; Cellular Neural Networks (CNNs) and Cellular Automata have been proposed using Single Electron Transistors (SETs) [18] etc.

These novel architectures are deemed feasible primarily due to advances in the areas of devices and manufacturing. For example [5] has shown the assembly of semiconductor nanowires into regular grids and the formation of useful devices such as diodes and Field-Effect Transistors (FETs) at the crosspoints. A ring oscillator based on carbon nanotubes (CNT) has been shown in [7]. In addition, there has also been a considerable research effort in the area of alternate state variables, i.e. novel methods to store and encode information such as molecular rotations [22], electron spins [23], magnetic spin wave bus [15] etc. as opposed to the traditional approach that uses electron charge. These devices and variables carry the promise of higher densities, better performance and possibly

lower power consumption.

However, there are many manufacturing constraints for systems utilizing such novel devices and variables. Firstly, non conventional methods of manufacturing such as based on DNA scaffolding [8], fluidic alignment [9], di-block copolymer [24] etc. favor the formation of regular structures. Arbitrary placement of devices and routing of signals is not likely possible in such systems. Also, these manufacturing routes are subject to far higher defect rates (few percent and greater) than seen in today's technologies. These constraints thus impact what architectures and systems can be built.

Nanoscale architectures proposed thus far can be categorized into three main groups: *General Purpose Computing Systems* such as based on NASIC [1], *FPGA style reconfigurable fabrics* such as CMOL [12], NanoPLA [2] and FPNI [6], and *Image processing systems* such as Analog Crossnets [17], and Digital Signal Processing (DSP) architectures [19] utilizing fabrics such as NASIC and CMOL.

In this paper, challenges facing nano-architects are discussed in the context of image processing architectures. For this purpose, the cellular neural network (CNN) [13] paradigm has been chosen for an initial evaluation since it lends itself to a multitude of possible implementations and fabrics. Comparing different nanoscale approaches targeting the same application can provide an insight into the issues and advantages for each as well as on how these approaches are fundamentally different from a traditional CMOS based design.

The rest of the paper is organized as follows: Section II provides a brief overview of Cellular Neural Networks. Section III discusses analog implementations and their feasibility at the nanoscale. Section IV discusses digital implementations, one possible CNN design using the NASIC fabric and also a proposed manufacturing route for such a design. A specialized architecture utilizing novel devices such as the Resonant Tunneling Diode (RTD) for CNN implementation is shown in Section V. Section VI summarizes the various approaches and concludes the paper.

II. OVERVIEW OF CELLULAR NEURAL NETWORKS

The Cellular Neural Network [13] (CNN, also Cellular Nonlinear Network) is a massively parallel computing system

The authors acknowledge support from the Center for Hierarchical Manufacturing (CHM), University of Massachusetts, Amherst. This work is also supported by the National Science Foundation under Grant No. 0541066. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

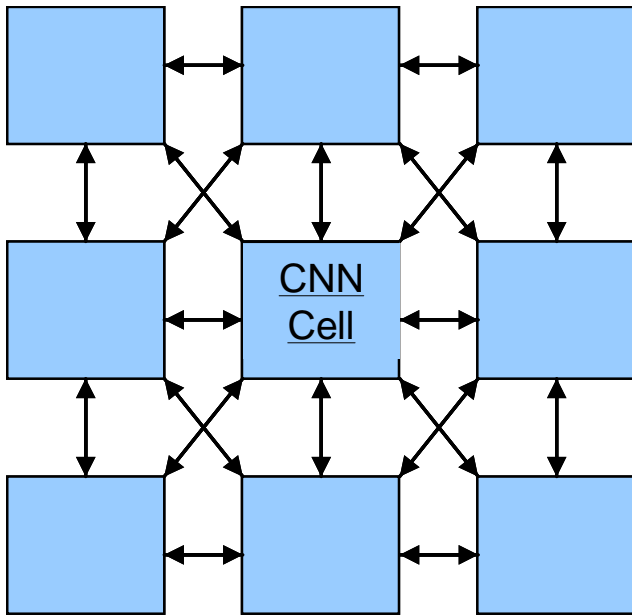


Figure 1. Schematic of a Cellular Neural Network System with local interconnections between identical units called cells

made of identical units called cells. Each cell processes one pixel of an image and has an input, an output and an internal state that evolves based on certain dynamic rules called templates. Most interconnections are local, and each cell exchanges signals with its nearest neighbors. This system as depicted in Fig.1 may be used for applications such as image processing where the values of the templates define the nature of the particular task being performed.

