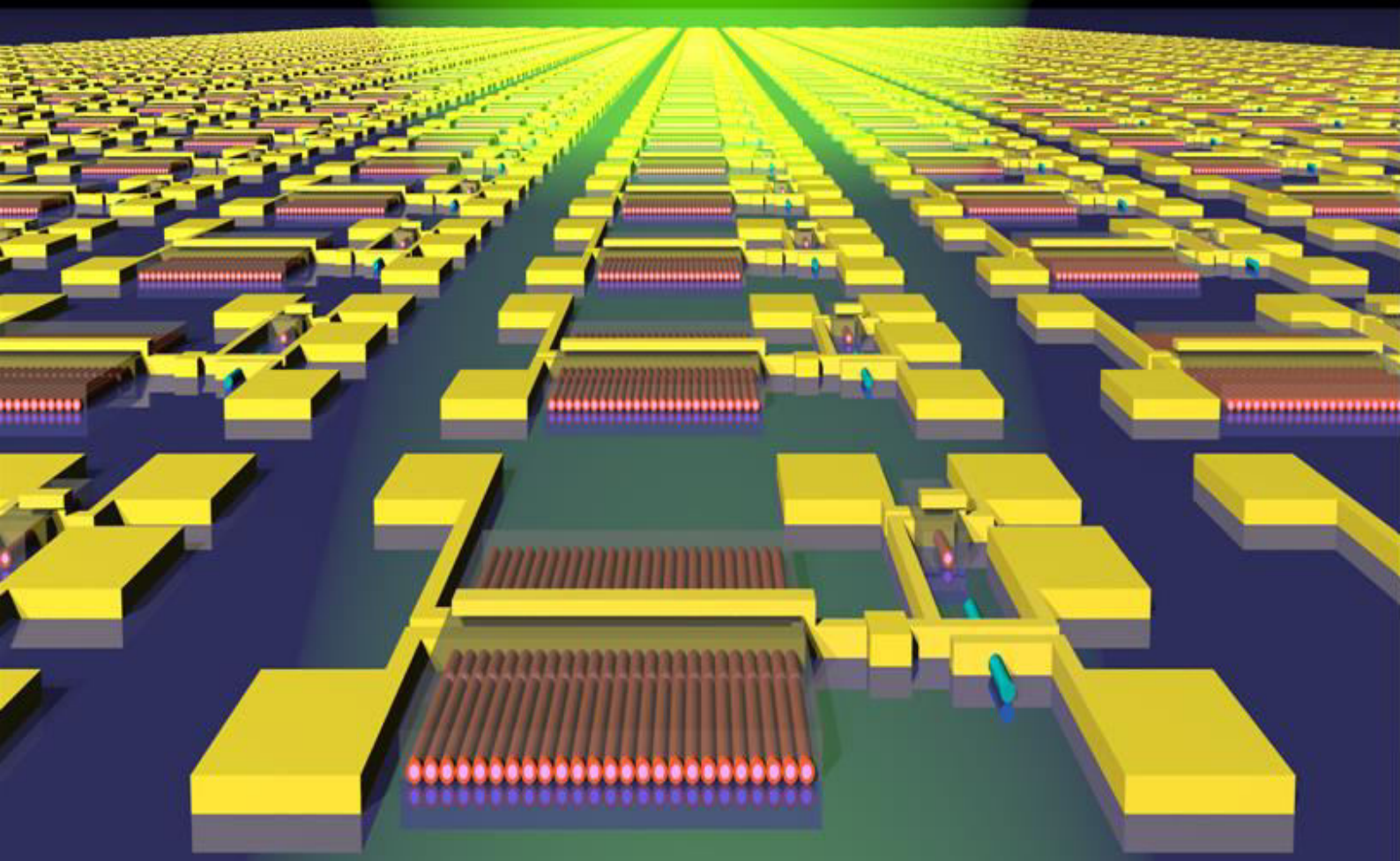
"EXPLORING THE RESEARCHES ON ELECTRON DEVICES"



TECHNICAL MAGAZINE ON ELECTRON DEVICES

ELECTRON DEVICES SOCIETY®

IEEE EDS KOLKATA CHAPTER
Vol. 1, No. 1, Jan-2016

ELECTRON
DEVICES
SOCIETY

Inaugural Editorial

We are honoured to welcome the readers and authors to the inaugural issue of the Technical Magazine, which is sponsored by IEEE EDS Kolkata Chapter.

The technical magazine is a new initiative of IEEE EDS Kolkata Chapter and is dedicated to under-graduate, post-graduate students and entry level research scholars who wish to pursue their academic and professional career in the domain of electron devices and their applications. The need for this type of technical magazine is realized during our attempts to reach the EDS student members in various seminars and workshops. The young students often come with several queries, which comprise of queries related to the topics of interest of EDS. Responding to the demands of these students and scholars, IEEE EDS Kolkata Chapter decided to publish a technical magazine with these students as the prospective readers. This technical magazine therefore attempts to present cutting edge technical ideas and important concepts relevant to the theory and applications, new research ideas and technical developments of electron devices in a lucid manner.

This first issue consists of five contributions related to topics of interest to IEEE EDS community. The first contribution is related to process variability which is a critical issue for low voltage VLSI design. The second contribution deals with Tunnel FET, which is a prospective device structure for ultra low power VLSI applications, through reduction of the sub-threshold swing below 60mV/decade barrier. The third contribution deals with resistive random access memory, which is an important non volatile memory technology. The fourth contribution is related to circuit design. It describes the VLSI design flow using electronic design automation tools. Finally, the fifth contribution describes a design flow of FPGA modelling and implementation. Therefore, it is seen that the contributions presented in this first issue encompasses a broad spectrum of electronic devices and circuits.

This technical magazine is a brain child of Prof. Soumya Pandit, of the University of Calcutta who is serving as the Editor-in-Chief. He along with a strong editorial board comprising of the following members, will work together related to decision about the relevance of individual submissions to an issue, review and approval of all reviewers' recommendations, and strategic or longer-term operational issues. The electronic version of the magazine will be initially published quarterly and will be available in the website of EDS Kolkata Chapter. We plan to publish the hard copy version soon. We look forward to your submissions, as well as your comments and support.

Dr. Soumya Pandit, *Editor-in-Chief*,
Assistant Professor,
Institute of Radio Physics and Electronics,
University of Calcutta, Kolkata, India
Email: soumya.pandit.rpe@gmail.com

Editorial Board

Advisory Board:

Message from President

Dear Fellow Calcutta EDS Chapter Members:

I am delighted to know the launching of the *Technical Magazine* by the IEEE Electron Devices Society (EDS) *Calcutta Chapter* aiming to publish technical articles related to electronic devices, technologies, and circuits. To the best of my knowledge, this is the first *chapter initiated* Technical Magazine outreach to *young researchers*, especially, undergraduate and graduate students with research interest in the EDS Field-of-Interest as described in the scope of this Magazine. Today's students are the engineering professionals of tomorrow. And, this Magazine is focused to benefit today's students in publishing their research ideas and help them prepare for top-tiered journal authors as well as entry into the industry or academia. The editorial board led by Professor Soumya Pandit is dedicated to help each student author to develop technical writing skills to be a successful author in the future publication endeavors. I encourage all young researchers to take advantage of this great opportunity to present their research work or ideas in this Magazine.

I wish a great success of the Calcutta EDS-chapter Magazine and set an example for the EDS chapters worldwide.

Prof. Samar Saha
President, Electron Devices Society, USA

Table of Contents

Process Variability: A Challenge to Low Voltage VLSI Design

Soumya Pandit, ***Member, IEEE, USA***

Assistant Professor, Institute of Radio Physics and Electronics,

University of Calcutta, Kolkata, India,

Email: soumya_pandit@ieee.org

I. Introduction

The growing market of portable electronic devices, like laptops, tablets, smart phone etc., demands electronic circuit design with ultra-low power dissipation. The various components of power dissipation in an integrated circuit are broadly classified into two categories: dynamic component and static component. The conventional approach of reducing the dynamic power dissipation of an integrated circuit is to reduce the supply voltage (V_{DD}). This is the reason for continuous downscaling of supply voltage from 5V to about 1V. The supply voltage was aggressively scaled from 5V in 0.5µm technology node to about 1.2V in 130-nm technology node. However, since then, the supply voltage nearly remains constant. The reason behind this is that the threshold voltage (V_T) of a MOS transistor does not scale down and the delay of a circuit increases (in other words, speed decreases) with the ratio V_T/V_{DD}. In CMOS technology, the material related parameters such as energy gap, work function etc., do not change with scaling. This restricts to a large extent, the scaling of threshold voltage. In addition, process variability, manifested as global/inter-die and local/intra-die process variability has limited the reduction of the threshold voltage of a MOS transistor. To realize further scaling of the supply voltage, device-level innovation is required.

II. Classification of Process Variations and Taxonomy

The two broad classifications of process variability are: inter-die/die-to-die and intra-die/within-die (see Fig. 1.). The parametric changes of identical MOS transistors separated by considerable distance, or fabricated at a different time are referred to as the inter-die process variations. Such variations can even be wafer-to-wafer or lot-to-lot. On the other hand, the parametric changes of identical MOS transistors located within a same die are referred to as intra-die process variations. The inter-die process variations affect all transistors on the same chip in a same way. This results in a shift in the mean value of the various design parameters such channel length, channel width, resistivity etc. On the other hand, intra-die variability may affect different devices differently on the same chip. The inter-die process variation effects are carefully considered by the VLSI designers by simulating the circuit not only at one design point but at small number of corners around the typical value of the nominal design point. These corners are often intelligently selected to faithfully represent the circuit behaviour under worst case deviations.

The various sources of process variations are used to create another taxonomy: systematic variations and random variations. The systematic variations include variations occurring from known physical phenomenon during manufacturing. The trend of variations across a chip is thus predictable. Examples are the variability caused by optical-proximity correction (OPC), phase-shift masking (PSM), layout-induced strain, and well-proximity effects. It is possible to address systematic variability through layout design and more controlled resolution-enhancement techniques (RETs). On the other hand, random variations include those variations that can be characterized in terms of a distribution. Such distributions can either be explicit in terms of samples collected from large number of measurements or implicit in terms of a probability density function that is fitted to a set of measurement data. Examples of random variability include random discrete doping, line edge roughness etc. The increasing amount of local/intra-die random process variability on the yield of nano-scale VLSI circuits, such as static random access memory (SRAM), has imposed an enormous

challenge in the conventional VLSI design methodologies. Addressing the impact of random local/intra-die variability requires innovative process and circuit design techniques and device modelling.

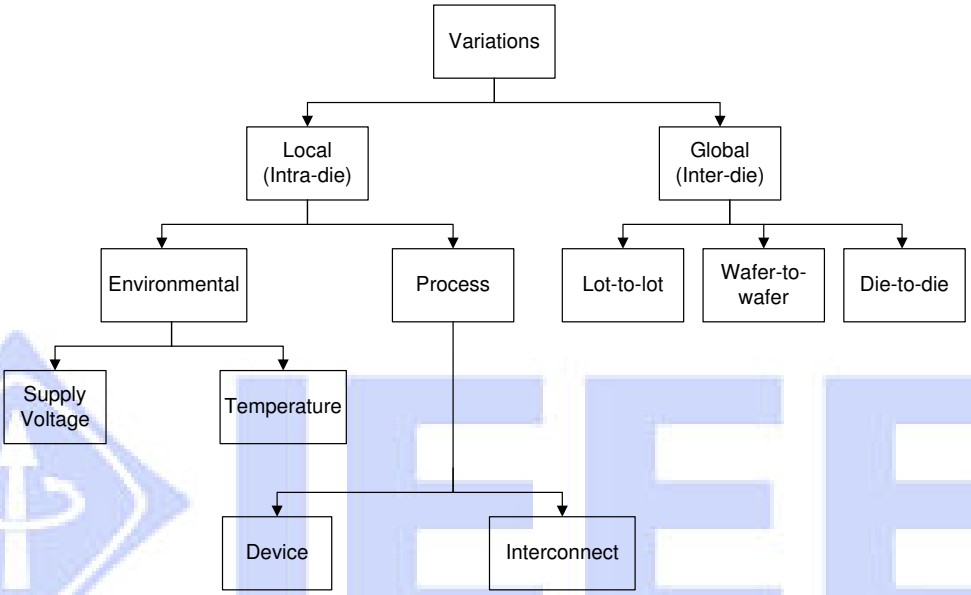


Figure 1: Classifications of Process Variations.

III. Random Local Process Variability

The three major sources of random local process variability are: random discrete doping, line edge roughness and oxide thickness variation. These are discussed below.

A. *Random Discrete Doping*

Random discrete doping (RDD) is a phenomenon which occurs due to discreteness of the dopant atoms present in the channel region of a MOS transistor. In a 1µm technology, a MOS transistor typically contains 5000 dopant atoms, whereas in a 45-nm technology, there are only about 100. Therefore, the number of dopant atoms is a discrete statistical quantity with probability to occupy any random location in the channel region (see Fig. 2.). This randomness of dopant atoms in the channel region of two identical MOS transistors placed side by side within a die results in variations of intra-die device and circuit performances. The major effect of RDD results in threshold voltage mismatch. The effect of RDD-induced process variability on threshold voltage mismatch is analytically modelled as

$$\sigma V_{T,\mathrm{RDD}} = \frac{q}{C_{ox}}\sqrt{\frac{N_A W_{dm}}{3LW}}$$

Here W_{dm} is the maximum depletion width of the MOS transistor; L and W are the channel length and width respectively. The threshold voltage mismatch leads to inversion charge mismatch which in turn leads to drain current mismatch between two otherwise identical MOS transistors sitting side by side on a die.

B. *Line Edge Roughness*

The critical physical dimensions of a semiconductor structure are defined through lithography process. The line edge roughness (LER) is caused by tolerances inherent to materials and tools used in the lithography process. It is another major source of random process variability and cause variations in critical dimension of the transistor feature size (see Fig. 3.). With the downscaling of critical

dimensions of MOS transistor, the amount of LER (typically 4-5nm) does not scale accordingly, becoming an increasingly larger fraction of the channel length of a transistor. The effect of LER includes degradation of the threshold voltage and higher sub-threshold leakage current.

