Variability in Nanoscale Fabrics: Bottom-up Integrated Analysis and Mitigation[1]

Pritish Narayanan[1], Michael Leuchtenburg[1], Jorge Kina[2], Prachi Joshi[1],
Pavan Panchapakeshan[1], Chi On Chui[2] and C. Andras Moritz[1]

[1]University of Massachusetts Amherst

Amherst MA 01003

[2]University of California Los Angeles

Los Angeles CA 90095

$\{pnarayan, andras\}@ecs.umass.edu$

Emerging nano-device based architectures will be impacted by parameter variation in conjunction with high defect rates. Variations in key physical parameters are caused by manufacturing imprecision as well as fundamental atomic scale randomness. In this paper, the impact of parameter variation on nanoscale computing fabrics is extensively studied through a novel integrated methodology across device, circuit and architectural levels. This integrated approach enables to study in detail the impact of physical parameter variation across all fabric layers. A final contribution of the paper includes novel techniques to address this impact. The variability framework, while generic, is explored extensively on the Nanoscale Application Specific Integrated Circuits (NASICs) nanowire fabric. For variation of σ=10% in key physical parameters, the on current is found to vary by up to 3.5X. Circuit-level delay shows up to 118% deviation from nominal. Monte Carlo simulations using an architectural simulator found 67% nanoprocessor chips to operate below nominal frequencies due to variation. New built-in variation mitigation and fault-tolerance schemes, leveraging redundancy, asymmetric delay paths and biased voting schemes, were developed and evaluated to mitigate these effects. They are shown to improve performance by up to 7.5X on a nanoscale processor design with variation, and improve performance in designs relying on redundancy for defect tolerance - without variation assumed. Techniques show up to 3.8X improvement in effective-yield performance products even at a high 12% defect rate. The suite of techniques provides a design space across key system-level metrics such as performance, yield and area.

1. INTRODUCTION

Emerging nano-materials and devices such as semiconductor nanowires [Lu and Lieber 2006; Cui et al. 2000], carbon nanotubes [Chen et al. 2006] and molecular devices [Collier et al. 1999] have been proposed for novel computational fabrics with density and performance potentially far exceeding the capabilities of scaled CMOS. However, reliable and deterministic manufacturing of such systems continues to be very challenging. Self-assembly based approaches and photolithography

---

[1]This work is partially based on a conference paper submitted to DFT-2010, Japan.

at feature sizes of few tens of nanometers and below are expected to introduce significant levels of permanent defects as well as large variations in physical parameters. While permanent defects have been extensively analyzed at circuit and system levels through approaches such as built-in defect tolerance [Wang et al. 2009; Moritz et al. 2007] and reconfiguration [Strukov and Likharev 2007; Snider and Williams 2007], the impact of parameter variability at all levels of a nanoscale fabric is relatively less well understood.

Parameter variations arise due to imprecision in the manufacturing process as well as fundamental atomic scale randomness. At nanometer dimensions where structures typically consist of tens of atoms/molecules, even a small absolute variation in the number of atoms causes a large shift in the electrical characteristics (e.g., random dopant fluctuation and V_{TH} [Wong et al. 1998]). This could potentially lead to performance deterioration and/or yield loss.

In this paper, we explore the impact of variability on a nanoscale fabric and present techniques tailored to mitigate this impact. We develop a detailed generic methodology that is integrative across device, circuit and architectural layers. This is necessary for emerging technologies that do not have built-in SPICE models, standard cell libraries or established CAD flows. We identify key sources of variability at the physical layer, such as channel and gate dimensions of transistors. We do detailed physics-based 3D simulation of multiple device configurations based on extent of variability in physical parameters and quantify electrical properties (e.g. on-currents, parasitic capacitance). We then characterize delay data for circuits incorporating these devices and use them in architectural simulations to evaluate performance impact on a nanoprocessor design. Results show that variation in physical parameters has a significant impact at higher levels. We discuss new techniques for mitigation of parameter variability in conjuction with high manufacturing defect rates arising from unconventional manufacturing and discuss tradeoffs involved.

The variability framework, while fully generic, is explored extensively on the Nanoscale Application Specific Integrated Circuits (NASICs) nanowire-based computational fabric [Wang et al. 2009; Moritz et al. 2007; Narayanan et al. 2008; Narayanan et al. 2009; Narayanan et al. 2009; Moritz et al. 2011]. NASICs consist of semiconductor nanowire grids with crossed nanowire field effect transistors (xnwFETs) functionalized at certain crosspoints and dynamic data-streaming circuits. Built-in defect tolerance schemes provide resilience against manufacturing defects such as stuck-on xnwFETs [Moritz et al. 2007; Wang et al. 2009]. The NASIC WIre Streaming Processor version-0 (WISP-0) [Wang et al. 2005; Wang et al. 2007; Moritz et al. 2011] is a stream processor on the NASIC fabric that is used as a test case for quantifying variability (specifically performance degradation) as well as for evaluating various techniques to mitigate these effects. While many prior publications, including those cited above, have discussed the principles of the NASIC fabric extensively, this is the first time that a detailed bottom-up exploration of parameter variability is presented and techniques for variability mitigation discussed.

The main contributions of this paper are: i) A novel methodology for integrated exploration of parameter variability across nanodevice, circuit and system levels is

presented; ii) Variability effects are analyzed in detail for xnwFET devices and associated NASIC circuits and systems; and iii) A new suite of built-in fault tolerance techniques is developed to mitigate the impact of variability in conjunction with permanent defects in nanoscale systems - yield and performance are evaluated.

The rest of the paper is organized as follows: Section 2 describes in detail sources of variation, variability models and a generic methodology for integrated explorations. Section 3 analyzes the impact of variability on xnwFET device characteristics, NASIC dynamic circuit delays as well as WISP-0 processor performance. New techniques for ameliorating the effects of variability as well as associated yield-area-performance tradeoffs are presented in Section 4. Section 5 concludes the paper.

A note on related work: There has been some previous work in characterizing properties of nanomaterials (e.g., distributions of nanowire diameters for a particular manufacturing setup [Lu and Lieber 2006; Cui et al. 2001]) and devices (e.g. on-current variation [Mehrotra and Roenker 2007]). [Sverdlov et al. 2003] investigates threshold voltage variations and power issues in sub-10nm n-MOSFETs. The device configurations for xnwFETs investigated in our paper is different from [Sverdlov et al. 2003]. Our work is also more comprehensive: it describes a generic methodology and evaluates parameter variability at all fabric levels, detailed physics-based 3-D simulations of device structures calibrated against experimental data are carried out as opposed to 1-D/2-D approximations used by [Sverdlov et al. 2003], and techniques for mitigation of variability are presented.

A more recent work [Gojman and DeHon 2009] investigates parameter variability for the NanoPLA fabric. [Gojman and DeHon 2009] similarly uses simplifying assumptions at the physical fabric and device levels. For example, as opposed to 3-D simulation, it assumes short channel current equations that incorporate a gradual channel approximation (GCA) which may not be valid for the dimensions under consideration. It is also not clear if mobility reduction due to increased scattering and interface-effects or velocity overshoot are accounted for. Furthermore, while physical sources of variation are mentioned, their impact on electrical characteristics is not clearly established. Physical parameter variations are abstracted into threshold voltage variations of the device, determined by ITRS, without extensive evaluation. To our understanding variation in other electrical characteristics (e.g. parasitic capacitance) is not accounted for. Only two operating points (fully-on or fully-off devices) are considered, as opposed to complete behavioral models explored in this work. Circuit evaluations are also different, with [Gojman and DeHon 2009] assuming simplified RC delay-models instead of detailed HSPICE based simulation. Finally, methods to mitigate parameter variation are also different; NanoPLA is a reconfigurable fabric with [Gojman and DeHon 2009] proposing algorithmic reprogramming, whereas our work uses built-in variation mitigation techniques without defect map extraction, complex micro-nano interfacing, or reconfigurable devices.

2. MODEL AND METHODOLOGY FOR VARIABILITY ANALYSIS

In this section we present the methodology for achieving integrated device-circuit-architectural explorations considering parameter variability. This methodology, while discussed in the context of the NASIC fabric, is fully generic and can be
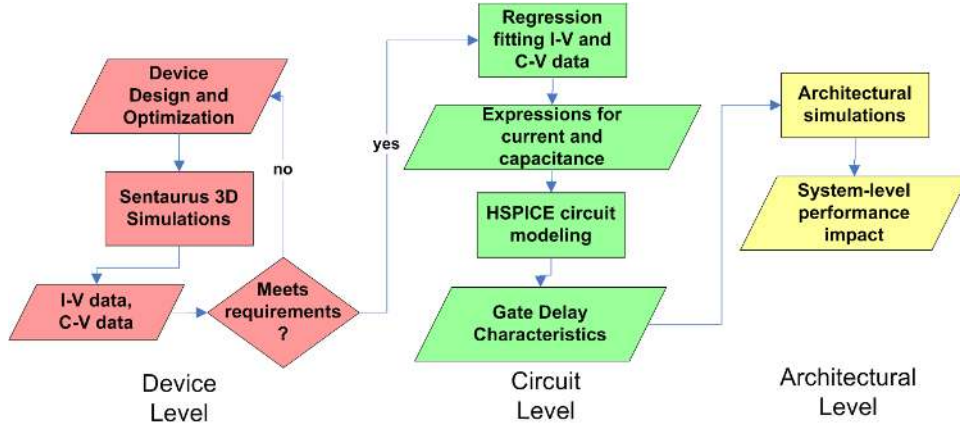
Fig. 1.   Methodology integrating device, circuit and architectural level explorations

applied to other emerging nanoscale computational fabrics for which analytical models of device behavior considering variations are not available. This integrated approach ties physical layer variability to circuit and system level metrics such as delay and performance.

The overall methodology for integrated exploration is presented in the flowchart on Fig. 1. Devices are characterized extensively using Synopsys Sentaurus to extract current-voltage and capacitance-voltage information. Different device configurations are investigated based on values of physical parameters and their behavior quantified. If the device does not meet circuit requirements for correct functionality, device design may be iteratively carried out. Otherwise, the current and capacitance data are fitted using a standard curve-fit tool to obtain mathematical expressions for the data. Using these, a unified behavioral model is created for a circuit simulator such as HSPICE. The unified behavioral model accurately describes the behavior of a single device across a range of input voltages and physical parameter values. Circuit level simulations incorporating Monte Carlo analysis may then be carried out to obtain distributions of circuit delays with parameter variation. This information is then used by a custom nano-architectural simulator to quantify the critical path delays and performance of large-scale designs. To our best knowledge, this framework is a first of its kind. Subsequent sections describe each phase in more detail.