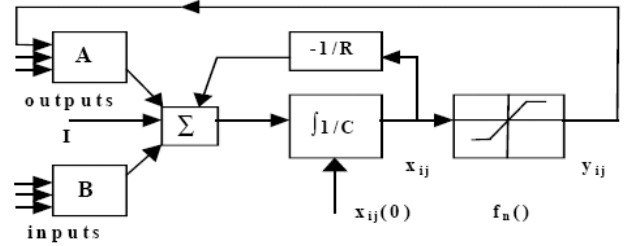
While the original CNN of Chua and Yang is an analog system [13], many digital implementations have since been proposed [14][16]. More recently, implementations involving novel devices such as Single Electron Transistors (SETs) [18] and spin-wave bus [15] have been discussed.

This type of collective computation system with parallel processing of image pixels may be a ‘natural fit’ for nanoscale architectures because:

- i) High densities are achievable for simultaneous processing of large pixel arrays for real time image processing tasks
- ii) Non-conventional manufacturing methods are more amenable toward the formation of regular structures
- iii) Local interconnectivity of cells with minimal global routing and
- iv) The nature of the application is such that achieving 100% fault free components is not a requirement, in contrast to general purpose computing.

However, there are many open issues that need to be addressed at all levels of the design including:

- i) A complex functional unit may be difficult to manufacture with non-conventional methods, even if the structure is uniform at the system level.
- ii) Choice of nanodevice: Should we ‘port’ a conventional system onto a nano-fabric and utilize devices such as CNTFETs and Nanowires, or should a transistorless design



$$C \frac{\partial x_{ij}(t)}{\partial t} = -\frac{1}{R} x_{ij}(t) + \sum_{c_{kl} \in N'_s} A(i, j; k, l) y_{kl}(t) + \sum_{c_{kl} \in N'_s} B(i, j; k, l) u_{kl} + I$$

$$y_{ij}(t) = \frac{1}{2} (|x_{ij}(t) + 1| - |x_{ij}(t) - 1|)$$

Figure 2. Schematic of Analog CNN from [20] with expression for current addition at the state node, and nonlinear output relation.

based on devices such as quantum dots be explored?
 iii) Suitable defect tolerance will still be required given the high defect rates in manufacturing. The impact of such redundant units on density and performance also needs to be studied.

III. ANALOG IMPLEMENTATION OF CNN

Figure 2 shows one possible analog implementation of the CNN cell. The cell is an electrical circuit implementing an analog transformation in continuous time. The input, output and state are all analog voltages. Nearest neighbor interaction is achieved by current flow and the dynamic rule of evolution of the CNN is the Kirchoff Current Law at the state node as shown in the expression on Fig. 2. Values ‘A’ and ‘B’ are the template values and define the nature of the image processing task.

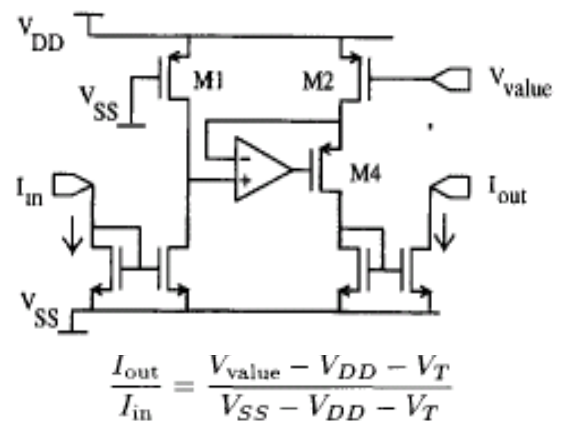


Figure 3. OPAMP based Circuit for implementing programmable templates from [21]. Externally controlled voltage determines scaling of currents

These are usually space invariant, i.e. the same template values are stored on all cells in the system. The output is a nonlinear function of the voltage at the state.