C. *Oxide Thickness Variations*

The oxide thickness variation is caused due to the atomic scale roughness of the silicon-gate dielectric and gate-gate dielectric. This variation becomes very significant when the oxide thickness is equivalent to only a few silicon atomic layers. The oxide thickness variation causes variations in voltage drop across the oxide layer, which leads to threshold voltage variations.

IV. Effect of Process Variation on Leakage Power Consumption and Speed of a Set of Sample Devices

An important component of static power dissipation in MOS transistor is sub-threshold leakage current. The magnitude of this current depends strongly on the threshold voltage. Therefore, threshold voltage variability has direct impact on leakage power consumption. Figure 4 shows the distributions of the threshold voltage of two sets of sample devices. Each sample set comprises of 1000 identical devices. The distributions are characterized by the variance/standard deviation, i.e., σV_T. The set of devices with larger σV_T has wider distribution from lower to higher threshold voltages. It is observed that the distribution is symmetric Gaussian in nature. The distributions of leakage power consumption of the same two sets of identical devices are shown in Fig. 5. The leakage power consumption is measured by the product of the Off current and the supply voltage. We observe that the peak of the leakage power distribution for the set of devices with larger value of σV_T is shifted towards the left of the mean value. This implies that the leakage power consumption of a set of devices is dominated by the transistors with lower threshold voltages. The dependence of the total leakage power consumption of a set of devices with the distribution of threshold voltage is shown in Fig. 6. We observe that as the variance of threshold voltage increases, the leakage power consumption increases exponentially. This is because of the exponential dependence of the transistor leakage current with the threshold voltage.

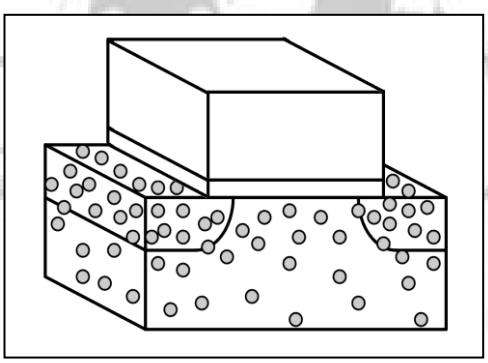


Figure 2: Illustration of discreteness of the dopant atoms in the channel region of a MOSFET.

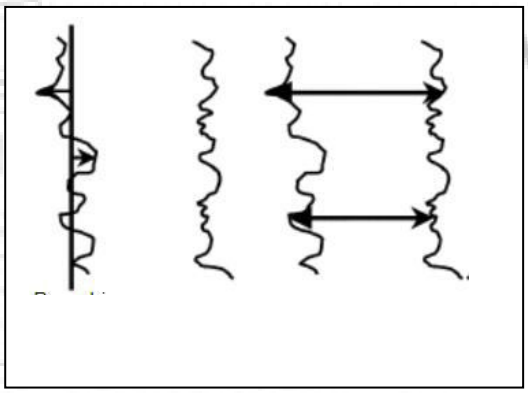


Figure 3: Illustration of roughness in channel length along the channel width direction.

For a constant supply voltage, with the increase of threshold voltage of a MOS transistor, the intrinsic speed of the transistor reduces and vice versa. The effect of process variation on the threshold voltage has significant impact on the speed performance of an overall IC. Higher device-to-device V_T limits the speed of the overall IC, which is due to the transistors with higher threshold voltage value. This is shown by right dotted circle in Fig. 7. In contrast, for same supply voltage, if such variations can be reduced, the path that limits the speed performance of IC becomes faster as shown by the left dotted circle in Fig. 7. Therefore, for same speed performance, (left circle overlaps with the right one) we can afford to reduce the supply voltage if it is possible to reduce the device-to-device threshold voltage variation. This is shown in Fig. 7.

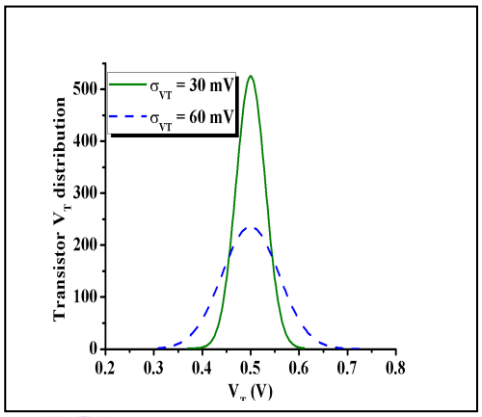


Figure 4: Distributions of two sets of sample devices

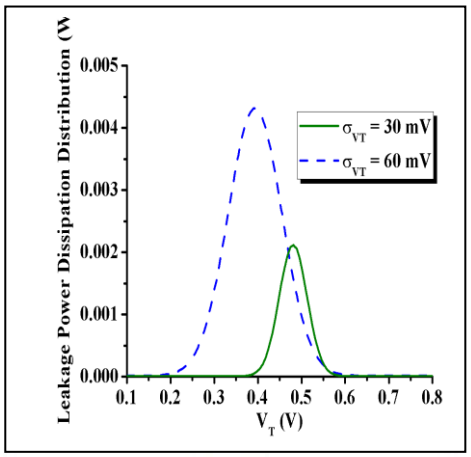


Figure 5: The distribution of leakage power of two sets of sample devices.

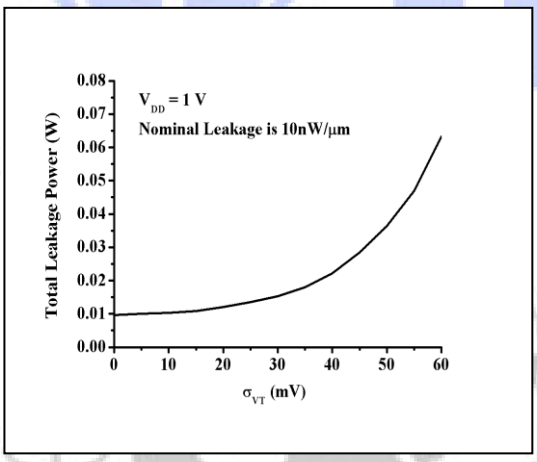


Figure 6: The dependence of total leakage power consumption on threshold voltage distribution.

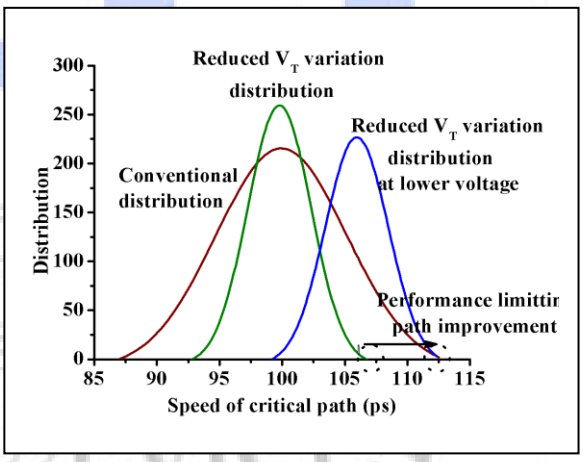


Figure 7: Benefit of lower threshold voltage distribution in achieving the same

V. Novel Device Structures for Mitigation of RDD Variability

To reduce device-to-device variations of threshold voltage and to achieve lower power dissipation the industry and researchers move from conventional doped channel transistor to undoped channel MOS transistor. However, with scaling, the doping concentration of the channel region must be made higher to reduce short channel effects. Therefore, undoped channel enhances on the other hand, the short channel effects. Therefore, novel structures must be designed to reduce RDD variability as well as control short channel effects. Three device structures are in development to achieve this goal: fully depleted silicon-on-insulator (FD-SOI), FinFET and epitaxial delta doped channel MOS transistor.

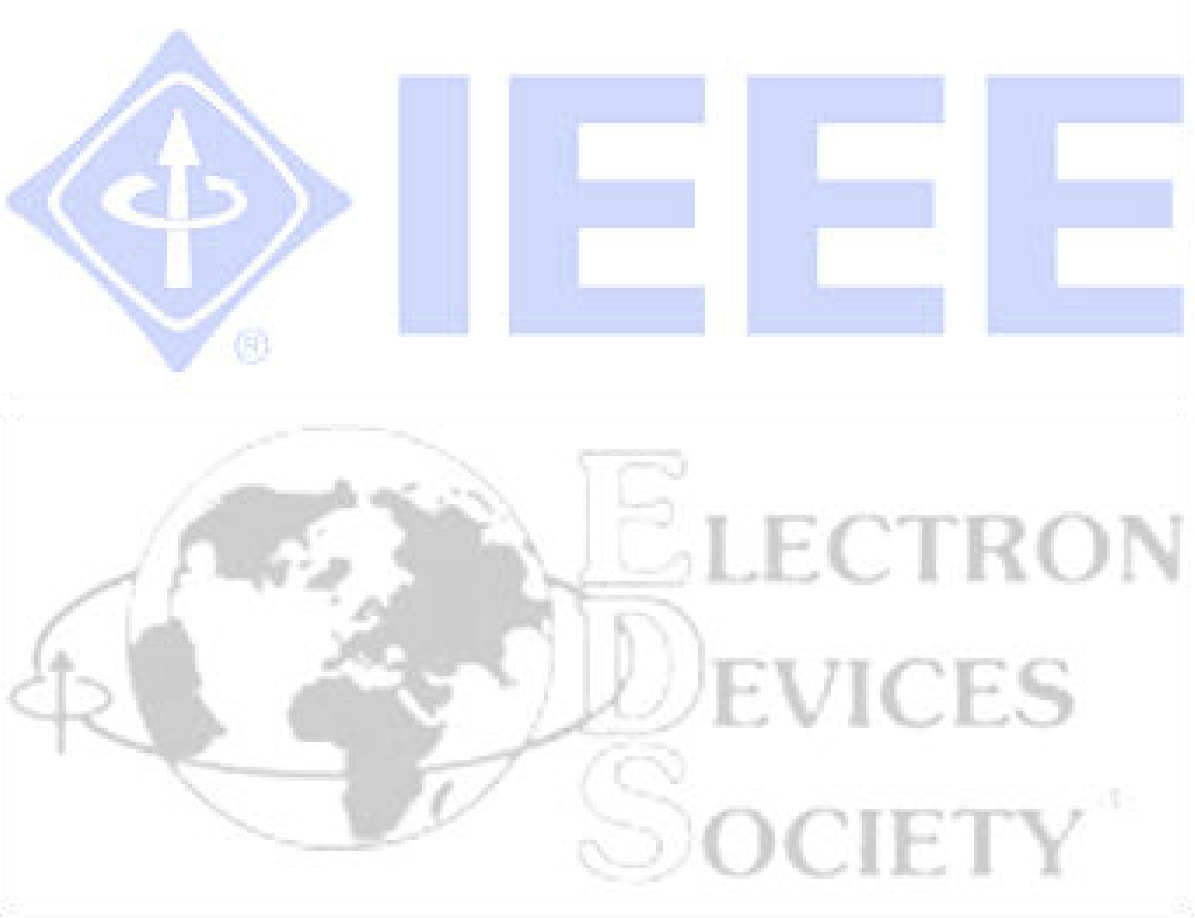
VI. Conclusion

Reduction of threshold voltage variability is the key for continued scaling of supply voltage. The major causes of random local process variability resulting in threshold voltage distribution are random discrete doping, line edge roughness and oxide thickness variations. Of these, the first one is the

dominant. The total leakage power consumption of a set of sample devices increases with increase of threshold voltage variance. It is possible to reduce the supply voltage of an integrated circuit without affecting the speed performance, if the threshold voltage variability can be reduced.

Further Reading

1. S.K.Saha, Modeling Process Variability in Scaled CMOS Technology, IEEE Design and Test of Computers, 2010 pp 8-16.
2. S.S.Sapatnekar, Overcomming Variations in Nanometer Scale Technology, Proc. Of IEEE JETCAS, Vol. 1, no 1. March 2011 pp 5-18
3. K.K.Kuhn and et al, Process Technology Variation, Proc. IEEE Trans. Electron Devices, Vol. 58, August 2011, pp 2197-2208.



Tunnel FET: A Perspective Review for Energy Efficient Applications

Dr. Angsuman Sarkar, ***Senior Member, IEEE***

Associate Professor, Dept. of ECE, Kalyani Government Engineering College, Kalyani, Nadia, India

angsumansarkar@hotmail.com

I. INTRODUCTION

The Complementary metal oxide semiconductor (CMOS) digital integrated circuits play a very crucial role in technology for modern information age [1]. Physical dimensions of metal-oxide semiconductor field-effect transistor (MOSFET) devices are being continuously scaled down over the past four decades according to Moore's law [2]. Aggressive downscaling in the deep sub-micrometer regime of the MOSFETs has led to enhancement of the detrimental short channel effects (SCEs) and leakage currents. [3]. In conventional MOSFETs, "Higher Subthreshold Conduction" can be identified as one of the fundamental SCEs that limits further downscaling due to its higher leakage current and power dissipation.

In the past, subthreshold leakage was small and ignored, but as transistors scales down a significant increase in subthreshold leakage has been observed. Subthreshold leakage current is the current that flows between the source and the drain of a MOSFET when the transistor is in sub threshold region, that is, for the gate to source voltages below the threshold voltage. The sub threshold region is often referred to as the weak inversion region. In digital circuits, sub threshold conduction is generally viewed as a parasitic leakage in a state that would have no current ideally. t is worth mentioning that subthreshold conduction is not the only one component of leakage, the other leakage components that can be roughly equal in size depending on the device design are gate-oxide leakage and junction leakage. It is known that with continuous downscaling, the supply voltages are also scaling down continuously in order to reduce the dynamic power consumption of integrated circuits, and to keep electric fields inside small devices low for maintaining device reliability. The reason for a growing importance of sub threshold conduction stems from this continual reduction of supply voltages. The amount of sub threshold conduction is set by the threshold voltage, and so has to be reduced along with the supply voltage. A reduction in threshold voltage means less gate voltage swing below threshold is required to turn the device "OFF" completely. As threshold voltage is reduced, the transistor cannot be switched from complete turn-off to complete turn-on with the limited voltage swing available. As a result, a compromise between strong current in the "ON" case and low current in the "OFF" case needs to be made in the circuit design, and the application determines whether to favor one over the other.