2.1   Physical Layer Aspects: Manufacturing and Devices

Crossed nanowire field-effect transistors (xnwFETs) are the active devices in NASIC designs. A typical xnwFET device structure targeting NASICs is shown in Fig. 2. In this, the top Silicon nanowire acts as the gate and modulates the conductivity of the bottom Silicon nanowire, which is the channel. In an n-type xnwFET, the gate, source and drain regions are doped n^+ and the channel is p-type. Applying a positive voltage on the gate causes inversion in the p-region creating an n-type channel. A thin layer of high-permittivity (high-k) dielectric material (HfO$_2$) separates the gate from the channel.
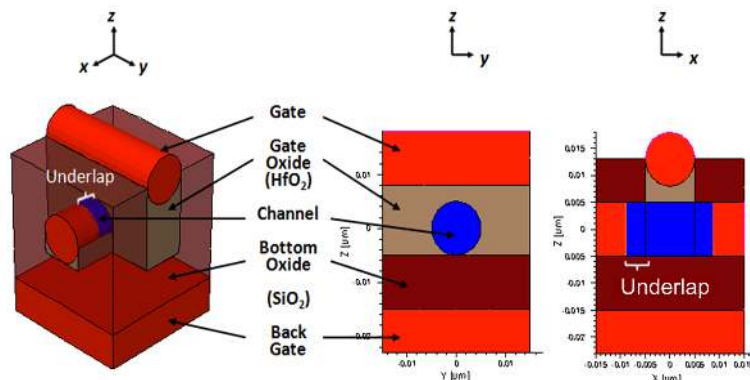
Fig. 2. Crossed Nanowire Field Effect Transistor (xnwFET) structure

A discussion of key steps in NASIC manufacturing and associated alignment considerations is described next to help readers understand process steps and sources of device parameter variability in NASIC fabrics.

2.1.1 *Manufacturing and Alignment.* NASIC manufacturing uses a combination of self-assembly or unconventional patterning, self-alignment and lithographic functionalization steps. Uniform parallel sets of nanowires are assembled on a substrate, followed by lithographic functionalization to define channel regions, interconnect and contacts. xnwFET gate dielectric and an optional underlap are created using self-alignment steps.

Initial Nanowire Alignment: A variety of nanowire alignment techniques including *in-situ* [He et al. 2005; Shan and Fonash 2008; Ural et al. 2002], *ex-situ* [Whang et al. 2003; Xiong et al. 2007; Liu et al. 2006] and unconventional patterning approaches are currently under investigation for the formation of aligned nanowire arrays. While *in-situ* and *ex-situ* nanowire alignment still pose very significant challenges in terms of reproducible, large-scale integration, direct-patterning approaches such as based on Nanoimprint Lithography (NIL) or Superlattice Nanowire Pattern Transfer [Melosh et al. 2003; Heath 2008] (SNAP) are very promising in terms of achieving intrinsic control over pitch and width. For example, SNAP has shown highly regular Silicon nanowire arrays at dimensions as small as 7.5nm width and 15nm pitch [Heath 2008].

Registration and Overlay: One important requirement for integrated systems is overlay alignment. In addition to creating parallel aligned nanowire arrays, it is necessary to create alignment markers for registration. In general, unconventional manufacturing techniques, such as based on NIL have poor overlay alignment (e.g. NIL has been shown to have a $3\sigma = \pm105nm$ overlay imprecision [Picciotto et al. 2009]) whereas conventional lithography has excellent overlay (ITRS 2009 [ITRS 2009a] projects CMOS to have $3\sigma = \pm3.3nm$ for the 16nm technology node). In NASICs, overlay and registration requirements are alleviated due to the following key reasons:

—Nanowire arrays can be direct-patterned on ultra-thin SOI substrates *a priori* to
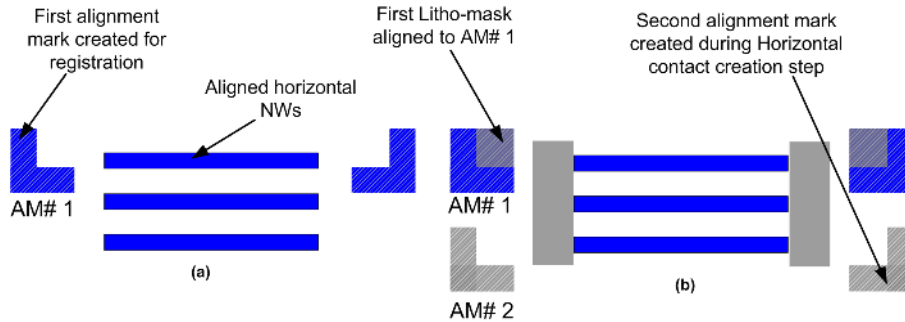
Fig. 3. a) Initial set of alignment marks for registration created simultaneously with nanowire pattern transfer (b) Alignment of first lithographic mask against initial set of alignment markers.

any lithography step using an approach such as NIL. Since this step is carried out before photolithography, no overlay alignment requirement exists.

—Alignment markers can be created on the substrate *at the same time* as the nanowire patterning step itself. For example, if NIL is used, alignment markers and nanowires can be part of the same imprinting mold. The features and markers transferred to the substrate would then be automatically self-aligned. These alignment markers would be used for registering the positions of the nanowires for the first photolithography step (e.g. to create peripheral power rails) with excellent overlay precision (Fig. 3).

—Given that the initial pattern of nanowires created on the substrate is regular (i.e. uniform parallel aligned nanowires), the initial photo-mask can be 'offset' on the grid with some degree of tolerance.

—After the first lithography step is completed, photolithographic alignment between successive masks is expected to be very precise.

Thus, fabric choices in NASIC including *a priori* transfer of nanowires and regular structures mitigate overlay requirements between nanoscale and lithographically defined features. Further information regarding NASIC manufacturing pathways is presented in [Narayanan et al. 2009][2,3].

---

[2]The *a priori* creation of nanoscale features in NASICs is in direct contrast to fabrics such as CMOL [Strukov and Likharev 2005] and FPNI [Snider and Williams 2007]. In these fabrics an underlying CMOS layer needs to be created with conventional lithography before unconventional process steps are used to define nanoscale features. This sets a dramatic limit on the actual overall features that can be accomplished with reasonably precise alignment and good yield. Furthermore, recent work [Vijayakumar et al. 2011] modeling overlay limited yields for NASICs has shown up to 75% yield for $3\sigma = \pm5.7nm$ (manufacturing solutions known according to ITRS 2009) and 100% yield for $3\sigma = \pm3.3nm$ (ITRS projection for 16nm CMOS.)

[3]A note on Stochastic Interfacing: Stochastic interfacing is not required because NASICs do not assume reconfigurable devices where each device/crossbar needs to be programmed. For reconfigurable fabrics, nanowire pairs in a crossbar need to be individually accessed and programmed by driving a potential difference. All functionalization in NASICs is done using a combination of lithography and a self-aligned process and lithography overlay requirements (defined by [ITRS 2009a]) are followed.
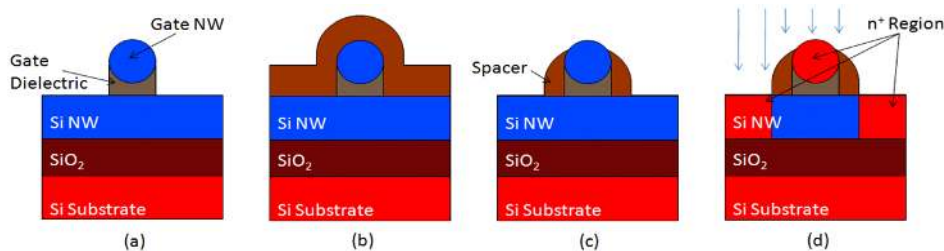
Fig. 4. Front view of the xnwFET during the formation of the source and drain underlap. (a) Initial structure right after channel nanowire, gate dielectric and gate nanowire have been placed into position. (b) A thin layer the spacer material (oxide or nitride) is conformally deposited. (c) The spacer material is anisotropically etched. (d) Ion implantation is performed to dope the source, drain and gate regions.

Self-alignment and Underlap: Gate dielectric and gate underlap are self-aligned against nanowire channels using self-aligned spacer technology (Fig. 4). This process is similar to what is used to form highly doped drain and source (HDD) in CMOS devices and does not need any lithographic masking or overlay. During the anisotropic etch step (Fig. 4c), deposited material on nanowire sidewalls is not completely etched owing to higher thickness (Fig. 4b). This technique provides an extremely good control in the size of the underlap. Furthermore, providing gate underlap is an optional step in the NASIC manufacturing sequence used for device optimization (e.g. V_{TH}, on/off current ratios).

2.1.2 *Device Parameter Variability.* Key sources of variability for a single device were identified based on device structure and manufacturing sequence. These include channel diameter and doping, gate oxide thickness, gate diameter as well as source-drain doping. Table I summarizes all parameters and their extent of variability. Variations in these parameters are dependent on the specific fabrication process used. For example, if a Vapor-Liquid-Solid (VLS) growth method [Lu and Lieber 2006] is assumed for nanowire growth, the gate and channel diameter parameters would be very strongly correlated to variations in the catalyst nanoparticles used as seeds. The standard deviation in wire diameter has been shown to be less than 10% in [Lu and Lieber 2006; Cui et al. 2001]. Similar deviation is seen for Silicon nanowires with SNAP [Heath 2008]. Atomic Layer Deposition for gate oxide formation has been shown to have spatial variability as low as $\sigma=1\%$ [McNeill et al. 2007].

xnwFETs need to be engineered to meet NASIC circuit requirements (e.g., threshold voltage, on-off current ratios [Narayanan et al. 2009]). Device level techniques such as gate underlap and substrate bias were applied in conjunction to achieve these targets. However, these techniques can be sources of additional variability. For example, variation in the length of the underlap can significantly affect I-V characteristics. Since this process step is identical to conventional spacer technology, the ITRS spacer requirements table [ITRS 2009b] estimates the extent of variability allowed for underlap. For a 16nm CMOS technology node this value is $3\sigma=\pm0.6$nm which is 50% of the extent of variability assumed in our work.