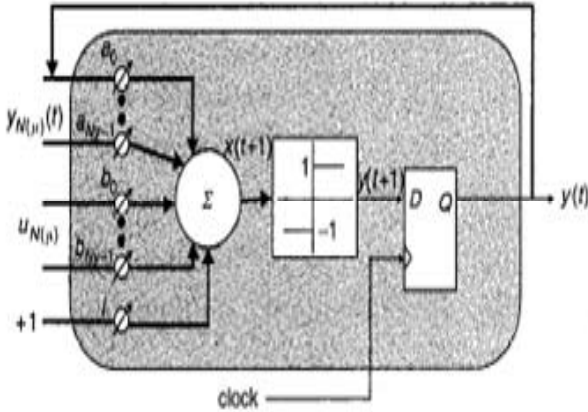
Manufacturing on the nanoscale is an immediate stumbling

block for the system shown in Fig. 2. Key components in the conventional analog design such as current mirrors, voltage-current converters etc require complementary doping at arbitrary locations, which is incompatible with a self-assembly based approach. Also, to achieve programmability of scaling factors circuitry requiring precise control of transistor operating points, such as shown in Fig. 3 [21] is required. This level of precision and programmability is unlikely to be available at the nanoscale in the range of 10nm. Also many routing issues exist, including the requirement for self feedback and complicated interconnections between various circuit components. It is therefore unlikely that an analog CNN can be directly ‘ported’ into any known nanoscale fabric.

IV. DIGITAL IMPLEMENTATION OF CNN

The digital implementation of the CNN is functionally identical to the analog implementation. However the cell in this case is a digital circuit evaluating a Boolean expression over discrete clock cycles. One possible implementation is shown in Fig. 4 [14]. Here the input, state and output are all binary values. The nearest neighbor interaction is achieved through data-buses connecting the cells. State evolution progresses through the execution of combinational logic followed by latching at a clock edge. The relevant expressions are shown in Fig. 4. Key components in the design include a digital accumulator for state accumulation and multi-bit registers for storing template values.

As opposed to the analog implementation that requires customization and complex devices, the digital implementation consists of relatively straightforward digital components. Also as discussed previously, some digital systems from nano-fabrics such as NASIC and CMOL have been shown. Therefore, it should be possible to build any purely digital system utilizing either of these fabrics that, while not



$$x_{\mu}(n+1) = \sum_{\lambda \in N(\mu)} a_{\lambda-\mu} y_{\lambda}(n) + \sum_{\lambda \in N(\mu)} b_{\lambda-\mu} u_{\lambda} + I$$

Figure 4. Schematic of Digital CNN from [14] with expression for the state evolution through accumulation of components

demonstrated yet, would likely have less manufacturing challenges.

In the following subsections we provide some insight into the tradeoffs involved in selecting a nanoscale fabric by comparing NASIC and CMOL, an initial design for a Digital CNN built using the NASIC fabric and also a brief description of a possible manufacturing route for such a design.

A. Comparison of NASIC and CMOL

NASIC (Nanoscale Application Specific Integrated Circuit) is a fabric architecture based on semiconductor nanowire grids with FETs at certain crosspoints. The nanowires are connected to microwires which provide control signals generated from CMOS circuitry. The nanowire grids are laid out in tiles, with each tile implementing two-stage logic with a dynamic control style to channel the flow of data through these tiles. Figure 5 shows a single NASIC tile implementing a full adder function using a 2-level AND-OR logic style. NASIC implementations are self-healing: defects are masked using built-in redundancy and error correcting circuits on the nanogrid coupled with system level voting in CMOS. Single FET designs are also possible.

