

# Design of 8T-Nanowire RAM Array

Vikram Suresh, Akshaya Shanmugam, Lekshmi Krishnan, Avinash Bijjal, Mostafizur Rahman and Andras Moritz  
Dept. of Electrical and Computer Engineering, University of Massachusetts, Amherst, USA

**Abstract**— SRAM based memory blocks constitute a major part of state-of-art processor architectures. Increasing complexity and variation in nanometer CMOS fabrication has prompted exploration of memory circuits based on emerging nanofabrics. In this work, we propose a new 8T-Nanowire based RAM (8T-NWRAM) circuit for high density memory arrays. The design is based on N<sup>3</sup>ASIC, a nanofabric using combination of crosspoint nanowire FETs and integration with metal interconnects. The layout implementation is optimized to reduce bitline load and achieve high performance. The upper bound on bitcell area is 0.1 $\mu$ m<sup>2</sup>, which is 50% more compared to conventional 6T-SRAM. However, both performance and leakage are significantly improved. Circuit simulation using N<sup>3</sup>ASIC 2C-xnwFET device models show improvements of >2X in read time and ~4X in write time compared to high performance SRAM. The average leakage power at 0.2nW is ~20X smaller compared to high performance SRAM. In comparison with existing 10T-NWRAM, the 8T-NWRAM provides a twofold improvement in read time and ~46% faster write time, but at the expense of ~30% increase in area and average leakage power. Thus, 8T-NWRAM is a viable alternative to 10T-NWRAM for performance vs area/leakage design requirement. We also study the impact of supply noise induced clock jitter on NWRAM circuits and propose adequate design margin to ensure bitcell stability.

**Keywords**- Memory, Nanowire, N<sup>3</sup>ASIC

## I. INTRODUCTION

Embedded memory constituted of SRAM is a critical component in modern day microprocessors and system-on-chip. As CMOS fabrication is approaching physical limits, SRAM array yield and reliability is of major concern [1][2]. Apart from issues in optical lithography, CMOS scaling has brought about numerous other challenges in SRAM design related to increasing leakage current, low noise margin, reduced stability and increased sensitivity to device aging [3]. Various novel devices and nanofabrics are being proposed to succeed in the post-CMOS era. The new paradigm of closely integrating circuit, architecture and lower complexity fabrication process has increased popularity in research of structured nanofabrics [4][5]. Carbon Nanotubes [6] and Silicon Nanowires [7] are seen as promising candidates for structured nanofabrics architectures. Designing memory circuits with nanofabric reduces complexity in fabrication due to the uniform grid based layout. The dynamic circuit based approach provides more control over leakage current. A 10T Nanowire RAM (NWRAM) has been proposed in [8][9] with significant improvement in performance and leakage compared to CMOS. The underlying fabric for NWRAM is N<sup>3</sup>ASIC [10].

N<sup>3</sup>ASIC is based on a dense Nanowire fabric with a kind of a dual-channel crosspoint FETs (2C-xnwFETs) [10] and an integration mindset to enable conventional CMOS interconnect fabrication process for signal routing. The 2C-xnwFET is fabricated by depositing Metal gates on silicon nanowire followed by self-aligning ion implantation to create enhancement mode n<sup>+</sup>/p/n<sup>+</sup> junction. The regular nanowire grid in conjunction with other fabric aspects provide the advantage of enabling simpler fabrication process with lesser overlay requirements while being compatible with design rules for integration with conventional metal stack. Using metal interconnects requires no changes in existing fabrication process and provides flexibility of connecting spatially separated Nanowire logic tiles.

In this work, we present a modified version of NWRAM cell with 8 xnwFETs. The 8T-NWRAM achieves better performance with minimal increase in area and power. Since stability is an important criterion in memory circuit design, we present a brief study of stability issues in NWRAM bitcells. The rest of this paper consists of a brief overview of existing 10T-NWRAM in section II; The proposed 8T-NWRAM circuit and array architecture in section III; Implementation and results in section IV; Discussion of NWRAM stability in section V; followed by conclusion and proposed future work in section VI.

## II. N<sup>3</sup>ASIC AND 10T-NWRAM

A 10T-NWRAM circuit has been reported with promising power-performance results compared to conventional CMOS circuits [8][9]. The 10T-NWRAM was implemented with two cross-coupled dynamic NAND gates to store the true and complementary values, fig. 1. Two sets of pre-charge and evaluate signals are used to write a bit into the cell and a separate read logic is implemented to retrieve the stored value. During write, the *bl* signal is either synchronized with *xpre*, *xeva* or *ypre*, *yeva* depending on the bit is written. To retain the value, the dynamic cell is restored before the charge is leaked from the node. During read, the pre-charged *bl* is either pulled down to V<sub>ss</sub> or remains high based on the stored value. The envisioned fabric level implementation is shown in fig. 2.