In the subthreshold region the drain current behavior is similar to the exponentially increasing current of a forward biased diode. Therefore, a plot of logarithmic subthreshold drain current versus gate voltage with drain, source, and bulk voltages fixed will appear approximately linear. Subthreshold swing S is the inverse of the subthreshold slope is defined as the change in voltage which must be applied in order to create a one decade increase in the output current and is given by

$$S = \left(\frac{dlog(I_d)}{dV_{GS}}\right)^{-1} \quad (1)$$

It can also be thought of as the voltage required for increasing or reducing subthreshold drain current by one decade. Where a dec (decade) corresponds to a 10 times increase of the drain current Id. A device characterized by steep subthreshold slope exhibits a faster transition between off (low current) and on (high current) states. While we know that MOS Scaling involves the scaling of the threshold voltage (V_{Th}), the above means that lowering V_{Th} by 60mV would increase the leakage current(power) by 10 times. In a conventional MOSFET, the subthreshold swing is given by

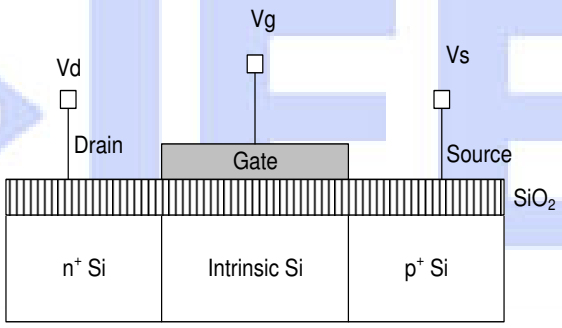
$$S = ln(10)\frac{kT}{q}\left(1 + \frac{C_d}{C_{OX}}\right) \approx ln(10)\frac{kT}{q} \equiv 60mV/dec|_{T=300K} \quad (2)$$

Where C_d= depletion layer capacitance and C_{OX}= gate-oxide capacitance of a MOSFET. In a ideal conventional long channel MOSFET, $C_{OX} >> C_d$, therefore subthreshold swing is limited to 60mv/dec at room temperature. A small subthreshold swing is advantageous because it supports a small voltage swing and to realize low power logic applications. It is here that we reach a roadblock in terms of MOS Scaling. MOSFETs are based on thermionic generation of electrons which restricts the subthreshold slope to a minimum of 60mV/decade [4]. Thus power supplies cannot be scaled down resulting in a low I_{ON}/I_{OFF} ratio. In other words, the problem with

60mv/dec subthreshold swing can be depicted as if one desires to shift V_{Th} by 60 mV, then the price to pay is an increase of one decade of off-current and in turn of static power. A better approach is to realize the steep SS devices by changing the current conduction mechanism while retaining the same materials or structure as that of a MOSFET. The lower SS results in improvement of the I_{ON} to I_{OFF} ratio, therefore reducing the static power dissipation at a given speed. There are plenty of interesting device structures being studied in multiple research groups that could reduce a switch's sub threshold swing to less than the 60 mV/decade limit of conventional MOSFETs at room temperature: the tunnel FET [4], the IMOS [5], MEMS/ NEMS switches [6]. Among these various possible alternative structures, tunneling field-effect transistors (TFET) emerges as a major candidate to replace conventional MOSFETs in order to continue downscaling, with a potential achievement of the realization of steep sub threshold swing.

II. STRUCTURE OF TUNNEL FET AND OPERATION

Tunnel FETs are gated p-i-n diodes, or less commonly, gated p-n diodes. To switch the device on, the diode is reverse biased, and a voltage is applied to the gate. Since a reverse bias is needed across the p-i-n structure in order to create tunneling, and since an NMOS operates when positive voltages are applied to the drain and gate, the n-region of a Tunnel FET is referred to as its drain, and the p+ region as its source for an n-type device as shown in Fig. 1. Unlike the MOSFETs, it is an asymmetric device.



Single Gate Tunnel FET

Fig 1: Structure of Single gate n-channel Tunnel FET

As the gate voltage is increased, the gate conduction band is goes down and this BTBT distance reduces. So BTBT current component dominates over the P-I-N leakage current as shown in Fig. 3. Now the electrons can directly tunnel through the barrier from the source VB to gate CB. If a positive drain bias is applied, these electrons will be drifted towards the drain and current will flow. Fig. 4 shows the effect of more gate bias and drain bias on the energy band profile in order to depict the increased tunneling. It is worth mentioning that In a TFET both the gate and source/drain voltages modulate the lateral electric field at the tunneling junction. The rate of change in the drain current with the gate voltage (subthreshold swing) depends on the rate of change of the tunneling distance with gate voltage and is not limited by kT/q thermal constraint as in conventional MOSFETs. Hence, it is possible to achieve a sub threshold swing of less than 60 mV/decade in TFET. This has also been verified experimentally [8].

TFET can be defined as "a semiconductor device in which the gate controls the source-drain current through modulation of Band-to-Band Tunneling (BTBT)". Band-to-Band Tunneling is a process in which electrons tunnel from the valence band through the semiconductor band gap to the conduction band or vice versa. When a Tunnel FET is OFF, only p-i-n diode leakage current flows between the source and drain, and this current can be extremely low (less than a fA/μm). For a symmetrical Tunnel FET (symmetry between the n- and p-sides with similar doping levels, similar gate alignment, etc.), ambipolar behavior has been observed in its drain current versus gate bias characteristics. In these ambipolar transfer characteristics, the negative gate bias resembles the characteriscs to a p-channel FET, whereby those of an nFET when a positive gate bias applied. The energy bands in the intrinsic region under the gate are lifted, and the energy barrier is now small enough for band-to-band tunneling to take place between the valence band of the intrinsic region and the conduction band of the n+-region. When a positive voltage is applied to the gate, on the other hand, the energy bands in the intrinsic region are pushed down and tunneling takes place between the valence band of the p+-region and the conduction band of the intrinsic region. The energy barrier width for band-to-band tunneling is the single most important factor that determines the amount of drain current through a Tunnel FET. The on-current of an n-type Tunnel FET depends on the width of the energy barrier between the intrinsic and p+ regions, and the current increases exponentially with a reduction in this barrier width.

At zero gate bias, the valence band (VB) of the source region and the conduction band (CB) of the intrinsic channel region are far apart as shown in Fig.3. So the tunneling distance and hence the probability of band-to-band tunneling (BTBT) is very high. This is the OFF-state. The current in OFF-state is mainly due to P-I-N leakage current. The OFF-current (IOFF) theoretically is in the range of fA/ μm. In real devices, IOFF is limited by the junction leakage and gate-tunneling current. Experimentally, IOFF in the range of hundreds of femto-amperes have been achieved [7].

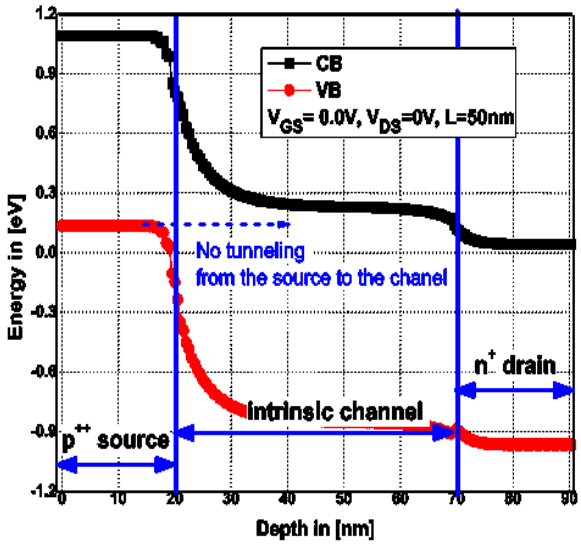


Fig. 2: Energy band profile of TFET at zero gate bias

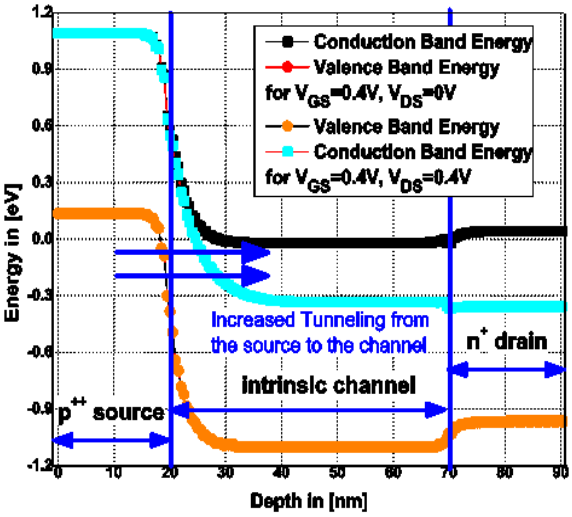


Fig. 3: Energy band profile of N-channel TFET for a small applied gate bias VGS=0.4V and for applied drain bias VDS=0V and 0.4V

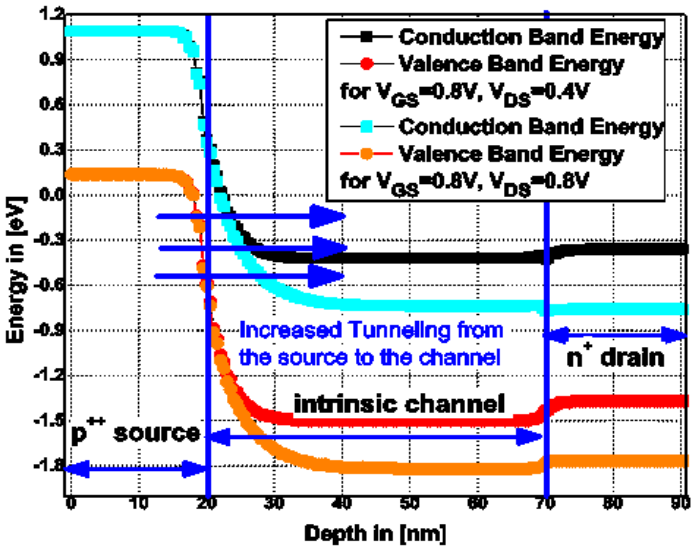


Fig. 4: Energy band profile for a N-channel TFET for applied gate bias VGS=0.8V and for applied drain bias VDS=0.4V and 0.8V

III. Recent Advancement

The last few decades have seen a burst of interest in the TFET [9-10] and a variety of approaches have been attempted. One of the basic problems of the Tunnel FET device is the low on current, which means the I_{on}/I_{off} ratio is very much less compared to the MOSFETs, so some actions needed to be taken in order to improve the on current of the device. To improve the ON-current and sub threshold swing, as in MOSFET, the double-gate structure [11-14] can also be applied to the TFET device. The performance of TFET can be improved if we use SiGe only at the source region [15-16] or both at the source and drain region. This gives rise to a new kind of TFET device which are called Hetero-junction Tunnel FET (HTFET) [17]. It has been demonstrated that lateral heterostructure TFET has the ability to be most effective to solve ambipolar issue with higher scalability [18].Recently, many researchers are using lower band gap material should be used in order to increase band to band tunneling rate. On the other hand, use of Halo doping [19] or dual material gate structure [20] has also been presented to increase the ON-current. In 2007, Bhuwalka et al. [21-22] published the first of many articles about their vertical Tunnel FET on silicon with a SiGe delta layer, grown by MBE. The SiGe can be replaced the silicon delta layer, and in theory, the smaller bandgap material can also be used to reduced the tunnel barrier width and increased tunneling current in the on-state as well as lowering the sub threshold swing. In 2007, Verhulst et al. at IMEC [23] showed by simulation that shortening Tunnel FET gate length, so that the gate covers the source-side junction where tunneling takes place, but does not cover the majority of the intrinsic region, has the benefits of decreasing off-current (tunneling through the drain-side junction and reducing speed, with a small or no reduction in the on-current, depending on the device design. Toh et al. [24] published a study of double-gate Tunnel FET silicon body thickness optimization for maximizing I_{ON}. Recently, several works that have attracted international interest and have been explored experimentally by multiple research groups in order to improve TFET performance are the use of bandgap engineering, hetero-structure material, strained semiconductors, novel architectures using carbon nanotube, graphene.

IV. Conclusion

The review of TFET demonstrates that TFET has distinct advantage of conventional MOSFET owing to its steep subthreshold swing resulting in much lower I_{OFF}. However, to increase I_{ON}, some novel architecture, materials has been the focus of research area in order to make TFET devices viable and attractive candidates for the future of digital logic designs, especially at ultra-low power applications.

References

[1] International Technology Roadmap for Semiconductors http://public.itrs.net,2009

[2] [2] "1965 – "Moore's Law" Predicts the Future of Integrated Circuits". Computer History Museum. http://www.computerhistory.org/semiconductor/timeline/1965-Moore.html

[3] R. Denard. "Design of Ion-Implanted MOSFETs with Very Small Dimensions", IEEE Journal of Solid State Circuits, 1974, p. 256

[4] Q. Zhang, W. Zhao, and A. Seabaugh, "Low-Subthreshold-Swing Tunnel Transistors," IEEE Electron Device Lett., vol. 27, no. 4, 2006, pp. 297-300

[5] W. Choi, J. Song, J. Lee, Y. Park, and B.-G. Park, "70-nm Impact-Ionization Metal-Oxide-Semiconductor (I-MOS) Devices Integrated with Tunneling Field-Effect Transistors (TFETs)," in IEDM Tech. Dig., 2005, pp. 955-958

[6] S. de Haan, "NEMS—emerging products and applications of nano-electromechanical systems". Nanotechnology Perceptions 2 (3): 267–275, 2006

[7] C. Le Royer and F. Mayer,"Exhaustive Experimental Study of Tunnel Field Effect Transistors (TFETS): From Materials to Architecture", in Proc. ULIS,2009,pp.53-56

[8] W. Y. Choi, B.-G. Park, J. D. Lee, and T.-J. K. Liu, "Tunneling fieldeffect transistors (TFETs) with subthreshold swing (SS) less than 60 mv/dec," Electron Device Letters, IEEE, vol. 28, pp. 743–745, Aug. 2007

[9] Koswatta, M. Lundstrom, and D. Nikonov, "Performance Comparison Between p-i-n Tunneling Transistors and Conventional MOSFETs," IEEE Trans. Electron Devices, vol. 56, no. 3, 2009, pp. 456-465

[10] T. Krishnamohan, C. Jungemann, D. Kim, E. Ungersboeck, S. Selberherr, A.-T. Pham, B. Meinerzhagen, P. Wong, Y. Nishi, K. C. Saraswat, "High performance, uniaxially-strained, silicon and germanium, double-gate p-MOSFETs," Microelectronic Engineering, vol. 84, 2007, pp. 2063–20661318-1323.