CMOL [12] is a 3-dimensional fabric consisting of an underlying CMOS layer for some logic operations and signal restoration functions along with a nanowire grid placed on top of this layer for interconnection (Fig. 6). Reconfigurable molecular switches are available at the crosspoints of the grid for performing signal routing and wired-OR logic implementation. 2-level logic implementations are possible in CMOL using a NOR-NOR scheme.

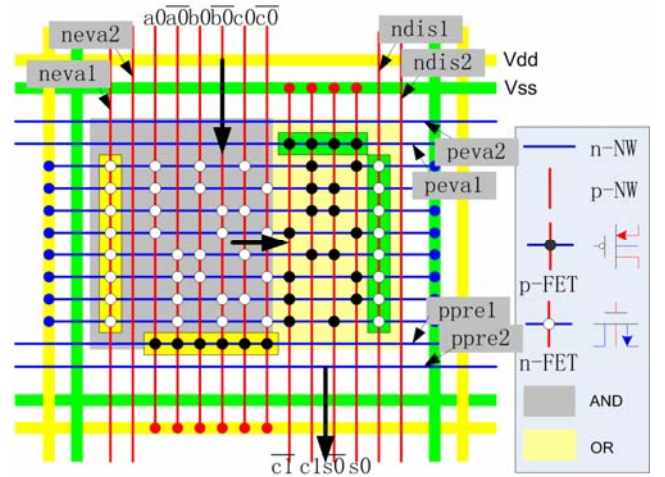


Figure 5. 1-bit full adder in NASIC. The thicker wires represent microwires (MWs) and the thin ones are NWs. The doping type of the wires (p-type or n-type) along source-drain of a FET transistor determines the type of the transistor. The black and white dots, at the crosspoints of NWs, represent p-FETs and n-FETs respectively. The function is based on AND-OR logic. Arrows show propagation of data through the tile.

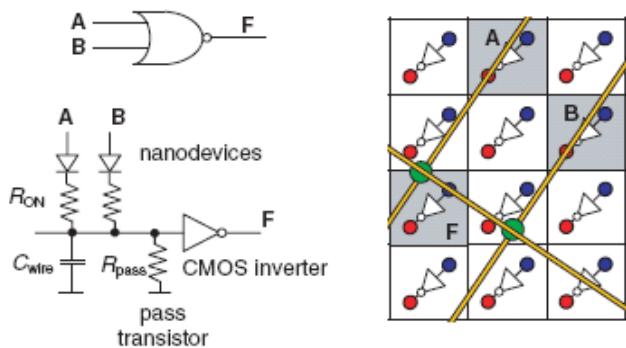


Figure 6. CMOL implementation of NOR gate from [12]. Diagram shows inverters in CMOS, along with metal nanowires and dots represent reconfigurable molecular switches.

Key tradeoffs between NASIC and CMOL include:

Density: NASIC density scales with improvements in nanofabrication whereas CMOL density is limited by the density of the underlying CMOS layer that implements part of the logic.

Micro-Interfacing: NASIC has a relatively simple interfacing with the microscale through peripheral microwires. Defect map extraction is not necessary since fault tolerance schemes are