Large-scale integrated manufacturing of nanoscale computing systems is still in

Table I.   Device parameters and extent of variation

Parameter	Nominal Value	Standard Deviation
Channel diameter (Cdiam)	10nm	10%
Gate diameter (Gdiam)	10nm	10%
Underlap (Ulap)	4nm	10%
Gate oxide thickness (Gox)	3nm	10%
Bottom oxide (Box)	10nm	10%
Channel doping (Cdop)	$10^{18} dopants/cm^3$	10%
Source-drain doping (Sddop)	$10^{20} dopants/cm^3$	10%

its infancy, and for NASIC system fabrication, different approaches are currently being investigated. Therefore, for our initial variability modeling, we conservatively model 10% standard deviation ($3\sigma=\pm30\%$) for all parameters[4]. Random variation in all parameters is assumed[5]. Furthermore, physical parameters are expected to be uncorrelated since they would be influenced by separate process steps. For example, the gate oxide may be created using Atomic Layer Deposition (ALD) [Ritala and Leskela 2004; McNeill et al. 2007]. There is no dependence of this parameter on any other process step. Similarly, variation in the underlap is purely dependent on the spacers used, and not on any other step.

As more experimental data on device characterization becomes available and detailed process models developed, the modes and extent of variation can be suitably altered.

Accurate 3D-physics-based simulations using Synopsys Sentaurus were carried out to characterize the electrical behavior of the xnwFET device structures. Depending on extent of variability in individual parameters, multiple device configurations were explored. Simulations were calibrated against published experimental data for nanowire FETs at similar dimensions to account for effects such as carrier scattering due to surface roughness and dielectric/channel interface trapped charges. Since parameters are assumed to be uncorrelaterd, in these simulations, each parameter was varied one at a time for $\pm3\sigma$ and the I-V and C-V data were obtained for all device configurations. This data was then used to construct unified behavioral models for circuit simulations.

2.2   Circuit-level simulations

In order to represent the behavior of the device accurately in a circuit simulator such as HSPICE, curve-fitting of the raw data obtained from device simulations

---

[4]For doping levels, each device simulation assumes a discrete number of dopants. 10% standard deviation represents the average deviation over multiple device simulations

[5]Scalable nanofabrication for emerging fabrics does not have the advantage of extensive foundry characterization or fully established process sequences available to CMOS. As such, it is not yet possible to separate die-to-die from inter-die variations. It is possible that systematic effects are likely to be seen for some more conventional process steps, e.g. deposition of gate oxides but less on the non-conventional self-assembly based steps. Self-assembly based approaches tend to have a more significant random variation component. For example, Vapor-Liquid-Solid growth for nanowires shows random variation in nanowire diameters, since these are strongly correlated to the size of gold nanoparticle precursors used for growth [Lu and Lieber 2006]. This phenomenon affects two of our most dominant parameters, the channel and gate diameters.
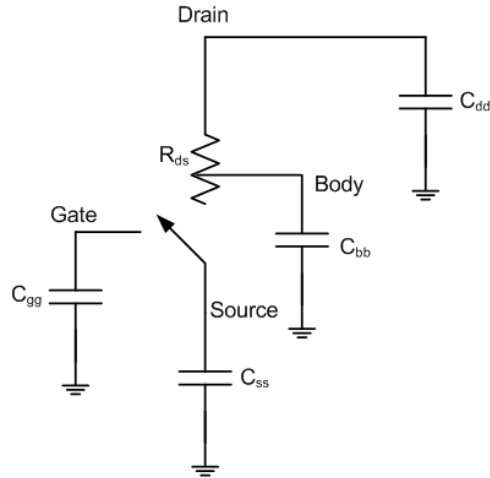
Fig. 5.   Equivalent circuit of a xnwFET showing capacitive and resistive circuit elements.

needs to be done. In this step, the current (and various parasitic capacitances) are fitted as a function of independent variables, i.e., input voltages (drain-source (V_{DS}) and gate-source voltages (V_{GS})) as well as the physical parameters described in Table I. This step was accomplished using the statistical computing tool R. Mathematical expressions describing the current (and capacitances) as functions of the independent variables are then obtained for various regions (see Fig. 1 for flow).

An equivalent circuit for the xnwFET (Fig. 5) was then built into HSPICE incorporating the current source and the parasitic capacitances using sub-circuit definitions. The values of individual elements in are calculated on-the-fly during simulations using the fitted mathematical expressions. The subcircuit definition in conjunction with the expressions for individual elements forms the unified behavioral model for the xnwFET device.

NASIC dynamic circuits were extensively characterized for delay using these models. A typical NASIC dynamic circuit is shown in Fig. 6. It has N inputs, as well as control xnwFET devices for precharge and evaluate. The output node is first precharged to logic '1', and then the *pre* signal is switched off and *eva* is enabled. If all inputs are logic '1', the output node will discharge to logic '0' accomplishing NAND gate functionality. The NAND gate is the universal building block for large scale designs, and its delay behavior needs to be extensively characterized for use in an architectural level simulator.

Delay characterization is done using NASIC dynamic NAND gate with number of inputs varying from 1 to 30. The Monte Carlo simulation framework available with HSPICE was used to vary parameter values and the delay to precharge and evaluate the output node was obtained. Parameters are assumed to follow a Gaussian distribution, with the mean and standard deviation values specified in Table I. They are varied independently for each device, except for the channel diameter which is assumed to be the same across all devices, since all devices are along the same nanowire. Since it may be very hard to do detailed circuit-level simulations on a larger design such as the WISP-0 processor, the delay information is abstracted
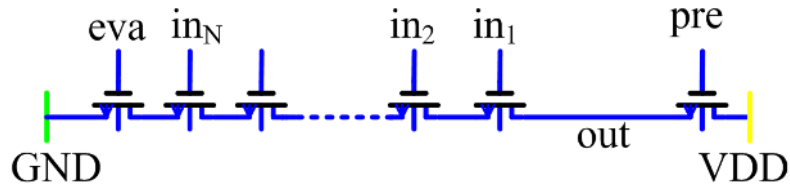
Fig. 6.   N-input dynamic NAND circuits characterized for delay distribution

and used in a higher level architectural simulator.

2.3 Architectural Simulations

The architectural simulations take as input gate delay characterizations obtained from circuit-level simulations, as shown in Fig. 1. We use a custom-written simulator called FTSIM (Fault Tolerance Simulator). FTSIM takes as input a NASIC design definition, gate timing characterizations, and defect models and simulates the operation of a large-scale design on a cycle-by-cycle basis, tracking values within the design logically.

FTSIM has several capabilities beyond simple logic simulation. It can additionally apply various types of defects to the circuit and test whether it is still operational. This is done via a Monte Carlo system. The user specifies the defect rate and how many different defect patterns to test and FTSIM simulates the system with random defect patterns and outputs the yield. For NASICs, a fairly generic defect model is used. Devices may be stuck-on, stuck-off or nanowires may be broken. A broken nanowire is equivalent to a stuck-off device, since the nanowire can no longer switch. Additional information on uniform and clustered defect models for NASICs can be found in [Wang et al. 2009; Moritz et al. 2007; Wang et al. 2005].

Additionally, it handles timing faults, specifically missed deadlines. In order to do this, it uses the gate delay characterizations that were obtained from HSPICE. For a gate with N inputs, delay characteristics are sampled from the distribution of delays obtained from the circuit simulator. Multiple trials are carried out and a different gate delay is sampled in each trial. A large gate delay due to parameter variation could cause a particular output to not evaluate to its correct value within a given clock period. This implies that a missed deadline, or in other words a timing fault occurs. For each trial FTSIM then adjusts the clock period to determine the maximum frequency at which correct final outputs may be produced. This may be a faster frequency than might be expected purely from the gate delay characteristics sampled for individual gates as timing faults may be masked either implicitly in the logic or by fault tolerance.

In this work, performance characterization was done in conjunction with device-level defect rates of up to 12%, 10 orders of magnitude higher than CMOS. We ran 1,000 trials for each defect rate which produces sufficient working circuits to give a sound idea of the performance distributions and yields. It is necessary to run the performance characterization together with non-zero defect rates as the presence of defects may cause correctly functioning circuits to run more slowly than they

Table II. Impact of physical parameters on device on-current

Parameter	% Change in I_{ON}	Correlation
Channel diameter	352.0	Positive
Underlap	181.2	Negative
Bottom oxide thickness	147.2	Positive
Gate oxide thickness	58.2	Negative
Source/drain doping	23.8	Positive
Gate diameter	16.2	Negative
Channel doping	11.7	Positive

would otherwise. For example, a faster path may contain a permanent defect that prevents its evaluation, and the performance of the design could be determined by a redundant slower path.

From architectural simulations, yield percentages and performance distributions for designs with and without fault tolerance are obtained.

3. VARIABILITY IMPACT ON XNWFET DEVICES, NASIC CIRCUITS AND SYSTEMS

In this section we present a study on the impact of variability on xnwFET on-currents (device level impact), circuit delays as well as system performance.

3.1 Device Level Impact – Variation in On-Current

At the device level, a key metric of interest in evaluating performance impact of variability is the on-current (I_{ON}) of the device[6]. This implies variation in the on-resistance leading to variations in delay and performance at higher levels.

In this study, physical parameters from Table I are varied one at a time, and the sensitivity of I_{ON} to parameter variation is measured. Parameters are varied across a $\pm 3\sigma$ range, assuming 10% standard deviation (i.e., parameters are varied from 70% to 130% of their nominal value).

Not all parameters have equal impact on I_{ON}. The percentage change in on-current between the lowest and highest sampled value for each physical parameter is shown in Table II. Channel diameter has the largest impact, with I_{ON} varying by 3.5X over a 7 nm to 13 nm range.