[11] P. Hashemi, L. Gomez, M. Canonico, J.L. Hoyt, "Electron Transport in Gate-All-Around Uniaxial Tensile Strained-Si Nanowire n-MOSFETs," IEDM Tech. Dig., 2008, pp. 865-868

[12] A.V.-Y. Thean, L. Prabhu, V. Vartanian, M. Ramon, B.-Y. Nguyen, T. White, H. Collard, Q.-H. Xie, S. Murphy, J. Cheek, S. Venkatesan, J. Mogab, C.H. Chang, Y.H. Chiu, H.C. Tuan, Y.C. See, M.S. Liang, Y.C. Sun, "Uniaxial-Biaxial Stress Hybridization For Super-Critical Strained-Si Directly On Insulator (SC-SSOI) PMOS With Different Channel Orientations," IEDM Tech. Dig., 2005, pp. 509-512

[13] K. Moselund, P. Dobrosz, S. Olsen, V. Pott, L. De Michielis, D. Tsamados, D. Bouvet, A. O'Neill, A.M. Ionescu, "Bended Gate-All-Around Nanowire MOSFET: a device with enhanced carrier mobility due to oxidation-induced tensile stress," IEDM Tech. Dig., 2007, pp. 191-194.

[14] N. Singh, W.W. Fang, S.C. Rustage, K.D. Budharaju, S.H.G. Teo, S. Mohanraj, G.Q. Lo, N. Balasubramanian, D.-L. Kwong, "Observation of Metal-Layer Stress on Si Nanowires in Gate-All- Around High-k/Metal-Gate Device Structures," IEEE Elec. Dev. Lett., vol. 28, no. 7, July 2007, pp. 558-561.

[15] S. Pidin, T. Mori, R. Nakamura, T. Saiki, R. Tanabe, S. Satoh, M. Kase, K. Hashimoto, T. Sugii, "MOSFET current drive optimization using silicon nitride capping layer for 65-nm technology node," Symp. on VLSI Technology, 2004, pp. 54-55.
[16] V. Nikam, K. K. Bhuwalka, and A. Kottantharayil, "Optimization of N- and P-channel heterojunction Tunnel FETs for sub 22nm gate lengths," in Device Research Conference, June 2008, pp. 77–78
[17] V. Nagavarapu, R. Jhaveri, and J. C. S. Woo, "The Tunnel source (PNPN) n-MOSFET: A novel high performance transistor," IEEE Trans. Electron Devices, vol. 55, no. 4, pp. 1013–1019, Apr. 2008
[18] Ashish Pal, Angada. B. Sachid, Harald Gossner and V. Ramgopal Rao " Insights into the Device and Optimization of Tunnel FET Devices and Circuits" IEEE Transactions on Electron Devices Vol. 58, no. 4, April 2011
[19] Sneh Saurabh , M. Jagadesh Kumar , "Novel Attributes of a Dual Material Gate Nanoscale Tunnel Field-Effect Transistor," IEEE TRANSACTIONS ON ELECTRON DEVICES, VOL. 58, NO. 2, FEBRUARY 2011, pp 404-410
[20] K. Bhuwalka, S. Sedimaier, A. Ludsteck, C. Tolksdorf, J. Schulze, and I. Eisele, "Vertical Tunnel Field-Effect Transistor," IEEE Trans. Electron Devices, vol. 51, no. 2, pp. 279-82, Feb. 2004
[21] K. Bhuwalka, J. Schulze, and I. Eisele, "Scaling the Vertical Tunnel FET with Tunnel Bandgap Modulation and Gate Workfunction Engineering," IEEE Trans. Electron Devices, vol. 52, no. 5, pp. 909-917, May 2005
[22] K. Bhuwalka, M. Born, M. Schindler, M. Schmidt, T. Sulima, and I. Eisel Verhulst e, "P-Channel Tunnel
[23] Field-Effect Transistors down to Sub-50 nm Channel Lengths," Jap. J. Appl. Phys., vol. 45, no. 4B, 2007, pp. 3106-3109
[24] E.-H. Toh, G. Wang, G. Samudra, and Y.-C. Yeo, "Device physics and design of double-gate tunneling field-effect transistor by silicon film thickness optimization," Appl. Phys. Lett., vol. 90, 2007, pp. 263507-1-3.

Resistive Random Access Memory: An Emerging Technology

Dr. Writam Banerjee, ***Member, IEEE***

Key Laboratory of Microelectronics, Devices & Integrated Technology and Lab of Nanofabrication and Novel Device Integration
Inst. of Microelectronics, Chinese Academy of Sciences, Beijing 100029 (China)

I. INTRODUCTION

Memory is an essential element of the electronic devices. Based on working procedure computer memory mainly devided into volatile and non-volatile form. Flash is a very common example of non-volatile memory (NVM). Due to fundamental scaling limitations and process complexity of baseline NVM devices, distinct and individual alternative approaches with exciting architectures are being explored. Recently, the emerging NVM devices are competing with baseline technologies. Among several alternatives, resistive random access memory (RRAM) is a potential candidate because of its simple design, high speed operation, magnificent scalability and good control over economic budget. Simple metal-insulator-metal based RRAM can work based on the changes of its resistance state. Cation or anion-like defects are controlling the resistive switching phenomena within a RRAM device. Along with the good performances, the RRAM devices offers the new applications in the form of high density 3D crossbar memory, transparent and flexible electronics and also neuromorphic devices. This article will lead the basic understanding of RRAM technology and will provide an overview of RRAM design, switching mechanism, electrical properties and applications.
Memory is the term which remind us about a special part of living body with receiving, storaging and recalling abilities. Which allow us to live in the past. Similar to our body, the memory is a vital part of computing systems. Memory in computer reffers to the computer hardware with the abilities to store information, which can be used if necessary. The computer memory usually refer as semiconductor memory and can be found in the form of semiconductor based integrated circuits (IC) often based on the silicon transistors. Due to its low cost and abundance in earth crust, silicon is used commonly and is a most fundamental element of modern electronic applications. Depending on the capabilities, the semiconductor memory can be catagorised into two parts: volatile memory (VM) and non-volatile memory (NVM).

A. Volatile Memory

In general, the form of memory which requires power to memorize the stored information, is known as VM. VM usually refers as the main memory of the storage system. Dynamic random access memory (DRAM), static random access memory (SRAM) are the types of VM. There are few VM also available in research devition such as thyristor random access memory (TRAM) and zero-capacitor random access memory (ZRAM).

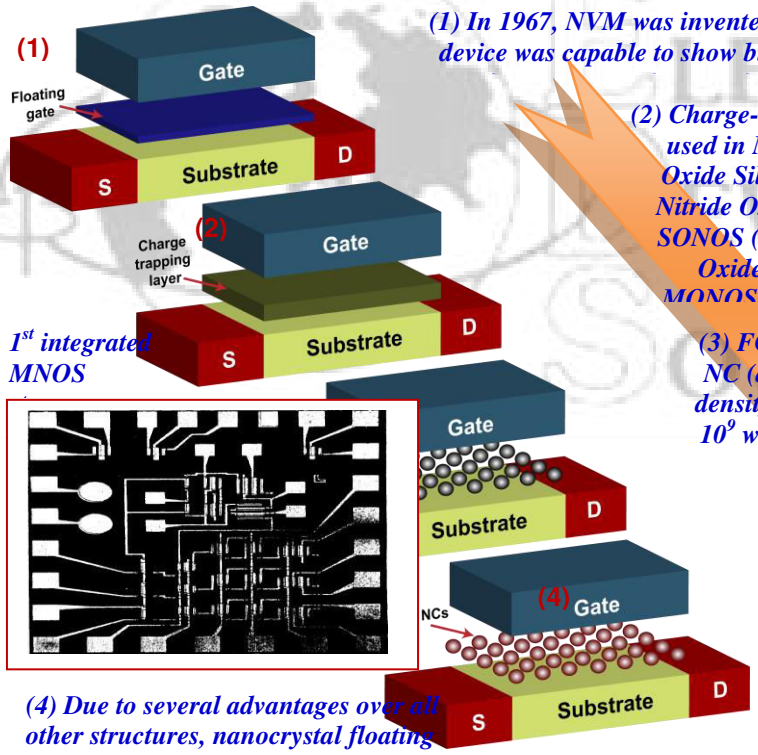
B. Non-Volatile Memory

Unlike VM, the NVM can memorize the stored information even after switch off the power. Flash memory is one of the most dramatic and long-term evolution in NVM field. Beside that, some prototype NVM devices are ferroelectric random access memory (FeRAM), spin-transfer torque magnetic random-access memory (*STT-RAM* or *STT-MRAM*), phase change memory (PCM). Research is going on to find the new emerging NVM such as electrochemical metalization system (ECM), valance change memory (VCM), three dimentional crosspoint memory etc. Generally, the switching mechanism of PCM, ECM and VCM are based on resistance and is known as RRAM. The standpoints of current baseline and prototype memories are summarized in Table 1 [1]. All memories have their own advantages and disadvantages. Depending on the application one can choose the type of the memory. Sometime it is easy to understand the concept from philosophical point of view rather than practical one.

It is very easy to realize the VM and NVM phenomena in our regular life. The similar way of remembering can also be found in human memory system. For example, we come across many persons in life. But we remember only a few of them. We always remember the person who is closer to our heart and/or mind without seeing them. This is synonymous with NVM. On the other hand, when we remember people upon meeting or seeing them, it is VM.

Table: 1

		Baseline technologies					Prototype technologies		
		DRAM		SRAM	Flash		FeRAM	STT-MRAM	PCM
		Stand-alone	Embedded		NOR	NAND			
Memory Type		Volatile Memory			Non-volatile Memory				
Cell Elements		1T1C		6T	1T		1T1C	1(2)T1R	1T(D)1R
Feature size F, nm	2013	36	65	45	45	16	180	65	45
	2026	9	20	10	25	> 10	65	16	8
Cell Area	2013	6 F^2	(12-30) F^2	140 F^2	10 F^2	4 F^2	22 F^2	20 F^2	4 F^2
	2026	4 F^2	(12-50) F^2	140 F^2	10 F^2	4 F^2	12 F^2	8 F^2	4 F^2
Read Time	2013	< 10 ns	2 ns	0.2 ns	15 ns	0.1 ms	40 ns	35 ns	12 ns
	2026	< 10 ns	1 ns	70 ps	8 ns	0.1 ms	< 20 ns	< 10 ns	< 10 ns
W/E Time	2013	< 10 ns	2 ns	0.2 ns	1µs/10ms	1/0.1ms	65 ns	35 ns	100 ns
	2026	< 10 ns	1 ns	70 ps	1µs/10ms	1/0.1ms	<10 ns	<1 ns	<50 ns
Retention Time	2013	64 ms	4 ms	-	10 y	10 y	10 y	>10 y	>10 y
	2026	64 ms	1 ms	-	10 y	10 y	10 y	>10 y	>10 y
Write Cycles	2013	>1E16	>1E16	>1E16	1E5	1E5	1E14	>1E12	1E9
	2026	>1E16	>1E16	>1E16	1E5	1E5	>1E15	>1E15	1E9



(1) In 1967, NVM was invented with a FG. The FG device was capable to show bi-stable performance,

(2) Charge-trapping are typically used in MNOS (Metal Nitride Oxide Silicon), SNOS (Silicon Nitride Oxide Semiconductor), SONOS (Silicon Oxide Nitride Oxide Semiconductor) and MONOS (Metal Oxide Nitride

1st integrated MNOS

(3) FG devices based on Si NC (diameter of 5 nm and density of $1\times10^{12}/cm^2$) with 10^9 write/erase cycles were investigated by

(4) Due to several advantages over all other structures, nanocrystal floating gate memory (NFGM devices) is in the

It is very easy to realize the VM and NVM phenomena in our regular life. The similar way of remembering can also be found in human memory system. For example, we come across many persons in life. But we remember only a few of them. We always remember the person who is closer to our heart and/or mind without seeing them. This is synonymous with NVM. On the other hand, when we remember people upon meeting or seeing them, it is VM.

The major focus of this article is RRAM technology, which is an emerging technology. But, before going to the emerging NVM it is important to discuss the baseline NVM i.e. flash memory.

II. NON-VOLATILE FLASH MEMORY

Flash memory or conventionally the non-volatile floating gate memory (FGM) was invented by D. Kahng and S. M. Sze in late 1960s at Bell Labs [2].

Not long ago, flash technology became a strong cost effective semiconductor storage medium widely used in electronic applications. Flash memory is usually two types, NAND and NOR. Generally, NAND flash is designed with a smaller cell size which can be used as a high-density storage medium where as NOR flash can be used as the code storage medium. Flash memory has offered mobility, variability, and a vision of scaling to the new age electronic devices. To meet the demand of the semiconductor industry it is essential to improve the performances, reliability with the devices scaling.