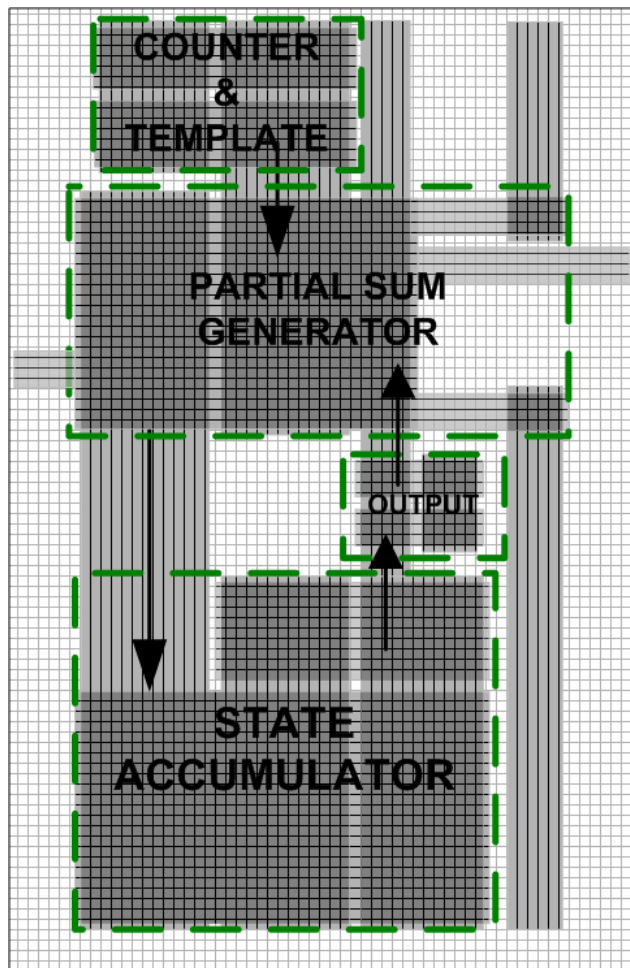


Figure 7. Floorplan of a single digital CNN cell implemented in NASIC and composed of four tiles. Arrows show propagation of data.

built into the fabric itself. CMOL on the other hand requires defect map extraction to reconfigure around defective crosspoints.

Nanogrid: NASIC nanogrid manufacturing is expected to be more challenging since the positions of transistors are fixed. The CMOL nanogrid is relatively simple to manufacture.

While there are important challenges for both the nanofabrics, it is expected that they may be realized in the near term with improvements in the areas of manufacturing and assembly.

B. Digital CNN design for NASIC

Figure 7 shows the floorplan for a NASIC design implementing one cell of the digital CNN. The system consists of four tiles, a counter cum template storage register, a partial sum generator that calculates a partial sum from any one neighbor depending on the counter value, a state accumulator that accumulates the partial sum at each iteration and an output generator that calculates the output once all the state values

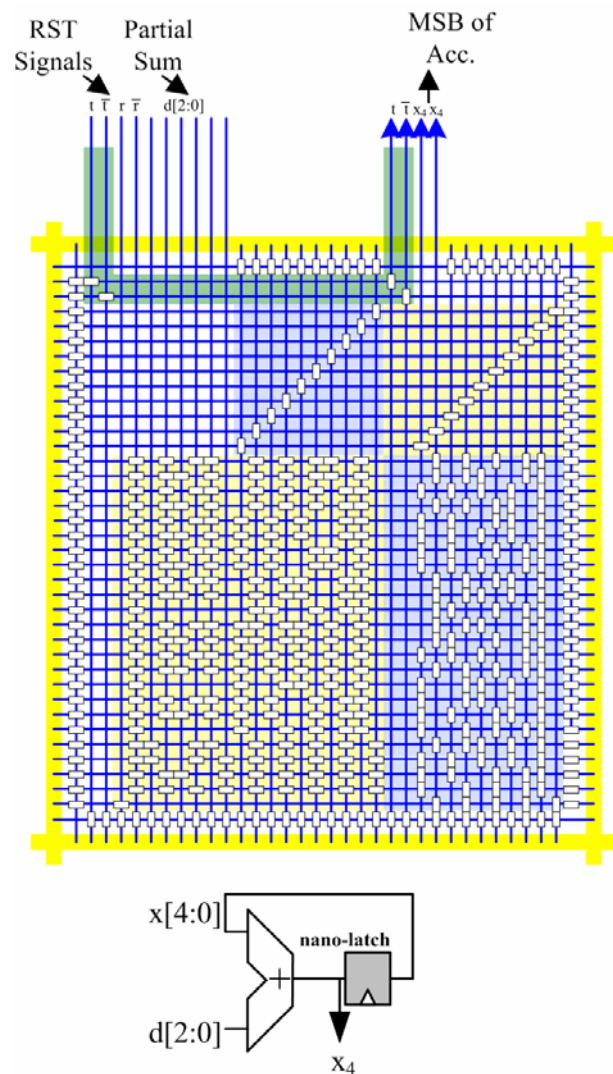


Figure 8. State Accumulator circuit implementation and schematic. White boxes represent n-FETs. Partial sum received from partial sum generator and accumulated. MSB after accumulation is sent to the output generator.

have been accumulated. This system exercises most NASIC principles such as a dynamic logic style for streaming datapaths, implicit latching on the wire and built in fault tolerance schemes for providing resilience against manufacturing defects, process variation and noise related issues.