Fig. 7 shows how I_{ON} varies as individual parameters of the xnwFET are varied. These graphs show clearly the direction and shape of the current variation with each parameter. For four parameters, positive correlation exists between the parameter value and I_{ON}. For example, as bottom oxide thickness increases, I_{ON} increases. The substrate bias is used to deplete carriers in the channel for reducing leakage and improving threshold voltage. However, the substrate bias also reduces I_{ON} due to a shift in the threshold voltage. As the bottom oxide is made thicker, the electrostatic control exerted by the back gate bias is reduced, producing a smaller

---

[6]Both on- and off-currents are captured in the device simulations and the circuit behavioral models. High variation in off-currents can cause loss of functionality. These are similar to manufacturing defects that can be masked by fault tolerance schemes. However, the delay/frequency impact of variability comes from the on-currents, given the circuit style with evaluation of a series stack of switched-on devices.
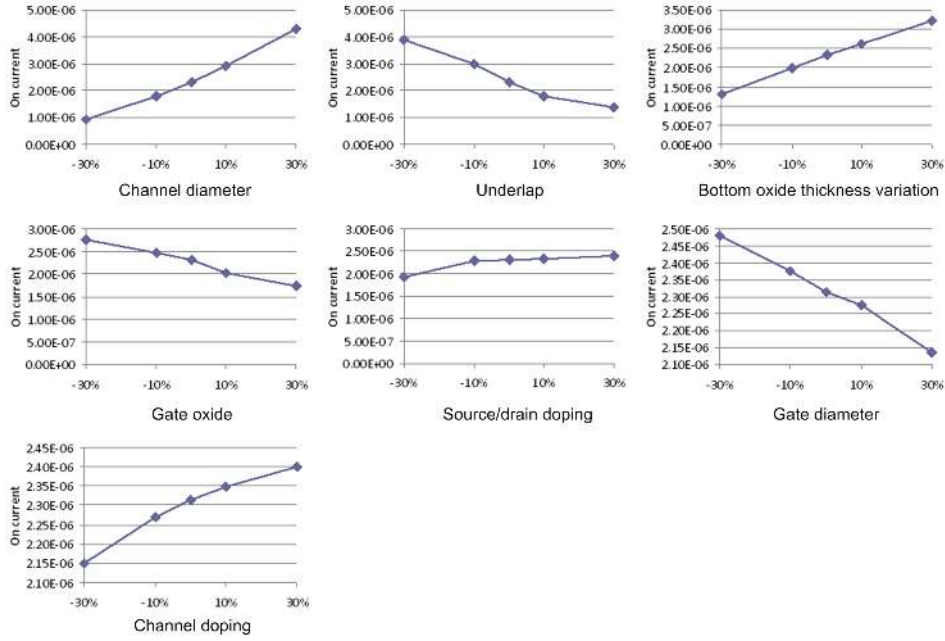
Fig. 7. Variation in I_{ON} as a function of percentage deviation of various physical parameters: Graphs are ordered in decreasing order of sensitivity

positive V_{TH} shift than expected, leading to larger I_{ON}. As channel diameter increases, the channel resistance decreases due to an increase in the cross-sectional area, leading to an increase in I_{ON}. Increasing the source and drain doping reduces the series resistance. Lastly, as channel doping increases, the short channel effects (SCE) are somewhat alleviated leading to larger I_{ON}.

The other parameters all correlate negatively with on current. Increasing the underlap increases the effective channel length, resulting in a decrease in I_{ON}. Similarly, increasing the gate oxide thickness decreases the gate capacitance and how well the gate can turn on the channel. Increasing gate diameter increases the length of the channel underneath, decreasing I_{ON}.

3.2   Circuit Level Delay Characterization

NASIC N-input dynamic NAND gates (Fig. 6) were simulated in HSPICE using unified behavioral models derived from device data. Delay characterization was done for fan-in varying between 1 and 30, which is the maximum fan-in for the NASIC WISP-0 processor, using the HSPICE Monte Carlo framework and Gaussian sampling of individual parameters. A single channel diameter value was sampled per Monte Carlo simulation for all devices, since all xnwFETs are on the same nanowire. Length-wise variation has been shown to be negligible for the nanowire lengths considered [Park et al. 2008] for a process such as VLS growth. All other parameters were varied independently for each device.

Fig. 8 plots the median of the precharge and evaluate times for fan-in varying
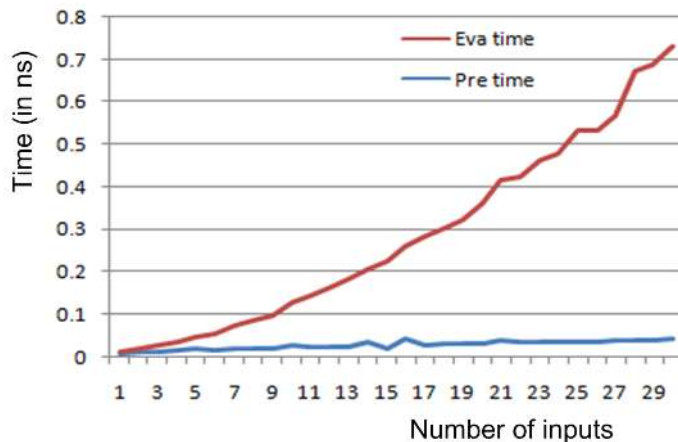
Fig. 8.   Median precharge and evaluate delays as a function of fan-in

between 1 and 30. As expected, there is an increasing trend in delay with respect to fan-in. Furthermore, evaluate time (i.e., the time taken to discharge the output node to logic '0' through a stack of xnwFETs – refer to Fig. 6) increases more rapidly than the precharge time with fan-in (evaluate time is 1.68X of precharge time for a 5-input gate and 10X for a 15-input gate). This implies that in a large scale design where high fan-in paths exist, the critical path delay will be dominated by the evaluate time for these gates. A variety of performance improvement techniques targeting shorter evaluate times in the presence of parameter variation will be discussed in Section 4.

The sensitivity of gate delay to individual parameters was also studied. We show the impact on delay for the four parameters that have maximum impact on I_{ON} at the device level. Representative results for fan-in of 15 and 30 are shown. Other fan-in gates were investigated and found to show similar trends.

Fig. 9(a) and (b) show the delay distributions for 15 input and 30 input NASIC dynamic NAND gates. The delay distribution due to channel diameter, underlap, bottom oxide and gate oxide thickness is studied. The following key observations are made -

Channel diameter has the maximum impact on delay distribution - 81% (71%) change in delay with respect to nominal for 15 (30) input gate. This is due to the high sensitivity of I_{ON} at the device level, and also due to the correlation of channel diameter across all devices for a single NASIC dynamic NAND circuit. These effects also imply a large percentage standard deviation - 18% (15%) for 15 (30) input gates - leading to a wide spread of delay values.

Underlap is negatively correlated with I_{ON}. This implies that delays will be less than nominal for shorter underlaps. Furthermore, from device level sensitivity analysis I_{ON} variation is asymmetrical with underlap. 30% negative (positive) deviation causes +74% (-43%) change in the I_{ON}. This would imply that in a circuit simulation, where underlap values for individual devices are independently sampled, the delay distribution should be left-shifted (majority of devices operating better
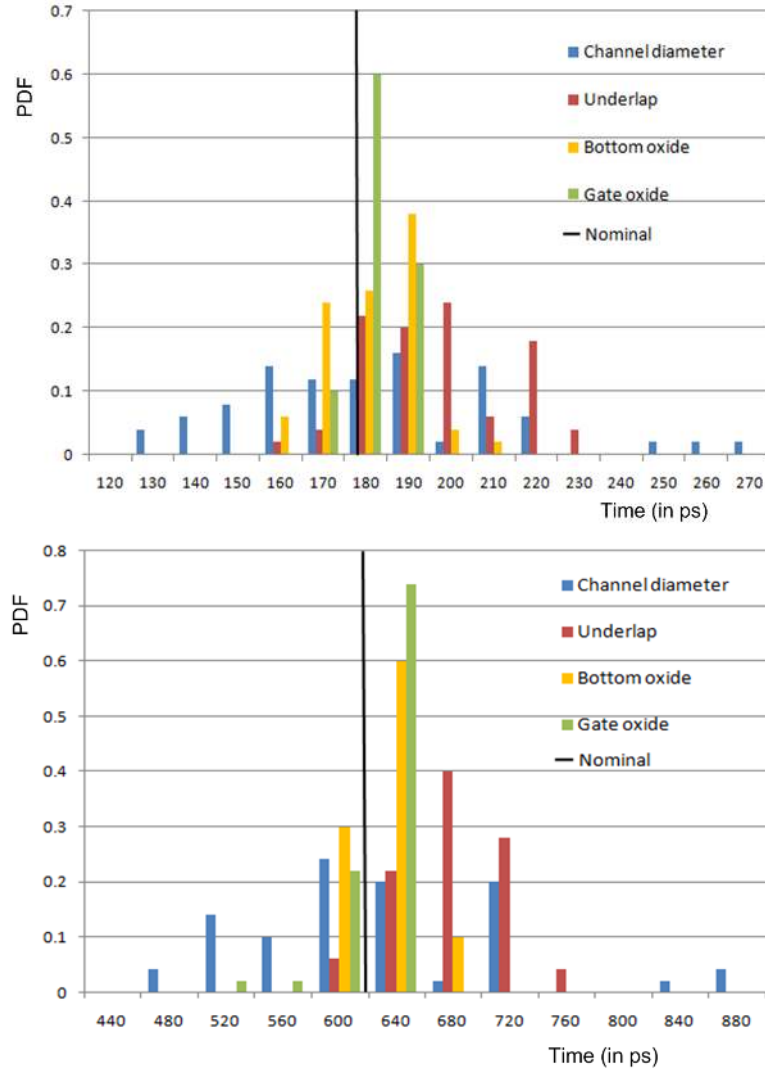
Fig. 9. Delay distributions for physical parameters with maximum impact on on-current for (top) 15 input and (bottom) 30 input NASIC dynamic NAND gates. Black line represents nominal.

than nominal). However, the opposite trend is noticed. This is because increasing trend in the I_{ON} with decreasing underlap is dominated by an increasing trend in the various capacitances as distances between terminals shrink.

The evaluation delays for **gate oxide** and **bottom oxide** are tightly distributed along the nominal, with mean values within 2% of nominal and standard deviation of 3% for the 30 input gate. Since these parameters are sampled independently, and there exist no appreciable asymmetries as compared to the underlap, variation in delays of individual devices tend to cancel out especially in higher fan-in designs.