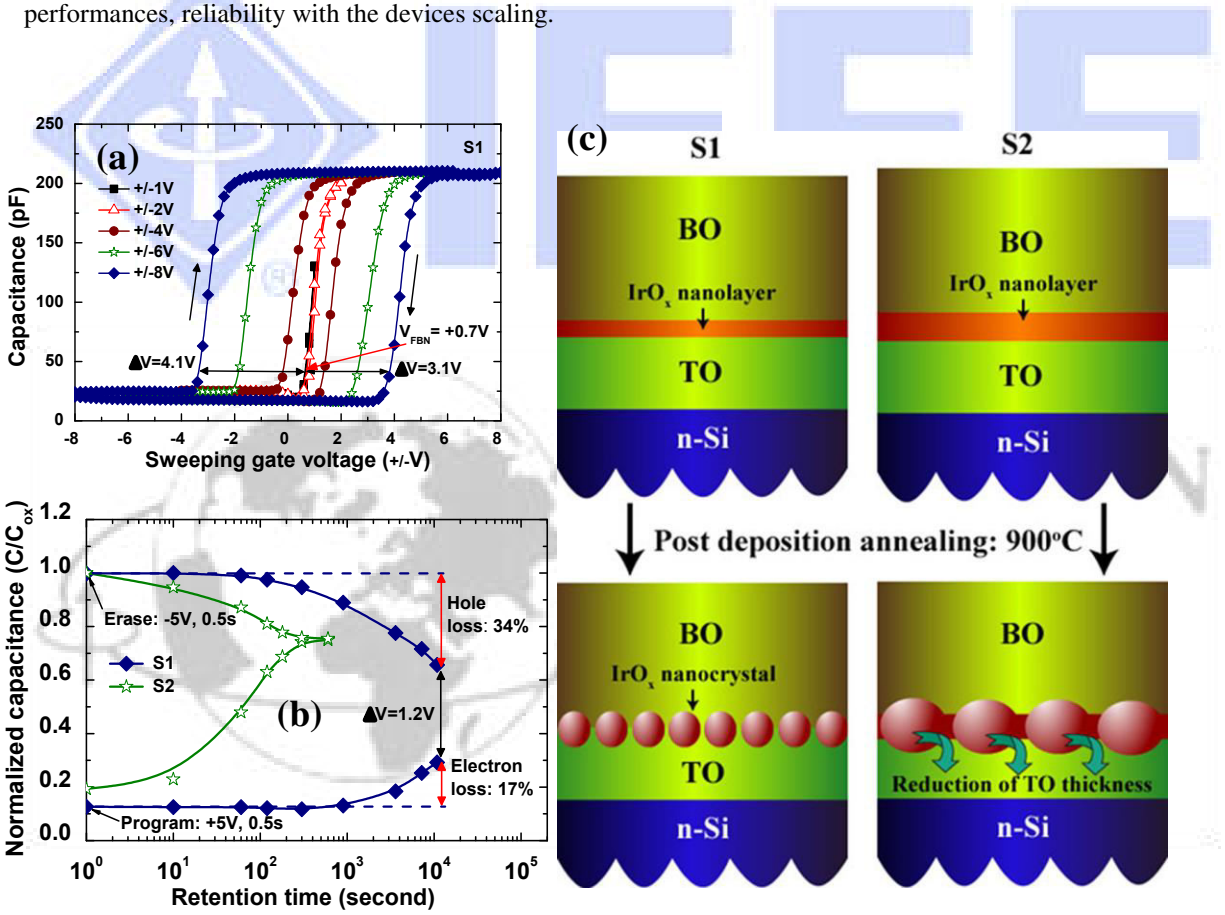


Figure 2: (a) Capacitance-voltage characteristic of IrO_x-NC based flash device. (b) Retention characteristics. (c) Schematic illustration of the retention failure model. The thicker the tunnelling oxide the better the device retention. [4]

Table 2

Parameter	Prototypical			Emerging RRAM			
	FeRAM	STT-MRAM	PCM	Conducting bridge	Metal Oxide		
					Bipolar	Unipolar	Interface
Scalability	⇩	⇔	⇧	⇧	⇧	⇧	⇧
MLC	⇩	⇩	⇧	⇧	⇔	⇔	⇔
3D integration	⇩	⇔	⇧	⇔	⇧	⇧	⇧
Fabrication Cost	⇔	⇔	⇔	⇔	⇧	⇔	⇔
Retention	⇔	⇔	⇔	⇔	⇧	⇔	⇩
Latency	⇧	⇧	⇔	⇔	⇔	⇔	⇔
Power	⇧	⇔	⇩	⇧	⇔	⇩	⇔
Endurance	⇧	⇧	⇔	⇔	⇔	⇩	⇔
Variability	⇧	⇩	⇔	⇔	⇩	⇩	⇧

Parameter	⇩	⇔	⇧
Scalability	F_{min} > 45 nm	F_{min} = 10 – 45 nm	F_{min} < 10 nm
MLC	difficult	possible	feasible
3D integration	difficult	possible	feasible
Fabrication Cost	high	medium	low
Retention	poor (<< 1 yr)	modest (> 1 yr)	long (> 10 yrs)
Latency	long (> 10 µs)	medium (0.3-10 µs)	short (< 300 ns)
Power	high	medium	high
Endurance	< 1E5 cycles	< 1E10 cycles	> 1E10 cycles
Variability	problematic	reasonable	low

A. Development of FGM:

Fig. 1 shows the development of FGM devices [3]. The basic structure was improved with the engineering the storage layer rafered as charge trapping flash and nanocrystal (NC) flash (NFGM) useing semiconductor or metal NC. Recently NFGM are getting huge attention because of several reasons such as, i) the formation of discrete NC can solve the leakage issue of a continuous storage layer, ii) higher available work function of the NC material, iii) good controllability of size and density of NC, iv) two bit per cell storage at the NC can improve the memory density. Several companies have been officially demonstrated NFGM devices. The NCs are used as the charge storage node where information can be stored in the form of binary digits "1" or "0". During programming, the charge can be injected to the NC and during erasing the charge can be removed from the NC or vice-versa. In NFGM devices, NC material, size, density, structure design and fabrication processes will greatly affect the device performances. It is found that the higher the annealing temperature, bigger the size of NC with lower density [4]. Several problems limits the further development of the flash memory with simple structure such as i) a rapid decrement of total number of storage node in single cell due to device scaling, ii) the reduced thickness of the tunnelling oxide layer will caused poor data retention problem [Fig. 2], iii) cell-to-cell interference problem will be severe with decreasing cell size etc.. Therefore, new emerging NVM devices are in the centre of attraction of the research. Several reasons put forward RRAM as compare to the other emerging devices as shown in Table 2 [1].

III. Emerging RRAM

RRAM or ReRAM is a type of emerging NVM technology in which the applied electrical stimulation can change repeatedly the basic resistance property of an insulating layer, from a high resistance state (HRS) to a significantly low resistance state (LRS) or vice-versa [5].

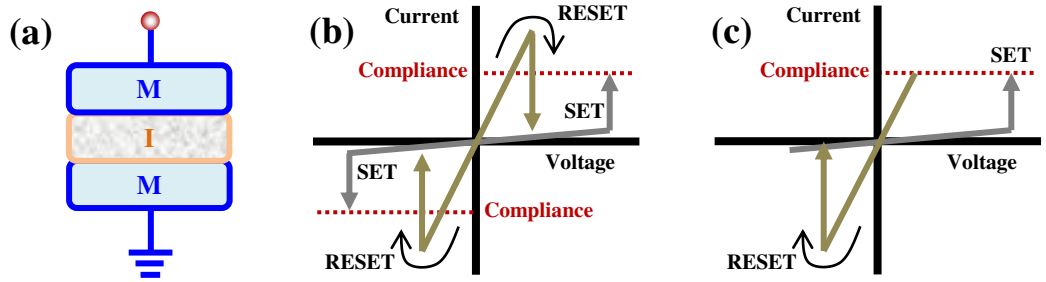


Fig. 3. (a) Basic structure of RRAM. Resistive switching devices usually show two types of current-voltage characteristics **(b)** uni-polar and **(c)** bi-polar.

The basic structure is simply in the form of metal-insulator-metal (MIM) as shown in Fig. 3. Several materials are available as insulating layer such as binary oxide/multinary oxide, transition metal oxide materials, higher chalcogenides and organic compounds.

A. Basic Mechanism of RRAM

The process to switch RRAM device from HRS to LRS i.e ON state is known as SET process and the in reverse the switch from LRS to HRS i.e OFF state is known as RESET process.

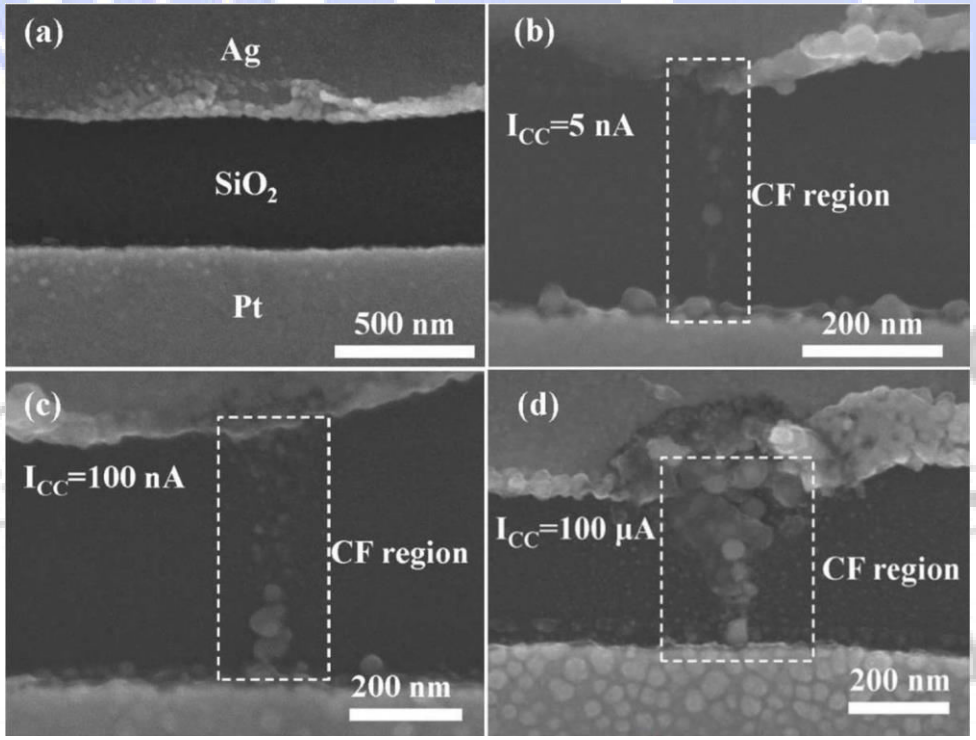


Fig. 4. (a) Ag/SiO_2/Pt planar RRAM structure. Electroforming with various compliances and morphologies of the conductive filament with different current compliances of **(b)–(d)** 5 nA, 100 nA and 100 μA, respectively.

Generally, the high initial resistance of a RRAM cell can be changed by using a voltage higher than the voltage necessary to SET the device. The process is known as electroforming process. The resistive switching mechanism of RRAM depends on several factors e.g. device structure, material, cell size etc. and it can be classified into three major categories i.e. ECM, VCM, and thermochemical systems [6]. An ECM cell usually known as conductive bridge random access memory (CBRAM). Structurally it consists with an electrochemically active metal (e.g. Ag, Cu, Ni) layer, an inert metal (e.g. Pt, Ir, Au) layer and solid electrolyte or oxide materials. Depending on current compliance the switching process will be in progress with dissolution of active metal to metal cations (i.e., Cu^{2+}, Ag^{+} or Ni^{2+}) → migration of cations towards the inert metal → reduction and deposition on the inert electrode [7]. Between two electrodes the conducing filament usually consists with isolated nanoparticles from the active electrode. Based on current compliance the RRAM device can show threshold switching (low compliance) and memory switching (high compliance) behavior [Fig. 4] [8].

VCM is a type of RRAM where none of the electrode materials inject metal cations. The oxygen vacancy like anion transportation is an essential requirement in this kind of devices. Therefore, it is very easily understood that the defects are the key factor for resistive switching. The nature of the filament is either metallic or semiconductor which can be easily figure out by a simple temperature dependent study of resistance. For metallic filament, the increase of temperature will increase the resistance but for semiconducting filament the resistance will be decreases with temperature. Generally, in metals the numbers of mobile carriers are high enough for conduction in normal situation. With increasing temperature the vibrations of lattice atoms will constantly interfere with the transportation path of those mobile carriers. Hence with increasing temperature, the resistance of metallic filament will be increased and the conductivity will be decreased. In case of semiconductor, insufficient mobile carriers are available at normal situation but the number of carriers will be increased with increasing the temperature as with temperature lightly bounded carrier will be free. Therefore, with increasing temperature the resistance of semiconducting filament will be decreased and the conductivity will be increased unless the maximum number mobile carriers become free to conduct. When the resistive switching based on thermal effect with uni-polar characteristics is known as thermochemical Systems. A common example of such type is PCM. Heating by electrical current dependent transformation from amorphous to crystalline or vice-versa is the basic switching mechanism of this kind of devices.

B. Electrical Characteristics:

During the measurements of RRAM devices the electrical stimulation is applied to one electrode and the other electrode remains grounded as shown in Fig. 3.