Figure 8 shows the circuit level implementation of the state accumulator using Field-Effect Transistors (FETs). The tile uses a two level NAND-NAND logic. The partial sum $d[2:0]$ is received from the partial sum generator and then accumulated with the previous value. These reset the accumulator as well as generate the outputs once all terms have been accumulated.

C. Proposed Manufacturing route for NASIC Digital CNN

Manufacturing of a NASIC CNN may be achieved through a combination of self-assembly and conventional top-down manufacturing steps.

- NWs may be grown using seed catalyst techniques or other methods that may ensure uniform NW diameters [3]. NWs may be aligned into parallel horizontal and vertical sets with Langmuir-Blodgett (LB) techniques or other approaches including nanopatterning and laterally aligned lamellar phase diblock copolymers.
- Regions on NWs where there should be no FET channels will then need to be metallized using a lithographic mask [11]. Though a 2NW pitch resolution is required, precise shaping is not needed, making this step less challenging as compared to a CMOS manufacturing step for a similar feature size.
- An oxide layer may then be grown over the gate regions of the NWs and a 2D grid formed by moving one NW set on top of the other.
- A fine grain metallization step is needed to demarcate FET channels, create metallic interconnect between neighboring FETs. This may be achieved by using the top NW as a self-aligning mask as shown in [10].
- Micro-Nano interfacing may be done in conjunction with lithographic process steps.

Although many of the key individual steps required have been demonstrated, combining the necessary steps for reliable manufacturing remains challenging and unproven. However, with improvements in these areas, a NASIC based CNN solution is feasible in the relatively near term. Ongoing research effort is also focused on theoretical comparisons of how such a system compares against equivalent CMOS implementations scaled to aggressive technology nodes.

V. RESONANT TUNNELING DIODE BASED CNN

It may be possible to leverage the unique properties of certain nanodevices to design CNN systems such as shown in [15][18]. As opposed to the previous approach involving the ‘porting’ of a digital design onto a transistor based nanofabric NASIC, this approach has no equivalent implementation at the microscale. The proposed system from [15] is shown in Fig. 9.

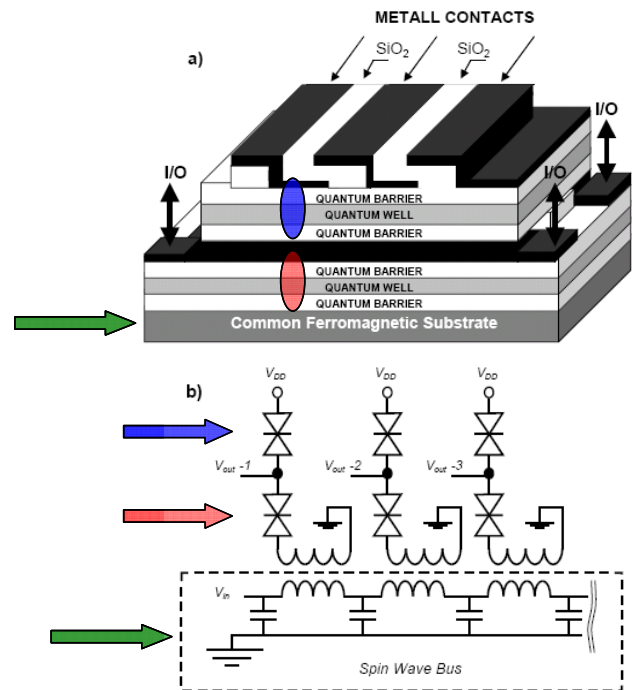


Figure 9. RTD based CNN system and equivalent circuit from [15]. Quantum layers represent RTD cells (blue and red arrows). The ferromagnetic substrate (green arrow) propagates magnetic spin waves and is inductively coupled to the RTD cells.