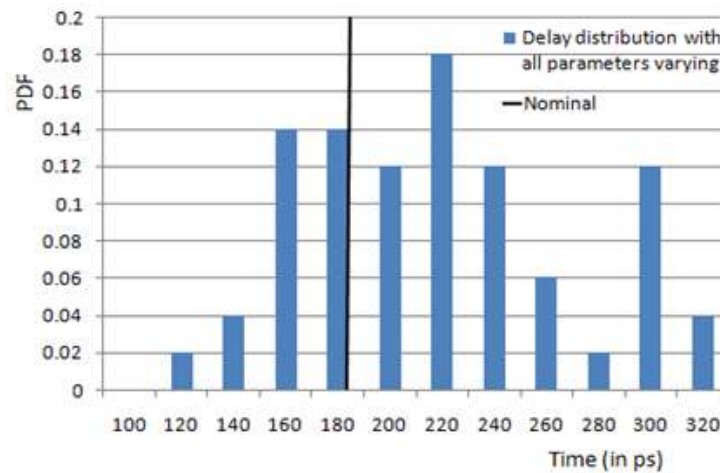Fig. 10 shows delay distributions for the 15 input NASIC dynamic NAND gate

Fig. 10. Delay distribution for 15 input gate with all parameters simultaneously varied: Nominal value is 174ps. Distribution is right-shifted due to asymmetric underlap effect

with all parameters varied simultaneously with $3\sigma=\pm30\%$. The mean is 20% higher than the nominal due to the underlap asymmetry effect that skews the distribution to the right. The same trend is observed in other fan-in gates as well. A 118% spread with respect to the nominal is observed for 15 input gates. The relative spread was found to be decreasing with increasing fan-in, as expected.

The gate delay distributions with all parameters varying for different fan-ins were modeled as gamma distributions and used in an architectural simulator to evaluate the process variation impact on a larger design.

3.3 System Level Performance Degradation

Architectural simulations of the NASIC WISP-0 processor [Wang et al. 2005; Wang et al. 2007] were carried out using the architectural simulation framework described in Fig. 1 and Section 2.3. Gate delay distributions obtained from Monte Carlo simulations of NASIC dynamic NAND gates were sampled for each gate in the design and the maximum operating frequency at which the processor functioned without missed deadlines was estimated.

The probability density function of operating frequencies obtained is plotted in Fig. 11. Also shown in the diagram is the nominal frequency for WISP-0 without any process variation. (Note: performance optimizations on device structure are currently ongoing - while we expect future devices to be considerably faster and thus the processor performance would be also much improved, it would not change the conclusions qualitatively). From the diagram, parameter variation causes performance deterioration in 67% of the samples investigated.

WISP-0 is not fully balanced with respect to timing and delay. The frequency is therefore determined entirely by a small number of high fan-in data-paths. The design consists of 221 gates in 16 stages. 11% of gates have a fan-in of 8 or higher. These are expected to form timing critical paths, depending on extent of variability in each of them. If the delays sampled from these paths are lower than nominal
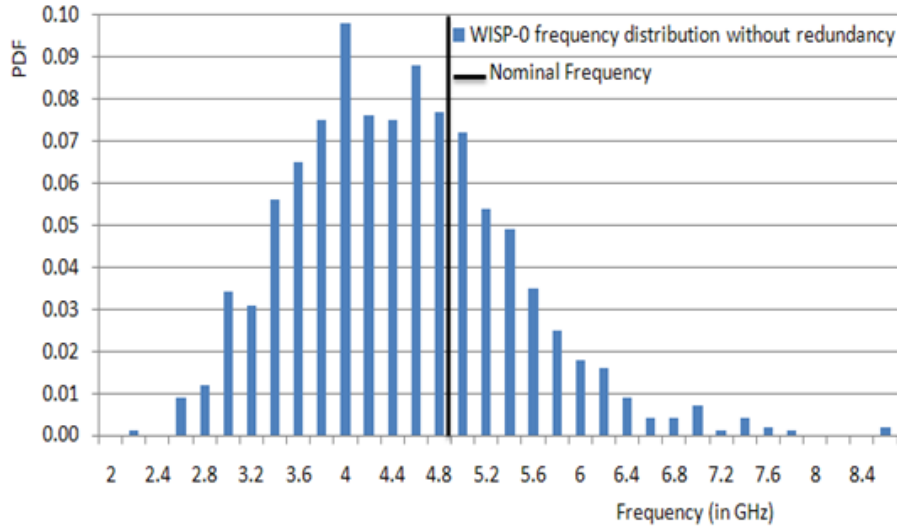
Fig. 11. Distribution of WISP-0 operating frequencies showing impact of parameter variations with no built-in fault tolerance incorporated. 67% of chips operate at frequency below nominal due to variations in device parameters.

then the performance of the entire design is not affected or may even improve. However, in designs balanced for timing, such as commercial processors where a lot of emphasis is typically put on timing path optimizations, there will be a large number of paths with similar nominal delay. The slowest path among these would determine the operating frequency. This implies that for balanced designs with process variation, a much larger fraction of chips will be slower than nominal, since data speed-up along some high fan-in paths will be entirely offset by others.

In the next section, we show how built-in fault tolerance techniques can also be used to ameliorate the effects of process variation. The key idea is to mask the timing fault due to a slower path using redundancy based schemes, leading to a majority of chips operating at a frequency higher than nominal. The challenge is, however, simultaneously managing performance impact due to redundancy.

4. TECHNIQUES TO MITIGATE IMPACT OF VARIABILITY IN NANO-DEVICE BASED SYSTEMS

While it is widely accepted that state-of-the-art CMOS designs need to deal with high levels of parameter variation, the defect rates expected are still fairly low. For example, at 65 nm, the expected defect rate is only 0.14 defects/cm^2. It may be possible to deal with parameter variation independently from defect tolerance and architectural level parameter variation resilience [Agarwal et al. 2005; Bennaser et al. 2007; Humenay et al. 2006] has been previously discussed.

In contrast, nanoscale fabrics based on self-assembly manufacturing processes are expected to have much higher defect rates (in NASICs we assume 10 orders of magnitude higher or 100s of millions to billions of defective devices per cm^2) in conjunction with high levels of parameter variation. These high defect rates

require a layered approach for fault tolerance and typically involve incorporating carefully targeted redundancy at multiple system levels, e.g., structural, circuit, and architectural levels, as has been shown previously [Moritz et al. 2007; Wang et al. 2009]. Given that this redundancy is already built in for yield purposes, it is imperative to understand whether it could be exploited to mask delay faults caused by process variation. The reasoning is that from a circuit/logic perspective both permanent defects and parameter variation lead to faulty outputs. For example, a NASIC stage is precharged to '1', and a missed deadline implies a faulty '1' at the output (i.e. the output has not evaluated to '0' within the clock period owing to parameter variability). This kind of fault may then potentially be masked by another copy of the signal which does not have the same fault. Furthermore, redundancy-based techniques may be developed that are more tailored towards process variation than defect tolerance, e.g., by trading off yield for higher performance, or the opposite, depending on system-level requirements. We can also apply redundancy non-uniformly, taking into consideration defect and process variation masking needs in specific circuits. For example, it might be more critical to have higher resilience against variations in circuit blocks that are part of the critical path, than in blocks that have plenty of timing slack. In non-critical regions, we might be able to apply techniques focusing on improving yield, without sacrificing overall performance.

4.1 Built-in Fault Tolerance for Parameter Variability

To investigate whether redundancy schemes can mitigate the impact of process variation, we ran architectural simulations of a WISP-0 NASIC streaming nanoprocessor design based on the NASIC fabric with built-in defect tolerance incorporated. The architectural simulation framework used is described in Section 2.3. The frequency distribution of the WISP-0 design was investigated for two techniques that were developed for defect tolerance: 2-way and 3-way redundancy. These redundancy techniques are based on replicating every nanowire and device on the grid, combining logic, and are optimized for yield improvement and not performance. An example of a 2-way redundancy scheme for a simple NASIC nanotile has been included in Fig. 13b. [Wang et al. 2009; Moritz et al. 2007] discuss defect tolerance for NASICs in detail.

Representative simulation results are shown in Fig. 12 . All speeds are normalized to the variation-free case with redundancy, i.e. the speed at which the design incorporating a particular redundancy scheme would run with zero process variation. This is done to isolate the impact of variation and allow comparing the distributions despite differences in their nominal frequency. The figure shows that the normalized frequency for circuits using redundancy tends to be significantly higher than without redundancy. In other words, introducing redundancy greatly increases the percentage of circuits that will operate at or above the nominal frequency in the presence of process variation.

These results attest to the potential for redundancy to counter the effects of process variation in addition to introducing defect tolerance. However, these techniques cause deterioration in the absolute frequency values owing to their higher fan-in due to replication, i.e., while the percentage of chips above nominal is much higher, the absolute nominal frequency is deteriorated (e.g. Nominal frequency for
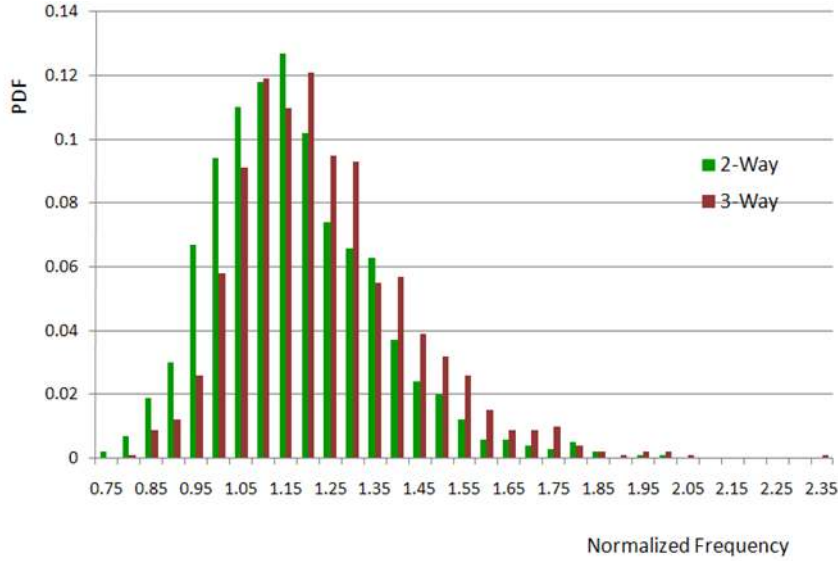
Fig. 12. Variation in maximum operating frequency of WISP-0 nanoprocessor with built-in fault tolerance

2-way (3-way) redundancy is 31% (15%) of the no-redundancy version). Therefore, new techniques tailored towards performance improvement and the associated yield tradeoffs need to be considered. A number of such techniques are presented next.