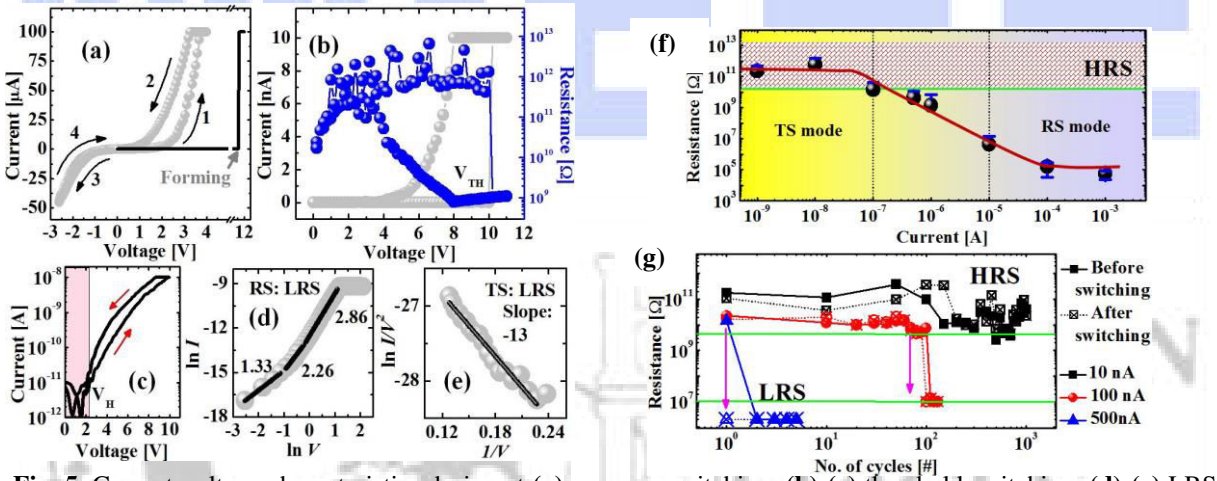


Fig. 5. Current-voltage characteristics during at **(a)** memory switching, **(b)-(c)** threshold switching. **(d)-(e)** LRS fitting. **(f)-(g)** Transformation from threshold to memory switch.

Usually, an initial forming process is necessary to break the virgin state of the device. Typical bi-polar current-voltage characteristic is shown in Fig. 5. The switching process can progress with the changing of applied voltage from 0 → +V → 0 → -V → 0 and the switching direction is indicated with arrows from 1 to 4. The LRS of the memory switching is very stable as compare to the threshold switching. Depending on current compliance level, a unidirectional change from threshold switching to memory switching is possible [9]. The resistance ratio (R_{OFF}/R_{ON}) of > 10 is necessary to allow sense amplifiers. Data endurance and retention are the vital properties for the adoption of new technology. To become a promising candidate in memory field the high temperature at 85°C data retention of > 10 years at a small electrical stress is necessary. So far RRAM is showing satisfactory testing results as compare to flash devices. Demonstrated parameters of RRAM devices are summarized in Table 3.

Table 3

	Electrochemical Metallization	Metal Oxide		
		Bipolar Filament	Unipolar Filament	Bipolar Nonfilamentary
Mechanism	Electrochemical filament formation	Valence change filament formation	Thermochemical filament formation	Interface switching
Feature size F	20 nm (GeSe), 30 nm (CuS)	5 nm (AlO_x)	35 nm	40 nm
Cell Area (2D)	4 F^2	4 F^2	4 F^2	4 F^2
Read Current (HRS/LRS)	-	100 pA/~10 nA	1 nA/20 nA	< 100 nA
Writa Erase time	< 1 ns	< 1 ns	10ns (W), 5ns (E)	10 ns
Retention Time	1000 hr 200°C	3000 hr 150°C	1000 hr 150°C	4 hr 125°C
Write Cycles	10^{10}	10^{12}	10^6	10^6
Key Advantages	scalability, endurance, speed, low energy operation	scalability, endurance, speed, CMOS compatibility	High R_{OFF}/R_{ON} ratio	uniform switching, adjustable current, no forming, nonlinearity

C. New Applications:

Several advantages put forward RRAM technology as compare to the other emerging NVM devices. RRAM devices with crossbar architecture are known as memristor which means memory resistor. Memristor is considered as the 4th fundamental circuital element was invented by Leon Chua in 1971 [10]. The memristor devices with cross-bar architecture are very useful for high density integration [11, 12]. Several studies has described 2D and 3D crossbar RRAM array for both inorganic and organic materials [13]. It has been tested for both of the non-flexible, flexible and trensperent applications [14]. Durable with faster operation is possible by 3D crossbar memory where the storage density is comparable with flash memory. The devices are also suitable for low power application as a lower voltage and current is sufficient to operate RRAM devices. Recent development shows that RRAM is very useful for neuromorphic applications [15]. The synaptic learning procedure i.e. synaptic plasticity, potentiation / depression, spike – time – dependent - plasticity, paired - pulse facilitation, short term memory and long term memory have already been adopted by RRAM devices. It is expected that RRAM technology with modified form will guide the emerging NVM devices to reach a new horizon.

III.. Conclusion:

In summary, RRAM is a type of emerging NVM technology. RRAM is a powerful technique with many advantages as compare to the baseline or prototype memory devices. In future this technology will open many possible applications including high density 3D crossbar storage and neuromorphic devices. However, the basic structure of RRAM is based on a very simple MIM stack. Based on cation migration, anion migration or thermal effect the RRAM devices can be classified into three major catagories ECM, VCM and thermochemical system. Electroforming process is the initial step to break the purity of the insulating film. A continuous application of positive and negative electrical stimulation can switch the devices OFF to ON or vice-versa. The initial HRS will change to LRS by a SET process and the LRS will change to HRS by RESET process. Generally a conductive filament can be formed by mobile ions (ECM) or oxygen vanancies (VCM).

I believe: Every mortal object in this living planet starts its new sun with the same magical word HOPE where H stands for Highly, O stands for Organised/Oriented, P stands for Positive and E stands for Energy.

Similarly the defects are the hope for the RRAM technology. Controlling of the defect or defect engineering may be a key parameter to achieve the best resistive switching condition.

References:

[1] http://www.itrs.net/
[2] D. Kahng and S. M. Sze, "A floating-gate and its application to memory devices", *The Bell System Technol. J.*, vol. 46, pp. 1288–1295, 1967.
[3] W. Banerjee, "Nonvolatile memories using nanoscale IrO_x metal nanocrystals (Doctoral dissertation)", 2012.

[4] W. Banerjee, S. Maikap, T. -C. Tien, W.-C. Li and J.-R. Yang., "Impact of metal nano layer thickness on tunneling oxide and memory performance of core-shell iridium-oxide nanocrystals"., *J. Appl.Phys.*, vol. 110, 074309, 2011.

[5] H. Lv, X. Xu, H. Liu, R. Liu, Q. Liu, W. Banerjee, H. Sun, S. Long, L. Li, and M. Liu., "Evolution of conductive filament and its impact on the reliability issues in oxide-electrolyte based resistive random access memory", *Scientific Report*, vol. 5, 7764, 2015.

[6] R. Waser, R. Dittmann, G. Staikov, and K. Szot, "Redox-based resistive switching memories-nanoionic mechanisms, prospects, and challenges", *Adv. Mater.*, vol. 21, pp. 2632–2663, 2009.

[7] H. Sun, Q. Liu, S. Long, H. Lv, W. Banerjee, and M. Liu, "Multilevel unipolar resistive switching with negative differential resistance effect in Ag/SiO_2/Pt devices", *J. Appl. Phys.*, vol. 116, 154509, 2014.

[8] H. Sun, Q. Liu, C. Li, S. Long, H. Lv, C. Bi, Z. Huo, L. Li, and M. Liu, " Direct observation of conversion between threshold switching and memory switching induced by conductive filament morphology", *Adv. Funct. Mater.*, vol. 24, pp. 5679–5686, 2014.

[9] W. Banerjee, X. Xu, H. Liu, H. Lv, Q. Liu, H. Sun, S. Long, and M. Liu., "Occurrence of resistive switching and threshold switching in atomic layer deposited ultra-thin (2 nm) aluminium oxide cross-bar resistive random access memory", *IEEE Electron Device Lett.*, vol. 36 (4), 333-335, 2015.

[10] L. O. Chua, "Memristor-the missing circuit element", *IEEE Trans. Circuit Theory*, vol. CT-18, pp. 507-519, 1971.

[11] D. B. Strukov, G. S. Snider, D. R. Stewart & R. S. Williams, "The missing memristor found", *Nat. Lett.*, vol. 453, pp. 80-83, 2008.

[12] W. Banerjee, S. Z. Rahaman, and S. Maikap, "Excellent uniformity and multi-level operation in formation free low power resistive switching memory using IrO_x/AlO_x/W cross-point", *Jpn. J. Appl. Phys.*, vol. 51, 04DD10, 2012.

[13] F. Pan, S. Gao, C. Chen, C. Song, and F. Zeng, "Recent progress in resistive random access memories: Materials, switching mechanisms, and performance, " *Mater. Sci. Eng. R-Rep.*, vol. 83, pp. 1–59, 2014.

[14] X. B. Yan, H. Hao, Y. F. Chen, Y. C. Li, W. Banerjee, "Highly transparent bipolar resistive RAM with IGZO semiconducting electrode in IGZO/Ga_2O_3/IGZO structure", *Appl. Phys. Lett.*, 105, 093502, 2014.

[15] Y. Wu, S. Yu, H. -S. P. Wong, Y.-S. Chen, H.-Y. Lee, S.-M. Wang, P.-Y. Gu, F. Chen, M.-J. Tsai, "AlO_x-based resistive switching device with gradual resistance modulation for neuromorphic device application," *IEEE International Memory Workshop*, 2012.

VLSI Design Employing EDA

Krishnendu Dhar[1], Chandan Kumar Sarkar[2], *Senior Member, IEEE*

[1,2]Dept. of ETCE, Jadavpur University. Kolkata, India

[1]krishnendu.dhar16@gmail.com, [2]phyhod@hotmail.com**

I. Introduction

ith rapid scaling in semiconductor technology and stiff competition prevailing in modern semiconductor industry, Electronic Design Automation (EDA) or Electronic Computer Aided Design (ECAD) or CAD tools have gained widespread attention. EDA tools focus on designing integrated circuits (ICs) and assesses an inflowing design for manufacturing promptness. They are also utilized for encoding design functionality into Field Programmable Gate Arrays (FPGAs).

EDA tools for VLSI have surfaced as a windfall in helping VLSI designers to select as well as optimize an assortment of design models with technology. In the midst of rising intricacy in VLSI circuits, the connotation of EDA tools in each facet of VLSI design have enlarged manifold. Enhanced performance and reliable systems with reduced power consumption, size and cost has resulted in today's electronics industry investing heavily on EDA tools in order to reap maximum profits within a taut market window and stay at par with the current technology. The worth of EDA tools can be comprehended by noting its data structures, multifarious algorithms, as well as modeling postulations utilized in logic synthesis and verification, layout synthesis along with timing verification.

II. EDA Design Flow and Approaches

Electronic Design Automation (EDA) has made a considerable influence on the progression of information technology especially by sustaining the fortunate scaling of Moore's Law over the past 4 decades leading to the design of high performance and cost-efficient systems [1].

EDA tools provide the designer the capability to-

- Maintain designs of subjective size and intricacy
- Authenticate functionality without the generation of an archetype
- Visualize and validate the design's functionality prior to completion of the underlying circuitry
- Insert electronic designs in text format and allow the computer to produce the circuit schematic
- Attempt various densities and package related options to monitor effect on performance, cost in addition to reliability
- Build in manufacturability as well as testability
- Make sure speed along with timing objectives are congregated
- Design without considering the steps needed to be taken into account during the fabrication procedure.
- Mechanize the tangible layout course of action

The EDA flow converts the design from an engineer's point of view to a logical representation and finally to physicalization. Typical EDA tool flow features include-

- Synthesis – This maps a hardware description language (HDL) into a standard gate netlist.
- Mapping – This maps the standard gate netlist to a technology.
- Timing Analysis – This evaluates and verifies the timing of a particular circuit.
- Physical Design – This maps the technology-mapped netlist to a geometric layout.

A typical VLSI design flow is shown in fig. 1

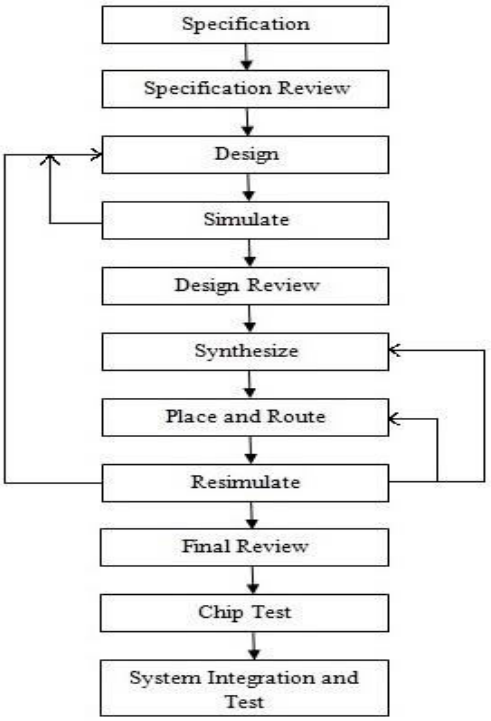


Figure 2 : VLSI Design Flow

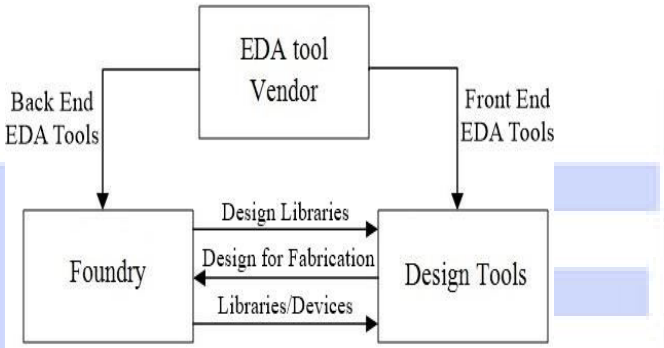


Figure 1: Relationship between EDA tools and design approaches

The liaison amid the EDA tools and design methods is illustrated in fig. 2 which shows that the EDA tool provider supplies both the front end as well as back end tools required for completion of any VLSI design [2]. The front end tools are utilized for design entry editors, simulation synthesis, timing study and test generation tools, whereas the back end tools comprise floor planning, place and route, extraction, and so on. Efficiency of a design depends on the design process employed and can be measured through the EDA tools supplied by the vendor. The design libraries provided by the EDA tools are in compliance with the data delivered by the foundry.