The CNN cells are RTD pairs connected in series. These act as a bistable element storing a (0/1) state. The input is an analog voltage, whereas the state and the output are binary. The nearest neighbor interaction is achieved through the common ferromagnetic substrate (spin wave bus) in which instead of electron charge, magnetic spin waves may propagate. The spin wave bus is inductively coupled to the CNN cells as shown in the equivalent circuit. State evolution is a voltage change caused by the inductive coupling.

The work in [15] is still in its early stages and there are many issues to be resolved. Some possible issues could include: *Spin wave bus*: Can this effectively replace a metallic interconnect? There are crosstalk and reliability issues, since undesirable coupling between cells must be avoided. As reported in [15] the signal attenuation on the bus is also very high.

RTD cells: Since tunneling is highly sensitive to the thickness, highly uniform layers of materials are required for uniform tunneling characteristics, the manufacturability of which has not been shown.

The advantages of such a system could be high density of RTD cells for large scale processing, as well as potential power density advantages of migrating from a charge based interconnect paradigm.

VI. CONCLUSIONS

Different approaches to creating a nanoscale CNN architecture and the challenges for each were discussed. These are summarized in Table I. An analog approach will require extensive CMOS support for functional parts of the system where self assembly based techniques are not feasible. The density of such a design would hence fundamentally be limited by these CMOS portions. The RTD based approach may be practical in the long term, but many open questions on manufacturability and reliability of such a device remain

TABLE I. SUMMARY OF APPROACHES

	Analog	RTD based	Digital
Devices/Fabric	CMOS FETs	RTD, Spin Wave bus	NASIC, CMOL etc
Manufacturing	Difficult with self-assembly	Not shown	Self-assembly + lithography
Density	Lower, as CNN cell is CMOS	High, if reliable assembly is possible (upto 10x)	Highest if using NASIC like fabric (3-10X)
Application Performance	Fast (Efficient for state transition but device is CMOS)	Fast (Phase velocity of spin wave 10^5 m/s, RTD switching THz)	Fast (Limited by NW FET switching in NASIC ~10GHz for processor)

unanswered. The important benefits of such a system could be high densities and performance. The digital approach may be more feasible on the near term utilizing nano-fabrics such as NASIC and CMOL. An example implementation on the NASIC fabric was shown. A potential manufacturing route for a digital CNN based on the NASIC fabric utilizing a combination of self-assembly was discussed. Very high density CNNs could be achieved using the NASIC fabric since all logic implementation is on the nanoscale. The performance of such a design would be limited by the Nanowire-FET switching time. Analysis of processor designs using NASIC [1] has shown operating frequencies in the order of 30-55GHz.