4.2 Leveraging Redundancy for Performance Improvement

New techniques based on inserting nanoscale voters at key architectural points in a design are explored in this section. These are combined with different redundancy techniques. The techniques can be classified into two categories - Biased voting schemes and FastTrack.

4.2.1 *Biased voting schemes.* Biased voting schemes leverage the property of NASIC circuits that logic '0' faults are much less likely than logic '1' related faults in high fan-in stages; therefore, we bias the voters towards logic '0'. This arises due to the unique combination of circuit and logic styles. In a NASIC dynamic stage, outputs are precharged to '1' and evaluation to '0' takes place through a series stack. Manufacturing defects can cause a faulty '0' in a high fan-in stage only when multiple devices in the stack are stuck-on and/or multiple faulty '1's are received at the input. A single correct input '0' to the stage is sufficient to turn off the series stack and prevent faulty '0' evaluation. Since performance of the system is determined by evaluation to '0' of high fan-in gates, and faulty '0s' are less likely in these gates, unbalanced or biased voting towards '0' can be leveraged for performance improvements and parameter variation resilience.

For example, consider two input blocks that are 2-way redundant, with 4 copies of a signal going to a voter biased to logic '0'. This voter outputs a logic '0' if any two of the inputs is '0', ensuring that the critical path delay is determined by the fastest two arriving '0' (which are presumed to be functionally correct). This
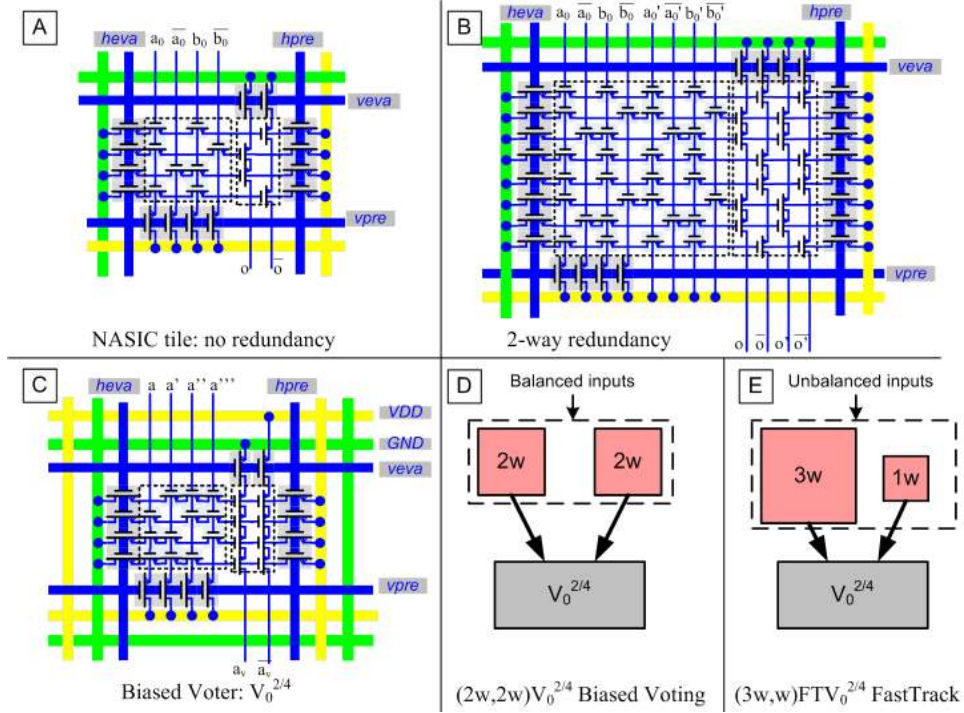
Fig. 13. NASIC fault tolerance schemes A) NASIC tile with no redundancy incorporated. B) 2-way redundant implementation of NASIC tile in A. C) Biased Voter: $V_0{}^{2/4}$ D) Block diagram showing biased voting scheme $(2w,2w)V_0{}^{2/4}$. E) FastTrack scheme $(3w,w)V_0{}^{2/4}$.

scheme is notated as $(2w,2w)V_0{}^{2/4}$ (2 input blocks with 2-way redundancy, voter biased to '0', requiring 2 of 4 inputs to be '0').

The voter and the scheme are shown in Fig. 13C and D. Voters are built like any other tile in the design using nanowire crossbars and cascaded dynamic logic. They do not have special manufacturing requirements or complex interfacing. Voter area and defects in the voter are factored into the overall effective yield calculation to assess if there is a net benefit of incorporating voters.

This biased voter is in contrast to a regular majority voter wherein a majority out of an odd number of input signals would need to be '0' to vote logic '0'. It is an instance of a plurality voter that requires a plurality of '0's (in this case two '0's) to output '0'.

4.2.2 *FastTrack: Leveraging asymmetric delay paths.* FastTrack schemes employ biased voters in conjunction with unbalancing the redundancy in each of the voter input blocks. The intuition behind these is to have some inputs (in some of the blocks) arrive faster than others to address the increased delay due to parameter variation and the delay due to the added redundancy itself. The combination of unbalanced input blocks and biasing define a variety of new techniques depending on the input configuration, redundancy levels for each input, voting type, and biasing applied. If a lesser redundancy block outputs a '0', there might be no need

to wait for slower outputs from higher redundancy blocks (the evaluate time for the lesser redundancy block, and consequently the critical path delay will be lower as shown in Fig. 8).

For example, along these ideas, a FastTrack voting scheme example and notation is shown in Fig. 13E . Note that the redundancy applied to input blocks to the voter is unbalanced. The figure implies a voter and two input blocks, one having 3-way redundancy, and one with no redundancy. Furthermore, the voter is biased towards '0' (see subscript on FTV): if 2 out of 4 inputs coming from the three blocks are '0' the output is voted as logic '0'. The combination of input organization and voter type results in a voting scheme notated, e.g., as $(3w, w)FTV_0^{2/4}$.

The notation used can be generalized to voting schemes with unbalanced redundancy in their input blocks and biasing in general. If there is no biasing of output the subscript and superscript after the voter type can be omitted. In the next section, yield-area-performance tradeoffs are discussed. This is followed by results for a variety of schemes based on the above concepts.

4.3 Design Choices and Tradeoffs

Using the framework for *Biased* voting and *FastTrack*, a variety of techniques can be developed by varying both the input configuration and voter bias. Techniques may be tailored to meet yield and performance targets depending upon manufacturing constraints as well as system level requirements. For example, a manufacturing process with high defect rates would require input blocks to have a higher level of redundancy incorporated for acceptable yield. The voter should also use more inputs to determine the final output. On the other hand, if manufacturing processes can be tailored to smaller defect rates FastTrack schemes can aggressively target performance: input blocks would need less redundancy and voting techniques would require a smaller number of inputs.

To illustrate this, consider two FastTrack schemes: $(3w,2w,1w)FTV_0^{1/6}$ and $(3w,2w)FTV_0^{2/5}$. In the first case, there are 3 input blocks with varying levels of redundancy. The voting block outputs logic '0' if one of the 6 inputs is '0'. Typically, it is expected that the no-redundancy version of the input would generate the fastest '0' based on . Therefore, in the absence of defects, the performance of the circuit is expected to be determined by the 1-way input block. However, in the presence of defects, the fastest arriving '0' may be faulty. Consequently, the voter propagates an incorrect value to the next stage resulting in yield losses. In fact, even for relatively small defect rates, faulty '0's in no-redundancy schemes are not uncommon. Therefore, the yield of this FastTrack scheme is expected to be relatively low with defects considered.

For the $(3w,2w)FTV_0^{2/5}$, the voter waits for the fastest two arriving '0's. This scheme offers significant benefits in terms of yield: i) voting decision is based on more input signals and ii) faulty '0's are much less likely in 2-way and 3-way redundancy schemes compared to 1-way (no-redundancy), since twice/thrice as many xnwFETs on a single nanowire would need to be simultaneously stuck-on to evaluate the stack. However, there is an area overhead (area of nanowire grid scales quadratically, peripheral microwire area scales linearly) and the performance of this design will also be degraded by the larger fan-in and linear increase in capacitance for switching more devices.

Similarly, other schemes can be developed trading off yield and performance. For example, a $(3w)V_0^{1/3}$ biased-voting scheme is expected to have an even better yield because of the higher redundancy in the input block. Note that the FT notation is dropped for this biased voting scheme since asymmetric path delays are not exploited. Consequently, the performance is expected to be lesser since it will be determined purely by the 3-way input block.

In the next section, we evaluate an ensemble of FastTrack and Biased schemes that were developed using different input configurations and voting bias and present further insight into the tradeoffs involved.

4.4   Results

FastTrack, biased voting, and other defect tolerance schemes previously developed for the NASIC fabric were evaluated for three key system level metrics: effective yield, normalized mean performance and normalized performance*effective yield (PEY) product.

Effective yield is defined as (Overall Yield)/Area. It captures the tradeoffs between area overhead and yield and represents the number of functional chips obtained for a fixed area (e.g. the number of functional chips obtained per wafer in a scalable manufacturing process).

The normalized mean performance across a range of techniques represents the average frequency across all architectural simulations normalized to the mean operating frequency for the slowest technique. This metric captures the effective improvement in performance for a design incorporating various redundancy schemes as compared to the slowest scheme. This metric is unitless.

The PEY metric is a composite one capturing tradeoffs between performance, area and yield. It is the product of the normalized performance with the effective yield and is useful for designs where both performance and effective yield constraints are of equal importance, or when specific yield and performance targets need to be simultaneously met.

The techniques studied include redundancy schemes optimized for yield: 2-way (2w), 3-way (3w), 4-way (4w) redundancy and majority voting ((2w)6MR) as well as different Biased voting and FastTrack schemes based on concepts described in previous subsections.

4.4.1   *Effective yield.* Fig. 14 shows the effective yield for the techniques considered. As expected, traditional defect-tolerance schemes such as 2-way and 3-way redundancy have the best yields across the range of defect rates studied. In these simulations, we assume a uniform distribution of manufacturing defects in the design. The dominant defect mode is stuck-on nanowire FETs arising out of ion implantation and/or metallization steps in the process flow. Diffusion of ions into channel regions or mask misalignment can cause nanowire channels to be incorrectly erased leading to always-on FETs.