Table I puts forward some of the commonly used EDA tools employed for VLSI design along with their license features in addition to their functionality [3]. One of the most commonly used EDA tools utilized for digital circuit simulation and which can operate in Windows environment is Tanner EDA tools. They have the ability to support every key industry standards embracing Graphic Database System (GDS), Caltech Intermediate Form (CIF), Drawing eXchange Format (DXF) and Gerber. Also, Tanner EDA tools support ad-hoc standards like Simulation Program with Integrated Circuit Emphasis (SPICE) and enable inter operability with diverse dialects such as HSPICE, PSPICE and Circuit Description Language (CDL) in addition to sustaining external formats like Virtuoso technology files, DRACULA DRC-rule decks, and Calibre DRC-rule decks. It is to be mentioned that Tanner EDA tools are the only EDA tools in the market that natively back Mentor Graphics Calibre and Cadence Dracula formats thereby enabling utilization of foundry rule decks devoid of translation of rules to a different format. This as a consequence reduces chances of data loss. Whilst the foundry revises its rule decks, Tanner EDA solutions instinctively absorb the modifications ensuring that the designer stays in row with the foundry

For any VLSI design, it is vital to evaluate its functionality along with its performance parameters. EDA tools play an important role in this regard. It helps to estimate the design quality metrics before actual fabrication of the proposed design. Some of the chief design quality measures that a VLSI designer must take into account while designing any circuit are listed as follows:

- Area – It is regarded as one of the most fundamental parameters of any VLSI design. Reduced area results in reduction of Silicon area required to implement a particular circuit.
- Congestion and Wireability – The routing phase is considered as the most time consuming step during the layout process. Improper routing can lead to ineffective management of the routing area thereby deteriorating circuit performance in terms of reliability and yield while increasing crosstalk and coupling noise. It is therefore suggested that wireability as well as wire congestion be considered beforehand in the design cycle.

- Crosstalk and Coupling Noise – Inappropriate interconnections result in undesirable crosstalk and coupling noise [4] which as a consequence adversely affects circuit performance. For instance, coupling noise might result in erroneous signal values on logic lines or increase the power consumption owing to instantaneous glitches on the logic lines. In deep submicron technologies, propagation delay also gets adversely affected due to crosstalk and coupling noise.
- Delay – In contemporary VLSI design routines, delay is contemplated as a vital performance parameter. Reduction in delay time entails faster propagation speed of the circuit. It is to be stated that optimization of delay is more difficult than area as the latter is deemed an aggregate metric where myriad over and under estimates tend to cancel one another ensuing in a cumulative result which is potentially more precise than the area guesstimates of the individual components. Conversely, delay is considered as a path based metric that is additionally more sensitive to its individual constituents.
- Power Consumption – Earlier, processor speed, performance, reliability, area, cost were of imperative concern. Power consumption was by and large of inferior angst. However, in modern day VLSI design, power consumption is being given equivalent importance due to the brisk escalation in portable and wireless systems which insist high speed calculations and complex operability with low power consumption. Currently, applications are intended at battery driven devices such that power dissipation turns out to be one of the prime design constraints [5]. Lesser the average power consumption, lesser will be the cooling and packaging expenses whereas reduction in peak power consumption signifies improved reliable functionality.

TABLE I.; COMPARATIVE STUDY OF DIFFERENT EDA TOOLS FOR VLSI DESIGN

Serial Number	EDA Tool	Open Source/ Licensed	Operation
1.	Tanner EDA	Licensed	Complete CAD Flow
2.	Cadence EDA	Licensed	Complete CAD Flow
3.	Synopsys EDA	Licensed	Complete CAD Flow
4.	Mentor Graphics EDA	Licensed	Complete CAD Flow
5.	Electric CAD	Open Source	Logic to Layout
6.	Alliance	Open Source	Logic to Layout
7.	Magic	Open Source	Circuit Layout
8.	SystemC	Open Source	Library for Digital Design
9.	myHDL	Open Source	Hardware Description Language

.For any VLSI design, it is vital to evaluate its functionality along with its performance parameters. EDA tools play an important role in this regard. It helps to estimate the design quality metrics before actual fabrication of the proposed design. Some of the chief design quality measures that a VLSI designer must take into account while designing any circuit are listed as follows:

- Area – It is regarded as one of the most fundamental parameters of any VLSI design. Reduced area results in reduction of Silicon area required to implement a particular circuit.
- Congestion and Wireability – The routing phase is considered as the most time consuming step during the layout process. Improper routing can lead to ineffective management of the routing area thereby deteriorating circuit performance in terms of reliability and yield while increasing crosstalk and coupling noise. It is therefore suggested that wireability as well as wire congestion be considered beforehand in the design cycle.

- Crosstalk and Coupling Noise – Inappropriate interconnections result in undesirable crosstalk and coupling noise [4] which as a consequence adversely affects circuit performance. For instance, coupling noise might result in erroneous signal values on logic lines or increase the power consumption owing to instantaneous glitches on the logic lines. In deep submicron technologies, propagation delay also gets adversely affected due to crosstalk and coupling noise.
- Delay – In contemporary VLSI design routines, delay is contemplated as a vital performance parameter. Reduction in delay time entails faster propagation speed of the circuit. It is to be stated that optimization of delay is more difficult than area as the latter is deemed an aggregate metric where myriad over and under estimates tend to cancel one another ensuing in a cumulative result which is potentially more precise than the area guesstimates of the individual components. Conversely, delay is considered as a path based metric that is additionally more sensitive to its individual constituents.
- Power Consumption – Earlier, processor speed, performance, reliability, area, cost were of imperative concern. Power consumption was by and large of inferior angst. However, in modern day VLSI design, power consumption is being given equivalent importance due to the brisk escalation in portable and wireless systems which insist high speed calculations and complex operability with low power consumption. Currently, applications are intended at battery driven devices such that power dissipation turns out to be one of the prime design constraints [5]. Lesser the average power consumption, lesser will be the cooling and packaging expenses whereas reduction in peak power consumption signifies improved reliable functionality.

III. CONCLUSION

Electronic Design Automation (EDA) tools assist in designing novel VLSI circuits in diverse process technology and also help in determining the circuit functioning in addition to estimating the corresponding performance factors like power consumption, delay, and area. It is through the aid of EDA tools that realization of integrated circuits (ICs) employing billions of transistors in scaled technology is made possible. EDA cultivates and embraces theories of computation and modeling thus transforming the procedure electronics engineers employ to design as well as concoct ICs. It is a product of inter-disciplinary collaboration between computer scientists, electronics engineers, electrical engineers, physicists, chemists, applied mathematics and optimization professionals, as well as application domain experts.

REFERENCES

[1] G Robert Brayton and Jason Cong, "Electronic Design Automation Past, Present, and Future", *NSF Workshop Report*, Arlington, Virginia, July, 2009.

[2] Course: "Trends in VLSI Design: Methodologies and CAD tools". Presenter Raj Singh. IC Design group, CEERI, Pilani-333031.

[3] http://www.vlsiacademy.org/open-source-cad-tools.html

[4] K.L. Shepard, V. Narayanan, and R. Rose. " Harmony: Static Noise Analysis of Deep Submicron Digital Integrated Circuits", IEEE Transactions on Computer Aided Design, 18(8):1132-1150, August 1999.

[5] Krishnendu Dhar[1], Aanan Chatterjee[2], Sayan Chatterjee[3], "Design of an Energy Efficient, High Speed, Low Power Full Subtractor Using GDI Technique", Proceeding of the 2014 IEEE Students' Technology Symposium, Indian Institute of Technology-Kharagpur (IIT-Kgp), pp. 199-204.

A Design Flow Illustrating the Phases of FPGA Modeling and Implementation

[1]Krishnendu Guha, [2]Amlan Chakrabarti, *Member, IEEE*

A.K. Choudhury School of I.T., *University of Calcutta, Kolkata, India*

[1] mail2krishnendu@gmail.com, [2]amlanc@ieee.org

I. NTRODUCTION

The advent of Field Programmable Gate Array (FPGA) architectures has been an important aspect in the evolution of the embedded domain. A traditional VLSI design flow is associated with the development of Application Specific Integrated Circuits (ASIC). However, this is costly and time consuming. Though suitable for mass production, it is ineffective for experimentation and reuse. FPGA's provides a solution to these problems which can be used as a prototype before actual ASIC realization is actuated. The architecture of a FPGA being reconfigurable makes it perfectly suitable for experimentation and reusability.

Xilinx and Altera are the top two companies taking up a possession of 67% of the FPGA market [1]. The top two have done a good job over the years in defending the duopoly in the market. However, the present era has witnessed a grasp of the market share by other companies which targets specific applications and sub markets. Fig. 1 provides a graphical representation of the market share of different FPGA vendors in 2013.

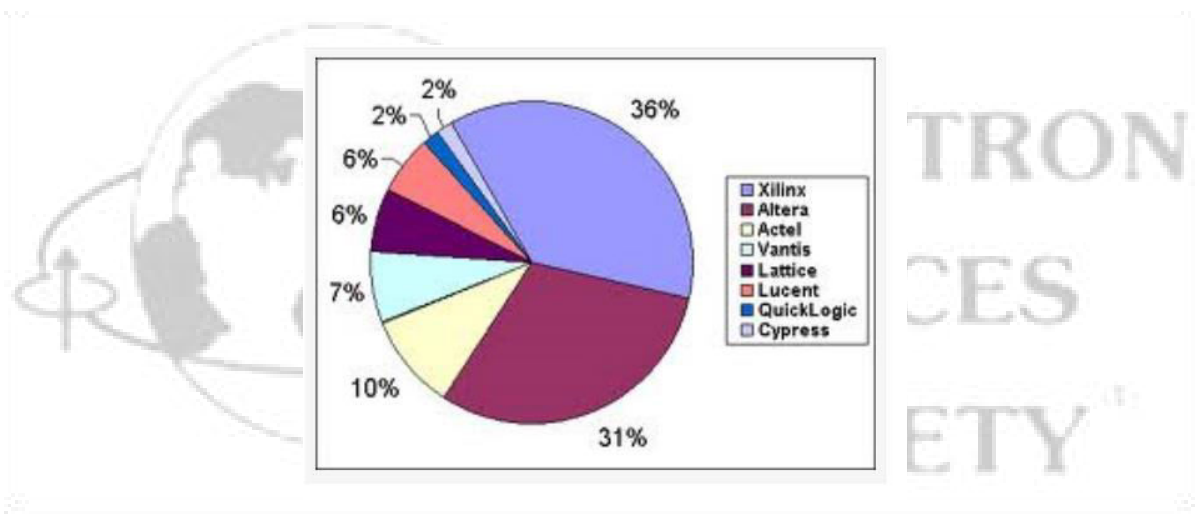


Fig. 1: FPGA market share, 2013 [1]

The present era has witnessed an extensive usage of FPGA architectures due to the following five factors:

(i) Performance: With the aid of hardware parallelism, FPGA's are able to break the paradigm of sequential execution of instructions and accomplish more per clock cycle. Thus, an increment in the computing power of Digital Signal Processors (DSP) has been accomplished [2].

(ii) Time to market: A prevalent challenge in the present era is to meet the stringent marketing deadlines. FPGA's being flexible with its capability of rapid prototyping is a good option to test a novel concept and implement it in hardware without going into the long custom ASIC design methodology [3].

(iii) Cost: The need of end users for custom hardware functionality is growing day by day. Moreover, the concept of programmable silicon means no fabrication costs. As system requirements change often over time in practical scenarios, the cost of making incremental changes to FPGA designs is negligible when compared to the large expense of ASIC architectures.

(iv) Reliability: The programming environment is provided by software tools while FPGA circuitry aids in the implementation of program execution. A processor based system often incorporates several layers of abstraction for scheduling tasks and sharing resources, while the driver layer controls hardware resources and the operating system manages memory and processor bandwidth. A processor core can execute one instruction at a time and hence, such processor based systems are always at risk of preemption. However, FPGA's minimize such reliability concerns as they do not use operating systems and facilitates parallel execution as deterministic hardware is dedicated to every task.

(v) Long term maintenance: As a system matures, a modification is inherent for the system. Considering a digital communication system, a change or update in the protocols is inherent. FPGA's being compatible for reprogramming is perfect for long term maintenance of such systems.

FPGAs are already being used in industry as an end product. Recent research has exposed its potential to be exploited in the domain of Internet of Things (IOT) [4]. Hence, experimentation and research on FPGA looks promising in near future.

This article is focused to serve as a preliminary for novice designers who wish to have a flair knowledge of modeling and implementing a design on FPGA. A brief introduction on the internal structure of a FPGA is provided in section II before diving into the intricacies of FPGA modeling and implementation. In section III, we illustrate the different phases associated in the design flow of modeling and implementing a design on a FPGA. Finally, we conclude this article in section IV.

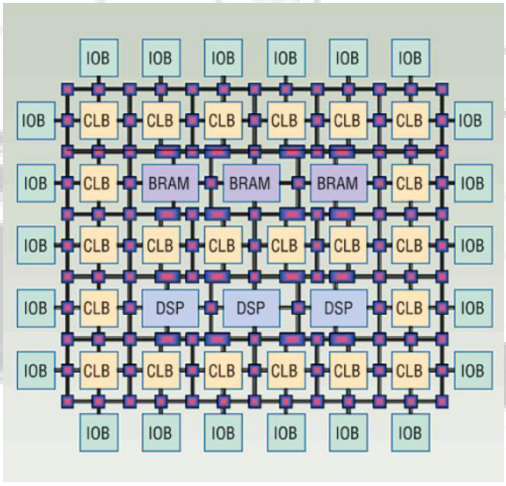


Fig. 2: Structure of a FPGA [7].

IV. STRUCTURE OF FPGA

A FPGA can be considered as a programmable semiconductor chip [5]. The elements of a FPGA comprise of a matrix of Configurable Logic Blocks (CLB's), memory elements (BRAM's), Digital Signal Processing (DSP) cores and a set of peripherals (IOB's). Interconnecting routing structures are present amongst the elements of the FPGA which facilitates communication and synchronization in the operations of different units of a FPGA. The basic logic unit of a FPGA is the CLB. The exact number of a CLB and its features varies from device to device but each CLB comprises of a configurable switch matrix with a set of inputs, a multiplexer and a set of flip flops. The flexibility of the switch matrix aids in

handling combinatorial logic. The on chip memory of a FPGA is enabled with the aid of embedded Block RAM (BRAM). The structure of a FPGA is shown diagrammatically in Fig. 2.