The memory array implementation of 10T-NWRAM is as shown in fig. 3. The pre-charge and evaluate signals are routed in Metal layer 1, while the bitline is routed along the horizontal direction in Metal 2. The rectangular shape of the bitcell increases bitline capacitance for large number of memory columns. In this work we propose a modified circuit that uses 8 2C-xnwFETs. The array architecture consequently is enhanced to reduce bitline loading.

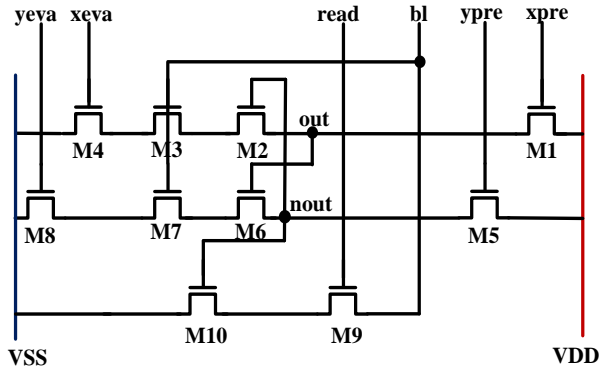


Figure 1: 10T-NWRAM Circuit [8]

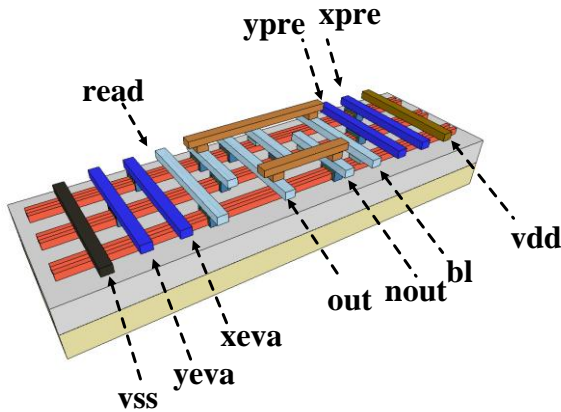


Figure 2: Implementation of 10T-NWRAM in  $N^3$ ASIC [8]

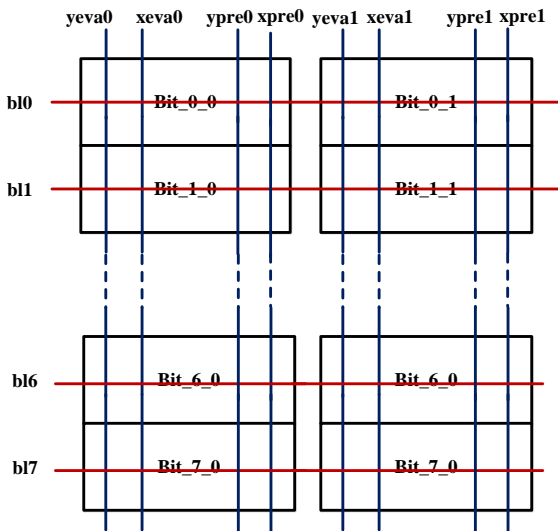


Figure 3: 10T-NWRAM Memory Array

### III. 8T NWRAM: CIRCUIT AND ARRAY

The 8T Nanowire RAM circuit is as shown in fig. 4. The storage element is implemented with a cross-coupled NAND structure using two 2C-xnwFETs, M2 and M5. Complimentary states are stored in out and nout.