2-way redundancy is best in terms of effective yield at lower defect rates owing to a much smaller area overhead (44% of 3-way redundancy); however beyond a 6% defect rate its yield drops implying that a 3-way technique is needed. The $(3w)V_0^{1/3}$ Biased voting scheme tracks, however, very closely with the 3-way redundancy scheme, implying that faulty '0's (which would be incorrectly propagated by
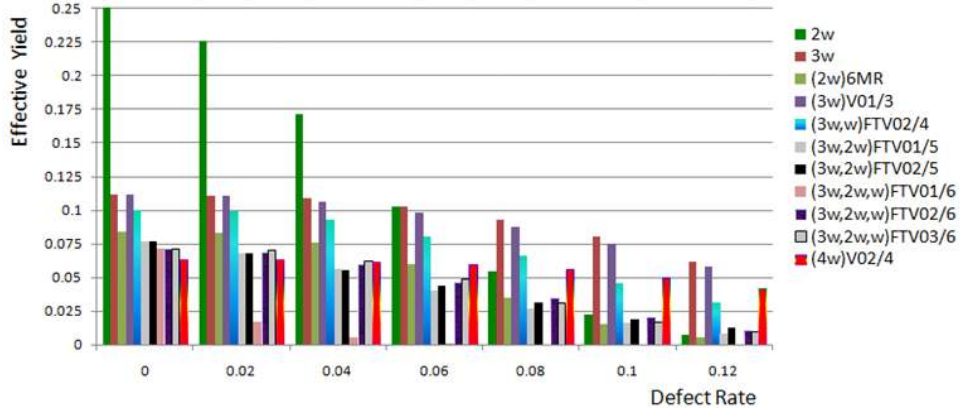
Fig. 14.   Effective yield vs. defect rate for redundancy, biased voting and FastTrack schemes

the voters) are relatively small for larger fan-in gates even at very high defect rates as expected. The FastTrack scheme with the best effective yield is $(3w,w)FTV_0^{2/4}$. While the no-redundancy version is prone to faulty '0's, the voter still requires 2 input versions to be '0' to propagate the value. Therefore, correct circuit functionality can be obtained if the inputs from the 3-way block are correct. On the other hand, a $(3w, 2w, w)FTV_0^{1/6}$ performs poorly. At lower defect rates its area overhead reduces the effective yield. At higher defect rates the yield is negligible since the voter propagates faulty '0's from the no-redundancy version and there is no other fall-back mechanism to mask it.

4.4.2   *Normalized Performance.* Fig. 15 shows the normalized performance for the techniques discussed and in essence capture a different perspective. The techniques are normalized against a simple 3-way redundancy scheme, which is the slowest of the techniques evaluated. The graphs show relative speed-up for the other techniques in relation to the slowest.

The architectural simulation framework described in Section 2.3 is used. A library of gate delay distributions is obtained from circuit-level simulation. The architectural simulator samples the gate delay distribution to ascertain if outputs evaluate correctly within a given clock duration. Otherwise, the clock period is lengthened and the minimum clock period (best frequency) at which outputs evaluate to their correct values is ascertained.

This data omits performance numbers for cases that do not yield. It is observed that in general the normalized performance reduces with increasing defect rates. As defects become more prevalent, faults on FastTrack designs will imply that paths with lower levels of redundancy may not switch correctly. Hence, the frequency may be determined by slower paths with higher degree of redundancy. For example, with the $(3w,2w)FTV_0^{2/5}$ scheme, we see that the performance is comparable to simple 2-way redundancy at lower defect rates. However, as defect rates increase, the normalized mean performance reduces, implying that the 2-way schemes become faulty and in many cases, the performance is determined by slower, albeit more defect tolerant 3-way blocks.
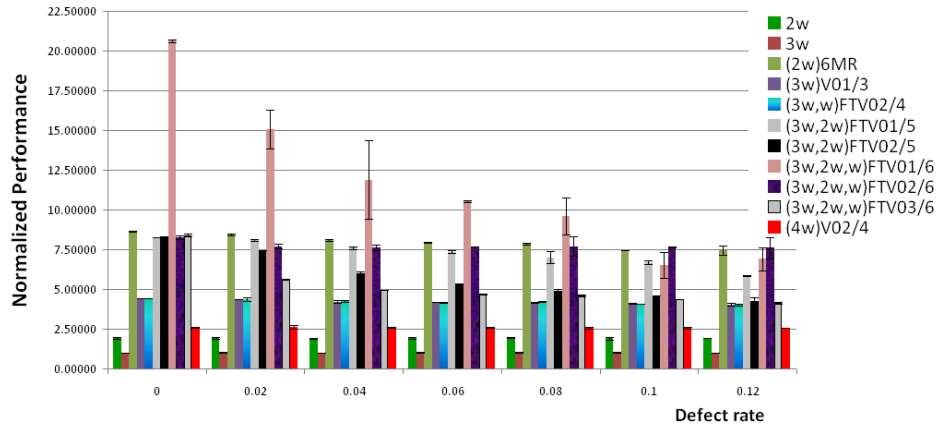
Fig. 15. Normalized performance vs. defect rate for redundancy, biased voting and FastTrack schemes

Somewhat unexpectedly, the 2-way majority voting scheme has a good normalized performance of around 7.5 for all defect rates. This may be attributed to the fact that irrespective of the defect rate, the critical path in a correctly functioning sample is always determined by a 2-way redundant path. While this is true even for simple 2-way redundancy, the key difference is that in the majority voting scheme, there are many more paths to choose from, implying that the delay is not sensitive to any one slow path. Similarly, the $(3w)V_0^{1/3}$ biased voting scheme has a consistent 4X improvement in performance compared to the 3-way redundancy scheme.

The $(3w, 2w, w)FTV_0^{1/6}$ scheme is fastest up to an 8% defect rate since its performance is determined by the fast no-redundancy input block. However, as previously mentioned, this scheme has very poor effective yield even for small defect rates. At higher defect rates, the $(3w, 2w, w)FTV_0^{2/6}$ scheme has the best performance. It can leverage the fast no-redundancy block when it outputs correct '0's, while also providing resilience at higher defect rates due to the redundant input blocks.

4.4.3 *Normalized Performance-Effective Yield (PEY) Product.* PEY product results are shown for the various techniques in Fig. 16. At zero defect rate, the $(3w, 2w, w)FTV_0^{1/6}$ works best owing to its large performance advantage (20X faster than the 3-way redundancy scheme). However, its PEY product falls off rapidly with increasing defect rates owing to deterioration in yield. This implies that while the $(3w, 2w, w)FTV_0^{1/6}$ technique has the best performance, it may not be suitable for cases where both yield and performance targets need to be met.

The 2-way majority voting scheme has the best PEY products for lower defect rates, since this scheme has good effective yield as well as good normalized performance in this range as discussed previously. However, at higher defect rates the benefits of this scheme drop off due to the reduced effective yield. The $(3w)V_0^{1/3}$ Biased voting scheme has a consistent 4X improvement in performance over the slowest scheme, and due to the triplication of signals can also handle high levels
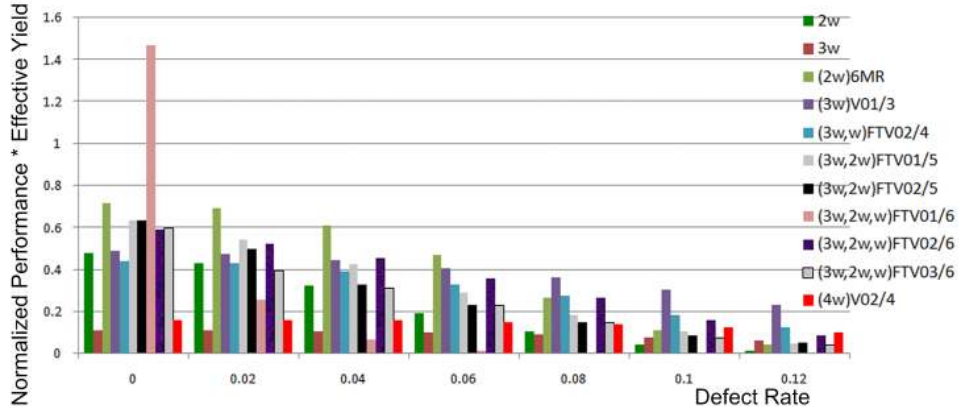
Fig. 16. Normalized performance * Effective yield vs. defect rate for redundancy, biased voting and FastTrack schemes

of defects. While this scheme is consistently good across all defect rates, it is the best scheme for defect rates higher than 6%. Among the FastTrack schemes, the $(3w,2w,w)FTV_0{}^{2/6}$ does also fairly well primarily due to having the best performance for high defect rates.

A variety of manufacturing processes are currently being explored to achieve high density integrated nanosystems. For example, for nanowire alignment some representative approaches being pursed include *ex-situ* crystallographic etch and transfer [Moritz et al. 2011], block-copolymer [Cheng et al. 2006] and other pattern transfer techniques [Shin and Chui 2011]. Depending on the specific processes used, different trade-offs across density benefits and requirements for mitigation of defects/variability are needed. However, for the foreseeable future, state-of-the-art unconventional manufacturing techniques will need to address defect masking in conjunction with parameter variability to achieve integrated systems. Therefore in this research study, a range of defect rates and conservative estimates for parameter variability are addressed. A suite of FastTrack and Biased voting techniques has been presented and implications analyzed. The choice of any one specific mitigation technique will ultimately depend on manufacturing and design-specific density/yield/performance trade-offs.

5. CONCLUSIONS

A novel methodology for integrated device-circuit-architectural explorations for analyzing the impact of parameter variability in nano-device based computing systems was developed. The methodology builds on accurate 3D physics based simulations of device structure to capture variations in on-current as a function of physical parameters. Circuit and architectural simulations evaluate the impact of this variability on gate delay and system level performance respectively.

The methodology was evaluated on the NASIC computational fabric with xnwFETs, NASIC dynamic NAND gates and a processor design. Key sources of variation at the device level such as channel diameter were identified and sensi-

tivity of I_{ON} was evaluated. I_{ON} may vary by up to 3.5X with variations in the channel diameter and by up to 1.5X with gate underlap.

Impact of device parameter variation on higher design levels was found to be significant, with simulations of a stream processor design showing 67% of chips operating at frequencies below nominal. As redundancy based techniques are necessary for providing resilience against permanent defects, they may be tailored to address variability in conjunction with defects.