V. PHASES IN THE DESIGN FLOW OF FPGA MODELING AND IMPLEMENTATION

In this section, we will illustrate the different phases associated with the modeling of a design on a FPGA. A graphical representation of the FPGA design flow is shown in Fig 3.

A. Phase 1: Analysis and Planning

A basic plan with novel concepts is a prerequisite for solving a problem and attaining a desired objective. Thus, it is a must for a designer to understand the requirements and select a FPGA board based on the analysis of its requirements.

The requirements for a design are generally documented in a System Requirement Specification (SRS). The choice of FPGA is based on the analysis of the SRS. It depends on the operating frequency, I/O voltage levels, type of processor interfaces, external peripheral interfaces and memory requirement needed by the intended design architecture. Determination of IP cores which are available with the tools or FPGA family is also important.

In addition to this, a conceptual level block diagram of the intended architecture is framed by the designer before beginning with the modeling of the architecture. This serves as a basic plan to explain the functionalities of different parts of the architecture and its associated timing constraints.

B. Phase 2: Design Modeling

Modeling a design enables a designer to structure the novel concepts and realize the desired objective. This is generally done with the aid of HDL (Hardware Description Language) programming or by designing a schematic which are enunciated below. A comparison between these two models of designing is tabulated in Table I.

(i) ***HDL Programming:*** In HDL Programming, the design is framed with the aid of hardware description languages like *VHDL (Very High Speed Integrated Circuit Hardware Description Language)* or *Verilog*. The structure of a HDL program varies with the choice of the designer. Dataflow, behavioral and structural modes of designing are commonly used. However, for practical applications, a combination of all is generally preferred by designers.

(ii) ***Schematic Design:*** This is a powerful design methodology which aids to create a design with the aid of a combination of FPGA library macros or HDL modules. Such macros may be custom made or provided by the vendors. It may include lower level HDL modules or other schematics or IP cores which it treats as black boxes. However, a synthesis tool converts such schematic files to structural HDL before implementation.

TABLE I: A Comparison of HDL Programming and Schematic Designing

	HDL Programming	**Schematic Designing**
Ease of Designing	Requires experience	Simple and novice
Portability	Portable	Not portable
Flexibility	Flexible (Easy to modify)	Not flexible
Debugging	Hard	Easy
Productivity (wrt gates and time)	Good	Poor

C. *Phase 3: Functional Simulation*

This phase is essential to verify the correctness of the design or to confirm that the HDL code written or the schematic drawn, functions as per specification. This phase is often termed as *Behavioral Simulation*.

In a practical scenario, architecture comprises of several modules. These modules might be designed by

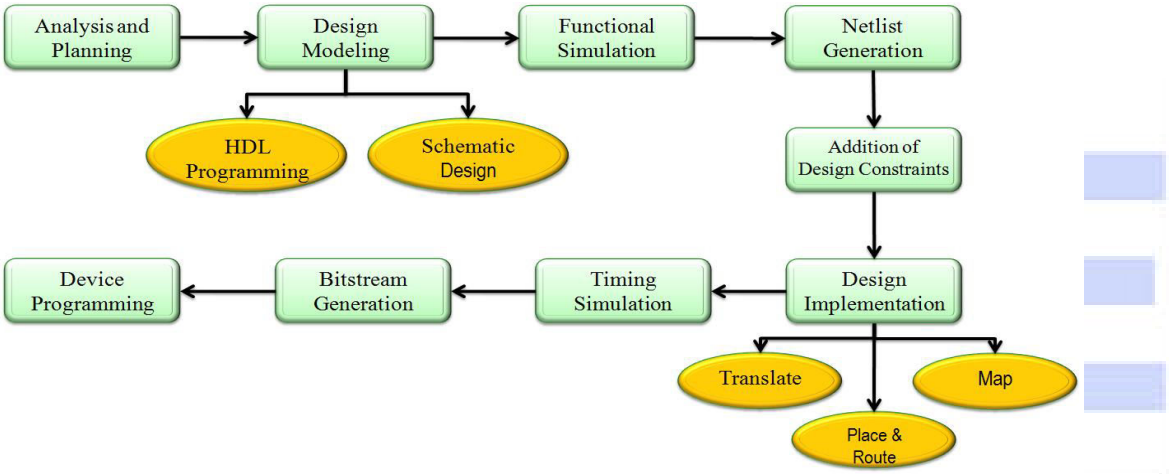


Fig.3: A diagrammatic representation of FPGA design flow

different designers which are finally integrated to generate a system. In this phase, the functionality of the different modules is verified separately and also after integration, the functionality of the entire system is verified.

HDL test benches or test bench waveform files like *Verilog Test Fixture* or *VHDL Test Bench* are commonly used in this phase to verify the correctness of the modeled architecture. A snapshot is provided in Fig. 4 depicting a test bench waveform which is used for functional simulation.

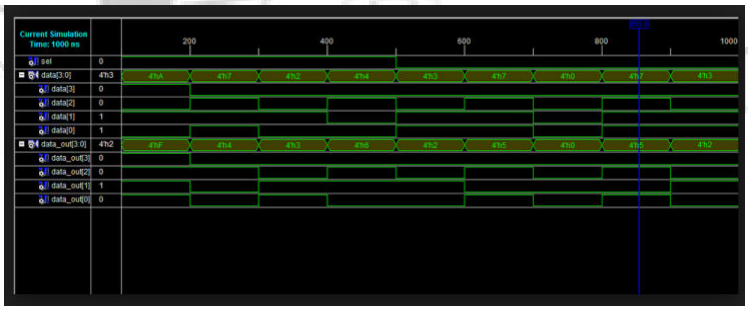


Fig. 4: Functional Simulation

D. Phase 4: Netlist Generation

This phase is entered after verification of each individual module as well as the entire system as a whole. Synthesis of the design in this phase is based on the libraries of the chosen vendor specific tool. The popular synthesis tools includes *Synplify, Precision, FPGA Compiler II and XST (Xilinx Synthesis Technology)*.

In this phase, the synthesis tool performs the following operations:

(i) Checks the syntax errors of the HDL files

(ii) Translates the HDL files into gate level netlists

E.g. An adder functional module is assigned to the architecture where a '+' operator is encountered in the HDL code, or a comparator functional module is assigned to the architecture where a '= =' operator is encountered in the HDL code.

(iii) Optimizes the design architecture based on the design constraints of the library associated with the chosen FPGA tool.

The outcome of the synthesis phase is a netlist file which comprises of a list of components and their interconnections as shown in Fig. 5. The common netlist formats are *EDIF (Electronic Design Interchange Format)* or vendor specific formats like *XNF (Xilinx Netlist Format)*.

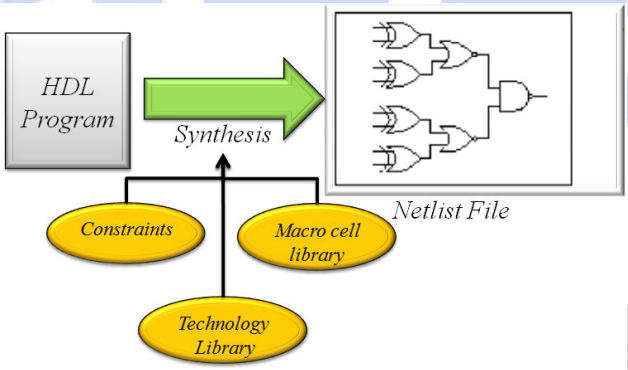


Fig. 5: Netlist Generation

E. Phase 5: Addition of Design Constraints

After completion of synthesis, constraints can be included with the aid of a separate file by the designer. This is commonly known as *User Constraint File (UCF)*. This file comprises of a list of signals with its corresponding FPGA pin number, I/O voltage levels, current driving strength for output signals, input clock frequency, timing constraints and other specific IP constraints which are recommended by the vendor. This information is crucial for implementing the design on the FPGA.

F. Phase 6: Design Implementation

The three principal components of this implementation phase are described as follows:

(i) **Translate:** In this phase, multiple design files are merged to form a single netlist.

(ii) **Map:** In this stage, the logical symbols of netlist files, i.e. gates are mapped into the physical components of the FPGA fabric, i.e. slices and IOB's.

(iii) **Place and Route:** Based on the specified user constraints and optimization rules, the synthesis tool places and routes the design on the FPGA fabric in this phase. The components are placed onto the chip and interconnection is made among them for communication.

Fig. 6 shows the implementation of a design in the FPGA fabric from a netlist file.

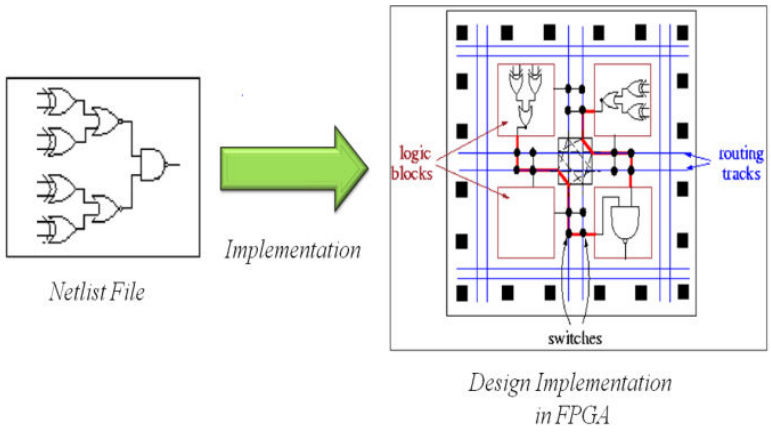


Fig. 6: Design Implementation

G. *Phase 7: Timing Simulation*

Verification of the design after the implementation stage (translation, mapping, placement and routing) is performed in this phase. Testing in this phase is considered successful if all the functional and timing requirements are realized.

This is often called a post layout simulation as shown in Fig. 7 which includes 'component delay', 'wire delay', 'clock skew' and 'setup and hold' time. The input vectors are same as for the functional simulations.

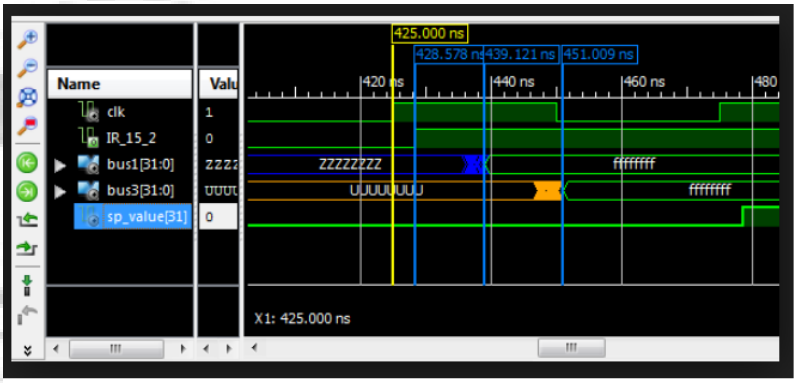


Fig. 7: Timing Simulation

H. *Phase 8: Bitstream Generation*

After successful implementation and verification of a design, a file must be created which the FPGA can understand. This is called a bitstream or a BIT file (a file with .bit extension). Generation of the FPGA programming file or the configuration bitstream of the designed architecture is obtained in this phase.

I. *Phase 9: Device Programming*

The generated BIT file can be directly downloaded to the FPGA via a download cable or can be converted to a PROM file which stores the programming information. This aids to test the functionality of the designed architecture on the FPGA board with actual peripherals as shown in Fig. 8.

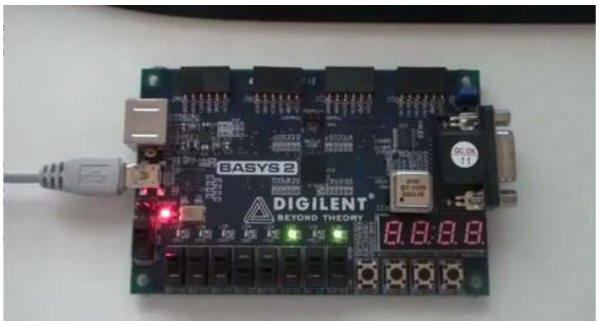


Fig. 8: Device Programming [6]

VI. CONCLUSION

Modeling and implementing a design on a FPGA is of utmost importance in the present, evolving era of embedded systems. Its property of being reconfigurable makes it perfectly suitable for experimentation and practical implementation where a system is subject to change with respect to time. Recent usage of FPGA in the domain of IOT is an indication to its bright future. However, a descriptive design flow illustrating the stages of modeling and implementing a design on a FPGA is of utmost importance to the novice designers. This article initially gives a basic introduction to the structure of a FPGA and then illustrates the phases involved in modeling and implementing a design on a FPGA.

REFERENCES

[1]. "Top FPGA Companies For 2013"', Source Tech411 (April 28, 2013), Source: http: //sourcetech411.com/ 2013/04/ top-fpga-companies-for-2013/.
[2] Berkeley Design Technology Inc., "FPGAs for DSP", BDTI Industry Report, 2nd ed., 2006.
[3] M. Thompson, ``FPGAs Accelerate Time to Market for Industrial Designs," EE Times (July 2, 2004), http://www.us.design-reuse.com/articles/8190/fpgas-accelerate-time-to-market-for-industrial-designs.html.
[4] A. P. Johnson, R. S. Chakraborty, D. Mukhopadhyay, "A PUF-enabled Secure Architecture for FPGA-based IoT Applications", IEEE Transactions on Multi-Scale Computing Systems, no. 99, 2015.
[5] Xilinx, "What is a FPGA", Source: http://www.xilinx.com/fpga/.
[6] Digilent, "Basys™2 Spartan-3E FPGA Board", Source: http://www.digilentinc.com/Products.
[7] D. Buell, T. El-Ghazawi, K. Gaj, V. Kindratenko, "Guest Editors Introduction: High-Performance Reconfigurable Computing", Computer, vol. 40, no.3, pp. 23-27, 2007.

IEEE

Electron Devices Society