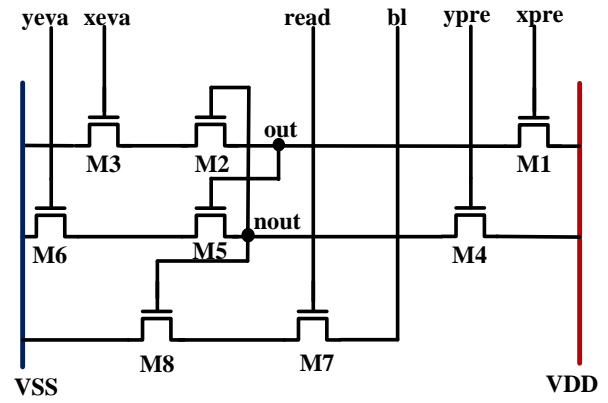


Figure 4: 8T-NWRAM Circuit

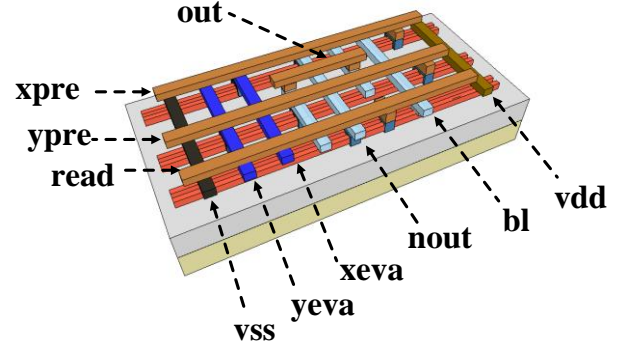


Figure 5: Implementation of 8T-NWRAM in  $N^3$ ASIC

Unlike 10T NWRAM, the bitline is not used to write into the bitcell. This reduces two transistors in the out and nout stack. A pre-charge and evaluate transistor in each path controls the write and memory restore operations through non-overlapping clock signals xpre, ypre, xeve and yeva. A two transistor stack, M7 and M8, pre-charged by bitline forms the single ended read circuit. The implementation of 8T NWRAM in  $N^3$ ASIC is shown in fig. 5.

#### A. Write

The write operation is performed in two cycles of pre-charge and evaluate of nodes out and nout. To write a bit "1", the out node is pre-charged using xpre clock. However, the xeve signal is gated to hold the state of out at 1. This is followed by pre-charge and evaluation of nout, using non-overlapping clocks ypre and yeva to discharge nout to 0. To write a 0 in out, while xeve is asserted after pre-charging out, yeva is gated after pre-charging nout to store a state 1, fig. 6.

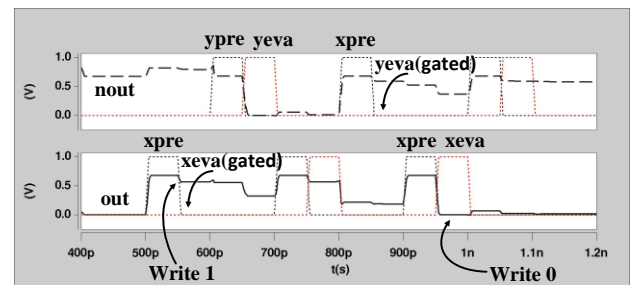


Figure 6: 8T-NWRAM Write Operation

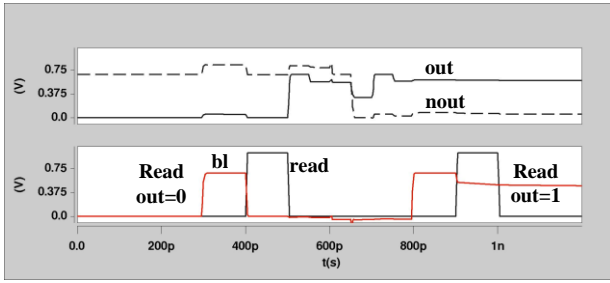


Figure 7: 8T-NWRAM Read Operation

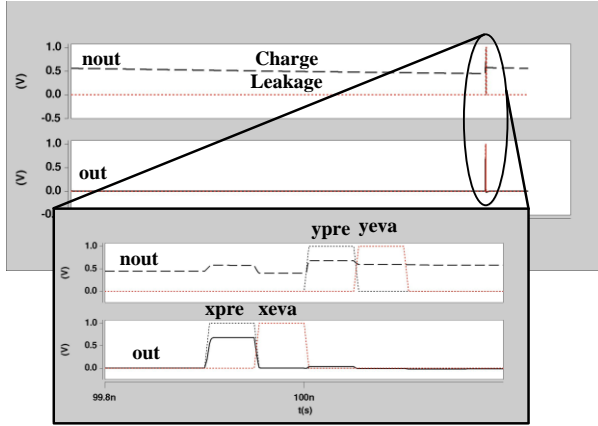


Figure 8: 8T-NWRAM Restore Operation

### B. Read

During read operation the bitline  $bl$  is pre-charged to VDD. The pre-charge is released and the  $read$  signal is asserted. Depending on the state stored in  $nout$ , transistor M8 is turned ON or OFF to read the bit stored on  $out$  through the bitline. Fig. 7 shows the read operation of 8T-NWRAM.