An ensemble of techniques to improve performance focusing on biased voting and 'fast tracking' signals was developed. FastTrack techniques show up to 7.5X performance improvement compared to more traditional redundancy schemes even at higher defect rates. In the absence of defects, a FastTrack scheme can be up to 22X faster than a traditional redundancy scheme.

Biased voting schemes such as $(3w)V_0^{1/3}$ show good balance over effective yield and performance for a wide range of defect rates. The normalized performance * effective yield (PEY) metric for this scheme was found to be 3.8X better than a highly defect resilient but slow redundancy scheme even at 12% defect rate. Among the FastTrack schemes, $(3w,w)FTV_0^{2/4}$ was found to be the best in terms of PEY product, with a 2.1X improvement over the baseline for 12% defect rates. $(3w,2w,w)FTV_0^{2/6}$ was the best performing in terms of speed at higher defect rates.

This array of techniques provides a framework for design space explorations towards simultaneously achieving yield and performance goals, as opposed to conventional techniques focused on redundancy alone. Depending on design requirements and defect rates, variation expected due to manufacturing, a suitable biased voting/FastTrack technique may be applied.

Acknowledgment

REFERENCES

AGARWAL, A., PAUL, B. C., MAHMOODI, H., DATTA, A., AND ROY, K. 2005. A process-tolerant cache architecture for improved yield in nanoscale technologies. *IEEE Trans. Very Large Scale Integr. Syst. 13,* 1, 27–38.

BENNASER, M., GUO, Y., AND MORITZ, C. A. 2007. Designing memory subsystems resilient to process variations. In *Proceedings of the IEEE Computer Society Annual Symposium on VLSI.* IEEE Computer Society, 357–363.

CHEN, Z., APPENZELLER, J., LIN, Y., SIPPEL-OAKLEY, J., RINZLER, A. G., TANG, J., WIND, S. J., SOLOMON, P. M., AND AVOURIS, P. 2006. An integrated logic circuit assembled on a single carbon nanotube. *Science 311,* 5768 (Mar.), 1735.

CHENG, J., ROSS, C., SMITH, H., AND THOMAS, E. 2006. Templated Self-Assembly of block copolymers: Top-Down helps Bottom-Up. *Advanced Materials 18,* 19, 2505–2521.

COLLIER, C. P., WONG, E. W., BELOHRADSK, M., RAYMO, F. M., STODDART, J. F., KUEKES, P. J., WILLIAMS, R. S., AND HEATH, J. R. 1999. Electronically configurable Molecular-Based logic gates. *Science 285,* 5426 (July), 391–394.

CUI, Y., DUAN, X., HU, J., AND LIEBER, C. M. 2000. Doping and electrical transport in silicon nanowires. *The Journal of Physical Chemistry B 104,* 22 (June), 5213–5216.

Cui, Y., Lauhon, L. J., Gudiksen, M. S., Wang, J., and Lieber, C. M. 2001. Diameter-controlled synthesis of single-crystal silicon nanowires. *Applied Physics Letters 78,* 15, 2214.

Gojman, B. and DeHon, A. 2009. VMATCH: using logical variation to counteract physical variation in bottom-up, nanoscale systems. In *Field-Programmable Technology, 2009. FPT 2009. International Conference on.* 78–87.

He, R., Gao, D., Fan, R., Hochbaum, A. I., Carraro, C., Maboudian, R., and Yang, P. 2005. Si nanowire bridges in microtrenches: Integration of growth into device fabrication. *Advanced Materials 17,* 17, 2098–2102.

Heath, J. R. 2008. Superlattice nanowire pattern transfer (SNAP). *Accounts of Chemical Research 41,* 12 (Dec.), 1609–1617.

Humenay, E., Tarjan, D., and Skadron, K. 2006. Impact of parameter variations on multi-core chips. *IN WORKSHOP ON ARCHITECTURAL SUPPORT FOR GIGASCALE INTEGRATION.*

ITRS. 2009a. International technology roadmap for semiconductors - table lith3.

ITRS. 2009b. International technology roadmap for semiconductors - table lith5b.

Liu, Y., Chung, J., Liu, W. K., and Ruoff, R. S. 2006. Dielectrophoretic assembly of nanowires. *The Journal of Physical Chemistry B 110,* 29 (July), 14098–14106.

Lu, W. and Lieber, C. M. 2006. Semiconductor nanowires. *Journal of Physics D: Applied Physics 39,* 21, R387–R406.

McNeill, D. W., Bhattacharya, S., Wadsworth, H., Ruddell, F. H., Mitchell, S. J. N., Armstrong, B. M., and Gamble, H. S. 2007. Atomic layer deposition of hafnium oxide dielectrics on silicon and germanium substrates. *Journal of Materials Science: Materials in Electronics 19,* 2, 119–123.

Mehrotra, S. R. and Roenker, K. 2007. Process variation study for silicon nanowire transistors. *Microelectronics and Electron Devices, 2007. WMED 2007. IEEE Workshop on,* 40–41.

Melosh, N. A., Boukai, A., Diana, F., Gerardot, B., Badolato, A., Petroff, P. M., and Heath, J. R. 2003. Ultrahigh-Density nanowire lattices and circuits. *Science 300,* 5616, 112 –115.

Moritz, C., Wang, T., Narayanan, P., Leuchtenburg, M., Guo, Y., Dezan, C., and Bennaser, M. 2007. Fault-tolerant nanoscale processors on semiconductor nanowire grids. *Circuits and Systems I: Regular Papers, IEEE Transactions on 54,* 11 (nov.), 2422 –2437.

Moritz, C. A., Narayanan, P., and Chui, C. O. 2011. Nanoscale application-specific integrated circuits. In *Nanoelectronic Circuit Design,* N. K. Jha and D. Chen, Eds. Springer New York, 215–275.

Narayanan, P., Leuchtenburg, M., Wang, T., and Moritz, C. A. 2008. CMOS control enabled Single-Type FET NASIC. In *Proceedings of the 2008 IEEE Computer Society Annual Symposium on VLSI.* IEEE Computer Society, 191–196.

Narayanan, P., Moritz, C. A., Park, K. W., and Chui, C. O. 2009. Validating cascading of crossbar circuits with an integrated device-circuit exploration. In *Nanoscale Architectures, IEEE International Symposium on.* IEEE Computer Society, 37–42.

Narayanan, P., Park, K. W., Chui, C. O., and Moritz, C. 2009. Manufacturing pathway and associated challenges for nanoscale computational systems. In *Nanotechnology, 2009. IEEE-NANO 2009. 9th IEEE Conference on.* 119 –122.

Park, W. I., Zheng, G., Jiang, X., Tian, B., and Lieber, C. M. 2008. Controlled synthesis of Millimeter-Long silicon nanowires with uniform electronic properties. *Nano letters 8,* 9 (Sept.), 3004–3009.

Picciotto, C., Gao, J., Yu, Z., and Wu, W. 2009. Alignment for imprint lithography using nDSE and shallow molds. *Nanotechnology 20,* 25, 255304.

Ritala, M. and Leskela, M. 2004. Atomic layer deposition. *High-K Gate Dielecrics,* 17–64.

Shan, Y. and Fonash, S. J. 2008. Self-Assembling silicon nanowires for device applications using the Nanochannel-Guided Grow-in-Place approach. *ACS Nano 2,* 3 (Mar.), 429–434.

Shin, K.-S. and Chui, C. O. 2011. Aligned assembly of nanowire arrays with intrinsic control. In *to be presented in the TMS Emerging Materials Conference.*

SNIDER, G. S. AND WILLIAMS, R. S. 2007. Nano/CMOS architectures using a field-programmable nanowire interconnect. *Nanotechnology 18,* 3, 035204.

STRUKOV, D. B. AND LIKHAREV, K. K. 2005. CMOL FPGA: a reconfigurable architecture for hybrid digital circuits with two-terminal nanodevices. *Nanotechnology 16,* 6, 888–900.

STRUKOV, D. B. AND LIKHAREV, K. K. 2007. Reconfigurable hybrid CMOS/Nanodevice circuits for image processing. *IEEE Transactions on Nanotechnology 6*, 696–710.

SVERDLOV, V., WALLS, T., AND LIKHAREV, K. 2003. Nanoscale silicon MOSFETs: a theoretical study. *Electron Devices, IEEE Transactions on 50,* 9, 1926–1933.

URAL, A., LI, Y., AND DAI, H. 2002. Electric-field-aligned growth of single-walled carbon nanotubes on surfaces. *Applied Physics Letters 81,* 18, 3464.

VIJAYAKUMAR, P., NARAYANAN, P., KOREN, I., MANI KRISHNA, C., AND MORITZ, C. A. 2011. Impact of nanomanufacturing flow on systematic yield losses in nanoscale fabrics. In *2011 IEEE/ACM International Symposium on Nanoscale Architectures (NANOARCH)*. IEEE, 181–188.

WANG, T., BEN-NASER, M., GUO, Y., AND MORITZ, C. A. 2005. Wire-streaming processors on 2-D nanowire fabrics. *NANOTECH 2005, NANO SCIENCE AND TECHNOLOGY INSTITUTE*.

WANG, T., NARAYANAN, P., AND MORITZ, C. A. 2007. Combining 2-level logic families in grid-based nanoscale fabrics. In *Proceedings of the 2007 IEEE International Symposium on Nanoscale Architectures*. IEEE Computer Society, 101–108.

WANG, T., NARAYANAN, P., AND MORITZ, C. A. 2009. Heterogeneous Two-Level logic and its density and fault tolerance implications in nanoscale fabrics. *IEEE Transactions on Nanotechnology 8,* 1, 22–30.

WHANG, D., JIN, S., AND LIEBER, C. M. 2003. Nanolithography using hierarchically assembled nanowire masks. *Nano Letters 3,* 7 (July), 951–954.

WONG, H. P., TAUR, Y., AND FRANK, D. J. 1998. Discrete random dopant distribution effects in nanometer-scale MOSFETs. *Microelectronics and Reliability 38,* 9 (Sept.), 1447–1456.

XIONG, X., JABERANSARI, L., HAHM, M. G., BUSNAINA, A., AND JUNG, Y. J. 2007. Building highly organized Single-Walled-Carbon-Nanotube networks using Template-Guided fluidic assembly13. *Small 3,* 12, 2006–2010.