REFERENCES

- [1] C. A. Moritz, T. Wang, P. Narayanan, M. Leuchtenburg, Y. Guo, C. Dezan, and M. Bennisner, "Fault-Tolerant Nanoscale Processors on Semiconductor Nanowire Grids", *IEEE Transactions on Circuits and Systems I*, vol 54, pp. 2422-2437, November 2007.
- [2] A. DeHon. "Nanowire-based programmable architectures", *ACM Journal on Emerging Technologies in Computing Systems*, vol 1, pp. 109-162, 2005.
- [3] W. Lu and C.M. Lieber, "Semiconductor Nanowires," *J. Phys. D: Appl. Physics*, vol. 39, pp. R387-R406, October 2006.
- [4] Y. Cui, X. Duan, J. Hu, and C. M. Lieber, "Doping and Electrical Transport in Silicon Nanowires", *Journal of Physical Chemistry B*, vol. 104, pp. 5213-5216, May 2000.
- [5] Y. Huang, X. Duan, Q. Wei, and C. Lieber, "Directed assembly of one-dimensional nanostructures into functional networks", *Science*, vol. 291, no. 5504, pp. 630-633, January 2001.
- [6] G. S. Snider and R. S. Williams, "Nano/CMOS architectures using a field-programmable nanowire interconnect", *Nanotechnology*, vol. 18, pp. 035204-035214, January 2007.
- [7] Z. Chen J. Appenzeller, P. M. Solomon and L. Yu-Ming, P. Avouris, "High Performance Carbon Nanotube Ring Oscillator", 6th Device Research Conference, June 2006.
- [8] S. Xiao, F. Liu, A. E. Rosen, J. F. Hainfeld, N. C. Seeman, K. Musier-Forsyth, and R. A. Kiehl, "Self-assembly of metallic nanoparticle arrays by DNA scaffolding," *J. Nanoparticle Research*, vol. 4, no. 4, pp. 313-317, August 2002.
- [9] D. C. Duffy, J. C. McDonald, O. J. A. Schueller, and G. M. Whitesides, "Rapid Prototyping of Microfluidic Systems in Poly(dimethylsiloxane)", *Anal. Chem.* vol. 70, pp. 4974-4984, 1998.
- [10] K. K. Likharev, "Single-electron devices and their applications", *IEEE Circuits Devices*, vol. 16, pp. 16-21, 2000.
- [11] Y. Wu, J. Xiang, C. Yang, W. Lu and C. M. Lieber, "Single-crystal metallic nanowires and metal/semiconductor nanowire heterostructures", *Nature*, vol. 430, pp. 61-65, 2004.
- [12] D. B. Strukov and K. K. Likharev, "Reconfigurable Hybrid CMOS Devices for Image Processing", *IEEE Transactions on Nanotechnology*, vol 6, pp. 696-710, November 2007.
- [13] L. O. Chua and L. Yang, "Cellular Neural Networks: Theory", *IEEE Transactions on Circuits and Systems I*, vol. 35, pp.1257-1272, October 1988.
- [14] M. Sindhvani, T. Srikanthan and K. Vijayan Asari, "VLSI efficient discrete-time cellular neural network processor", *IEE Proceedings on Circuits, Devices and Systems*, vol. 149, pp. 167-171, June 2002.
- [15] A. Khitun and K. L. Wang, "Cellular Nonlinear Network based on semiconductor tunneling nanostructure", *IEEE Transactions on Electron Devices*, vol 52, pp. 183-189, Feb. 2005.
- [16] M.-D. Doan, M. Glesner, R. Chakrabaty, M. Hidenreich and S. Cheung, "Realization of a cellular neural network for image processing", *Proc. of the 3rd IEEE Intl. Workshop on Cellular Neural Networks and their Applications*, pp. 85-90, December 1994.
- [17] K. K. Likharev et. al., "CrossNets: High-Performance Neuromorphic Architectures for CMOL Circuits", *Molecular Electronics III, (Ann. New York Acad. Sci.)* vol. 1006, pp. 146-163, 2003
- [18] C. Gerousis et. al. "Modeling Nanoelectronic CNN cells: CMOS, SETs, and QCAs", *Proceedings of ISCAS*, Vol. 1, pp. 274-277, 2000
- [19] D. B. Strukov and K. K. Likharev, "Image Processing with CMOL Circuits", in *Proc. NSC'07*, pp. 9-16.
- [20] M. Milanova, et.al "Applications of cellular neural networks for shape from shading problem," in *Proc. Int. Workshop Machine Learning and Data Mining in Pattern Recognition, Lecture Notes in Artificial Intelligence*, Leipzig, Germany, September 1999, pp. 52-63.
- [21] P. Kinget and M. S. J. Steyaert, "A programmable analog cellular neural network CMOS chip for high speed image processing," *IEEE J. Solid-State Circuits*, vol. 30, pp. 235-243, Mar. 1995.
- [22] J.E. Green et. al. "A 160-kilobit molecular electronic memory patterned at 1011 bits per square centimeter", *Nature*, vol. 445, pp. 414-417, Jan. 2007
- [23] J. Nitta, "Semiconductor Sprintronic", *NTT Technical Review*, vol. 2, no. 6, June 2004
- [24] Thurn-Albrecht, T. et al. "Ultrahigh Density Nanowire Arrays Grown in Self-Assembled Diblock Copolymer Templates", *Science* 290, pp. 2126-2129