### C. Restore

The NWRAM circuit is based on NMOS devices implemented in the form of 2C-xnwFET's. As a result, the cross-coupled structure does not have a static source of charge. The bit-cell potentially loses state due to charge leakage in nodes  $out$  and  $nout$ . NWRAM requires frequent restore cycles to re-charge the nodes similar to Dynamic RAM (DRAM). The restore cycle consists of a sequence of pre-charge and evaluate operations on  $out$  and  $nout$ . Depending on the state of the bit-cell either one of the nodes automatically gets pre-charged to VDD, while the other node is discharged during evaluate operation. Unlike DRAM restore, no additional read operation is necessary for restore. Charge leakage over time and restore operation are shown in fig. 8. The physical layout of 8T-NWRAM bit cell is as shown in fig. 9. The pre-charge and read signals are routed in Metal 2. Evaluate and bitline signals are routed in the orthogonal direction in Metal 1. Assuming a scaled design rule in 16nm technology, shown in Table 1, the upper and lower bounds on bit-cell area are  $0.10\mu m^2$  and  $0.07\mu m^2$  respectively with an aspect ratio (W/L) of the 2.42.

The aspect ratio of 8T-NWRAM bit-cell is greater than unity. The upper bound in design rule results in a bit-cell 510nm wide and 210nm tall, while the lower bound results in 425nm wide and 175nm tall. As a result, the internal routing

of signals in the horizontal direction experience larger capacitive load as compared to signals in vertical metal layer. This has significant impact in designing the array for 8T-NWRAM. In the proposed array architecture, bitline is routed in the vertical direction. This reduces bitline capacitance resulting in faster read operation. The architecture for 8T-NWRAM arrays is as shown in fig. 10.

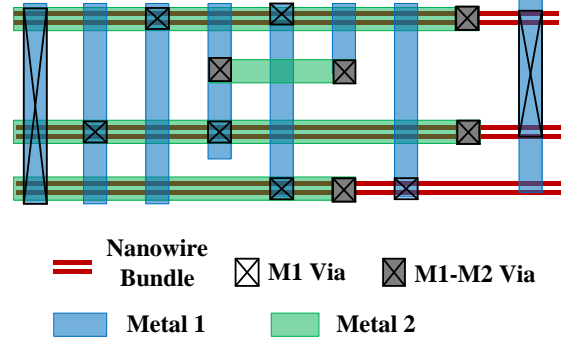


Figure 9: Physical Layout of 8T-NWRAM

Table 1: Interconnect and Via Design Rule

Pitch (16nm Tech)	Metal 1, Metal 2	Contact
1D-Gridded Design	40~60nm	50nm

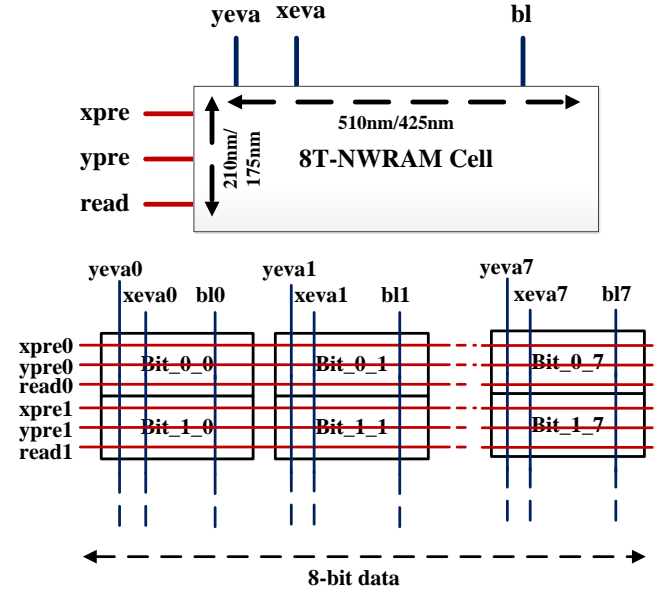


Figure 10: 8T-NWRAM Memory Array

## IV. IMPLEMENTATION AND RESULTS

The proposed 8T-NWRAM circuit was implemented using 2C-xnwFETs transistor models. These device models were generated from TCAD Sentaurus [10]. Interconnect parasitic models from Predictive Technology Models (PTM) [12] were used for RC delay calculation. Circuit simulations were performed in HSPICE. Since the area, performance and power of memory circuits are highly correlated with the physical implementation, we present a lower and upper bound for each of these parameters.

Table 2: 16nm Scaled Area for 6T-SRAM

Scaling Factor	2.45	2.02	1.75	1.64
Area in $\mu\text{m}^2$	0.028	0.042	0.056	0.064
	0.026	0.038	0.051	0.058
	0.025	0.037	0.049	0.056

Table 3: Design Rules for SRAM

Scaling Factor	0.76	0.7	0.72	0.76
M1x half pitch (nm)	32.49	27.56	29.16	32.49
	28.88	24.50	25.92	28.88
	28.88	24.50	25.92	28.88
N+/P+ spacing (nm)	43.32	36.75	38.88	43.32
	33.79	28.66	30.32	33.79
	37.54	31.85	33.369	37.54
Via spacing (nm)	32.49	27.56	29.16	32.49
	28.28	24.5	25.92	28.28
	28.28	24.5	25.92	28.28

The different bounds for SRAM are based on the benchmarking methodology proposed in [8][9]. Scaling factors and design parameters for 16nm SRAM were derived from scaling trends reported in literature [13-22]. For instance, [20] and [21] indicate a scaling of 2.02 in SRAM bitcell area from 45nm technology to 32nm. We assume the scaling factor to estimate SRAM bitcell area in 22nm and then in 16nm technology. Table 2 shows the different scaled area. The scaled design rules in Table 3. are used in conjunction with the scaled area to estimate parasitic values for power and performance measurements. In case of NWRAM circuits, there are no impediments in achieving the lower bound on interconnect dimensions. However, we present both upper and lower bounds based on 1D-Gridded Design rule for metal interconnects and contacts, table 1, for the sake of comparison with scaled SRAM designs. Based on the 1D-Gridded Design rules in Table 1, the upper and lower bound for 8T-NWRAM area are  $0.10\mu\text{m}^2$  and  $0.07\mu\text{m}^2$ . A comparison of 8T-NWRAM area with 6T SRAM, 8T SRAM and 10T-NWRAM is as shown in fig. 11. The increase in bit cell area is due to the additional Metal 2 routes and asymmetric aspect ratio.

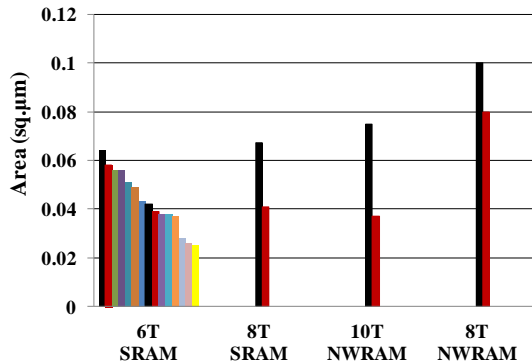


Figure 11: Comparison of Area

The write time of 8T-NWRAM is 4X smaller compared to high performance SRAM cell. Reduced Nanowire transistor in the sink path results in ~46% improvement in write time compared to 10T-NWRAM, fig. 12. One of the major improvements of 8T-NWRAM over conventional SRAM and 10T-NWRAM is access time. Unlike 10T-NWRAM, the bitline is routed on Metal layer 1 along the height of the bit-cell. Due to the larger than unity aspect ratio, the bitline length and hence the bitline load capacitance is significantly reduced. The 8T-NWRAM has 2X faster read access compared to high performance SRAM and 10T-NWRAM, fig. 13.

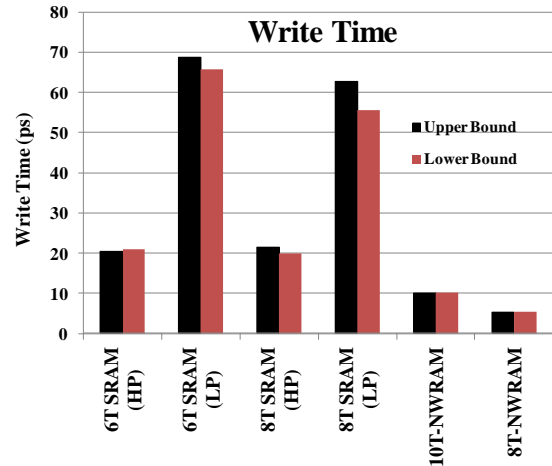


Figure 6: Comparison of Write Time

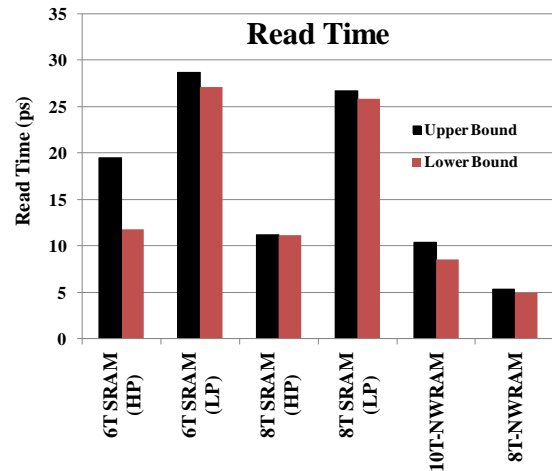


Figure 7: Comparison of Read Time

The active power during read operation reduces by 10% compared to 10T-NWRAM due to lower bitline capacitance, fig. 14. However, the active power is higher compared to conventional SRAM bit-cell. NWRAM circuits do not have a static power source during stand-by and hence lose charge over time. As a result, the net leakage power of NWRAM consists of the transistor leakage and the power for restore operation. The additional power for restore operation includes the power during two cycles of pre-charge and evaluation operations.

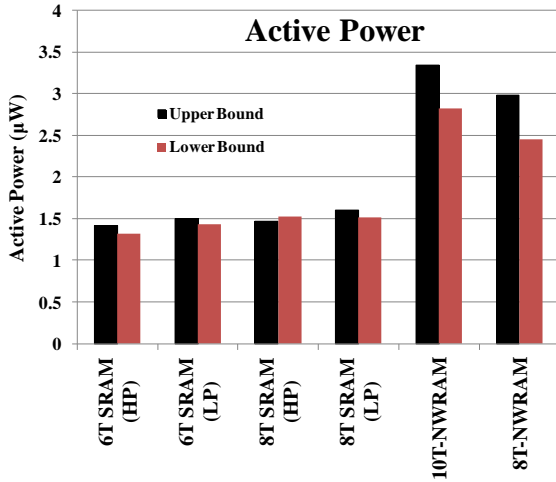


Figure 8: Comparison of Active Power

The 8T-NWRAM has  $\sim 20X$  smaller leakage power compared to high performance SRAM. However, it has a shorter sink path compared to the 10T-NWRAM circuit due to fewer stacked transistors. The internal nodes lose charge faster compared to 10T-NWRAM necessitating more frequent restore cycles. As a result, the leakage power of 8T-NWRAM is observed to be 33% more compared to 10T-NWRAM, fig. 15. Performance of SRAM may be expected to match 8T-NWRAM by trading-off for larger area. However, 8T-NWRAM provides significant leakage power improvement for the same performance targets.

The proposed 8T-NWRAM has significant improvement over 10T-NWRAM in access time due to reduced bitline capacitance. The improvement is further highlighted in NWRAM array architecture. A read time comparison of 10T and 8T-NWRAM for different array sizes (assuming equal number of 8-bit rows and columns) is shown in fig. 16. The 8T-NWRAM layout implementation provides  $\sim 2-3X$  improvement in read time across different memory bank sizes.

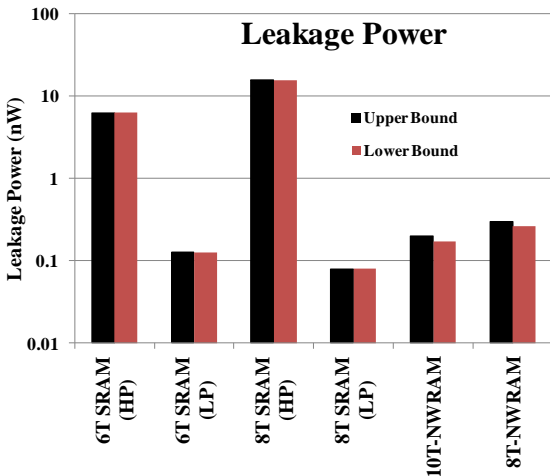


Figure 9: Comparison of Leakage Power

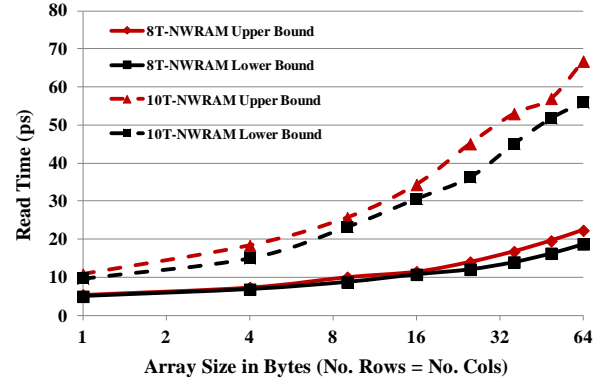


Figure 10: Comparison of Array Read Access Time

## V. STABILITY ANALYSIS

Stability is a critical evaluation parameter for storage elements. Unstable bit-cells lead to data error and require expensive error detection/correction techniques. In multi-level cache architectures, bit errors result in large performance penalty to fetch correct data from higher level memory entities. In this section we present a brief discussion and stability analysis of Nanowire based RAM circuits.

Conventional CMOS SRAM cells have stability issues due to local process variation [11]. Transistor threshold variation due to lithography effects and Random Dopant Fluctuation (RDF) cause device mis-match between pull-down, access and pull-up transistors. NWRAM circuits do not depend on matched device strengths for read, write and restore operations. However, since the memory access operations use free running on-overlapping clocks for pre-charge and evaluate, the potential stability issue in NWRAM arises from clock jitter. Clock jitter is perturbation in clock period or duty-cycle. Assuming the four clock signals  $xpre$ ,  $ypre$ ,  $xeva$  and  $yeva$  are derived from a common ideal PLL, the source of jitter is power supply noise. Power supply noise or droop occur due to RLC resonance in the power grid and local switching activity. Since read operation in NWRAM circuits are single ended and do not depend on clock signals, we study impact of jitter on bitcell stability during write and refresh operations.

### A. Effect of Clock Jitter during Write Operation

In a 10T-NWRAM, the write operation is performed by sequential pre-charge and evaluate of nodes  $out$  and  $nout$ . The value to be written is controlled by  $bl$ . Keeping  $bl$  low during  $xeva$  writes a value of "1" in  $out$ .

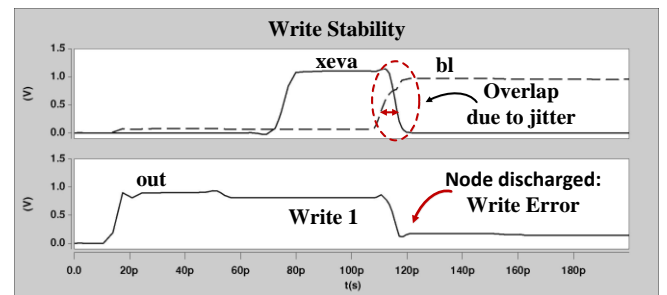


Figure 11: Write Error due to Clock Jitter

However, clock jitter can cause an overlap between the falling *xeva* and rising *bl* signals. This creates a sink path through transistors M2, M3 and M4, writing a bit “0” in node *out*. Fig. 17 demonstrates a write error in node *out* due to overlap between *bl* and *xeva*. Simulations indicate that overlap of 2ps or more between evaluate clock and *bl* causes write error. Unlike 10T-NWRAM, the 8T-NWRAM circuit is always write stable since *bl* is not used for controlling the write bit and the corresponding evaluate clock is gated to write a “1” into *out* or *nout*.

### B. Effect of Clock Jitter during Refresh Operation

A similar stability issue is observed due to overlap between an evaluate clock and the following complimentary pre-charge clock. Restore stability issue can occur in both 10T-NWRAM and 8T-NWRAM. Fig. 18, shows an example of internal bit swapping. The state of *out* is lost due to false evaluation because of overlap between *xeva* and *ypre*. An overlap of 2ps in case of 10T-NWRAM and 1ps in case of 8T-NWRAM causes bit flip. The 8T-NWRAM is more sensitive due to a faster pull down path.

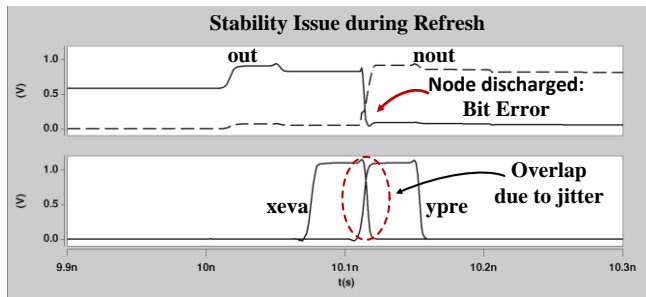


Figure 12: Bit Instability during Restore Operation

The stability issue due to clock jitter can be addressed by providing adequate design margin in the form of increased skew between clock edges. By separating the falling edge of evaluate clock from rising edge of corresponding pre-charge clock or *bl*, the potential overlap between signals can be avoided. The appropriate margin will depend on expected clock jitter and hence the overlap between the ideally non-overlapping clocks. Simulations in 16nm predictive CMOS models indicate that for a low frequency supply noise of 10% of VDD and frequency 400MHz [12], the maximum overlap between consecutive clock edges is ~26ps for 10-stage buffered clock network. This necessitates a design margin of 26ps between fall edge of evaluate clock and rise edge of the following pre-charge clock. Adequate design margin can ensure stability of bitcell in the presence of supply noise variation.

## VI. CONCLUSION AND FUTURE WORK

In this work we present a 8T Nanowire RAM circuit implemented in 2D Gridded N<sup>3</sup>ASIC nanofabrics. The circuit is based on existing 10T-NWRAM with dynamic 2C-xnwFET cross-coupled NAND structure. The N3ASIC has a reduced requirement on fabrication vs CMOS. An array architecture for 8T-NWRAM is proposed with minimized bitline load and high performance. Simulation results using 3D TCAD based device models indicate a bitcell read time

of ~4.9ps and write time of ~5.3ps. This accounts to ~2X improvements in performance compared conventional SRAM as well as to existing 10T-NWRAM. The bitcell area is 0.1 $\mu\text{m}^2$  and leakage power is 2.63nW – the latter a significant reduction vs all high performance SRAMs we could compare with. Stability analysis of NWRAM bitcells indicates sensitivity to supply noise induced clock jitter during write and restore operations. A design margin of >26ps between evaluate and pre-charge clock edges mitigates the impact of clock jitter. Future work will include a more detailed bitcell stability analysis.

## REFERENCE

- [1] Y. Taur, “CMOS design near the limit of scaling,” *IBM Journal of Research and Development*, vol. 46, no. 2.3, pp. 213–222, 2002.
- [2] S. Khan and S. Hamdioui, “Trends and challenges of SRAM reliability in the nano-scale era,” in DTIS, 2010.
- [3] S. V. Kumar, K. H. Kim, and S. S. Sapatnekar, “Impact of NBTI on SRAM read stability and design for reliability,” in *ISQED*, 2006
- [4] T. Wang, Z. Qi, and C. A. Moritz, “Opportunities and challenges in application-tuned circuits and architectures based on nanodevices,” in *First ACM Conference on Computing Frontier*, 2004, pp. 503–511.
- [5] A. DeHon, “Array-based architecture for FET-based, nanoscale electronics,” *IEEE Transactions on Nanotechnology*, 2003.
- [6] P. Avouris, et. al, “Carbon nanotube electronics,” *Proceedings of the IEEE*, vol. 91, no. 11, pp. 1772–1784, 2003.
- [7] A. DeHon, “Nanowire-based programmable architectures,” *ACM Journal on Emerging Technologies in Computing Systems*, 2005.
- [8] M. Rahman, P. Narayanan, and C. A. Moritz, “N3asic-based nanowire volatile RAM,” in *IEEE-NANO*, 2011.
- [9] M. Rahman and C. A. Moritz, “Nanowire Volatile RAM as an Alternative to SRAM,” *IEEE Transactions on Nanotechnology*.
- [10] P. Panchapakeshan, et. al, “N3ASICs: Designing nanofabrics with fine-grained CMOS integration,” in *NANOARCH*, 2011.
- [11] A. Agarwal, et. al, “Process variation in embedded memories: failure analysis and variation aware architecture,” *IEEE JSSC* 2005.
- [12] J. Jang, et. al, “Compact Expressions for Supply Noise Induced Period Jitter of Global Binary Clock Trees,” *IEEE TVLSI* 2012.
- [13] P. Bai, et. al, “A 65nm logic technology featuring 35nm gate lengths, enhanced channel strain, 8 Cu interconnect layers, low-k ILD and 0.57  $\mu\text{m}^2$  SRAM cell,” in *Electron Devices Meeting, 2004.*
- [14] X. Chen, et. al, “A cost effective 32nm high-K/ metal gate CMOS technology for low power applications with single-metal/gate-first process,” in *2008 Symposium on VLSI Technology*, 2008, pp. 88–89.
- [15] K.-L. Cheng, et. al, “A highly scaled, high performance 45 nm bulk logic CMOS technology with 0.242  $\mu\text{m}^2$  SRAM cell,” in *Electron Devices Meeting, 2007.*
- [16] C. H. Diaz, et. al, “32nm gate-first high-k/metal-gate technology for high performance low power applications,” in *Electron Devices Meeting, 2008. IEDM 2008. IEEE International, 2008*, pp. 1–4.
- [17] B. Greene, et. al, “High performance 32nm SOI CMOS with high-k/metal gate and 0.149  $\mu\text{m}^2$  SRAM and ultra low-k back end with eleven levels of copper,” in *Symposium on VLSI Technology*, 2009.
- [18] C.-H. Jan, et. al, “A 65nm ultra low power logic platform technology using uni-axial strained silicon transistors,” in *Electron Devices Meeting, 2005. IEDM Technical Digest. IEEE International, 2005.*
- [19] E. Leobandung, et. al, “High performance 65 nm SOI technology with dual stress liner and low capacitance SRAM cell,” in *Symposium on VLSI Technology*, 2005.
- [20] K. Mistry, et. al, “A 45nm Logic Technology with High-k+Metal Gate Transistors, Strained Silicon, 9 Cu Interconnect Layers, 193nm Dry Patterning, and 100% Pb-free Packaging,” in *Electron Devices Meeting, 2007.*
- [21] M. A. S. Natarajan, “A 32nm logic technology featuring 2nd-generation high-k metal-gate transistors, enhanced channel strain and 0.171 $\mu\text{m}^2$  SRAM cell size in a 291Mb array,” pp. 1 – 3, 2009.
- [22] A. Steegen, et. al, “65nm cmos technology for low power applications,” in *Electron Devices Meeting, 2005.*