Journal of Applied Sciences Research, 8(8): 4138-4146, 2012
ISSN 1819-544X
This is a refereed journal and all articles are professionally screened and reviewed

ORIGINAL ARTICLES

Cmos Downsizing: Present, Past And Future

Siti Sarah BintiMdSallah, Habibah Mohamed, Md. Mamun, Md. Syedul Amin

Department of Electrical, Electronic and Systems Engineering Universiti Kebangsaan Malaysia, 43600 UKM Bangi, Selangor, Malaysia.

ABSTRACT

The evolution process of Complementary Metal Oxide Semiconductor (CMOS) is very important for the modern technology. CMOS beyond 22nm and toward 7nm faced many challenges and opportunities in design.The downsizing evolution is reviewed on scaling theory as well as limitation issues focusing on performance, power utilization, economical, technological and reliability issues.It is expected that the CMOS sizelimit of 7nm of physical gate length will be overcome in 2018 by using High-*k* materials. Besides, high-*k* dielectric materials can reduce the current leakage problem. Behind the transistor downsizing, the lithography technology is one of the key important processes. Several issues are being highlighted and discussed in term of performance, power utilization, material, economical, and technological limitations.

Key words**:** 7nm Gate Length, CMOS, Downsizing, High-k, VLSI.

Introduction

The invention of the CMOS Large Scaled Integrated (LSI) into nanometer regime has become a serious issue in integrated circuit(IC) of modern human society (Akter *et al*. 2008a, b; Reaz *et al*. 2007a, b; Marufuzzaman *et al*. 2010; Reaz *et al*. 2003; Reaz *et al*. 2005; Iwai, 2012).Today's advanced communication and engineering technology could be never imagined without the latest massive evolution of the integrated circuit(Iwai, 2003; Reaz *et al.* 2006; Reaz and Wei 2004; Mohd-Yasin *et al*. 2004; Mogaki *et al*. 2007). Besides, daily living, manufacturing, commerce, transportation, medical treatment, education etc cannot functionwithout thesupport of the CMOS technology(Iwai, 2008).Thus, the evolution process of CMOS technology is veryimportant from the semiconductor industry and global economical point of view.The electronic circuit evolution with the component dimensionis shown in Figure 1(Iwai and Ohmi, 2002).

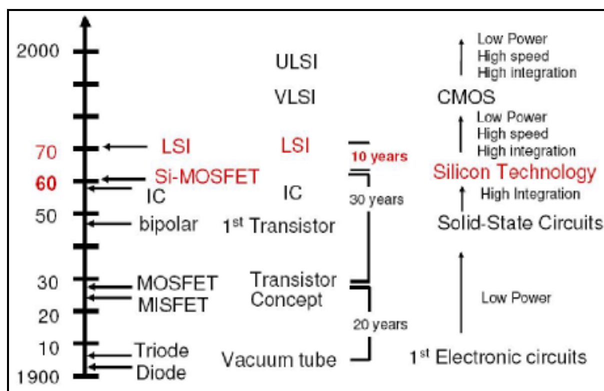

Fig. 1: The downsizing of component dimension.

There are lot of benefitsfrom CMOS downsizing which includeshigherperformance, lower power consumption, higher density and lowercost(Kin and Park, 2011). LSI products are sold well in the retail rapidly to replace the products of the previous generation.In this paper, the evolution of CMOS downsizing technology is reviewed on scaling theory as well as limitation issues focusing on performance, power utilization, economical and technological and reliability issues.

Corresponding Author: Md. Syedul Amin, Department of Electrical, Electronic and Systems Engineering Universiti Kebangsaan Malaysia, 43600 UKM Bangi, Selangor, Malaysia.
E-mail: amin.syedul@gmail.com

Background:

For more than 40 years, the semiconductor industry has recognized itself by the fast trackof advancement in its products. In 1965, Gordon E. Moore, the co-founder of Intel,expected that the number of transistor which can be placed in cutting edge IC chips is doubling approximately every two years without correspondingly increasing the cost of the chip (Moore, 1965). Figure 2shows the Moore's Law which relate to the number of transistors in Intel's Microprocessors. This prediction continuespractically true as microprocessors observedtheir number of transistors rise from 2250 in Intel's 4004(1971) to 731 million in Intel's Pentium iCore 7(2008).

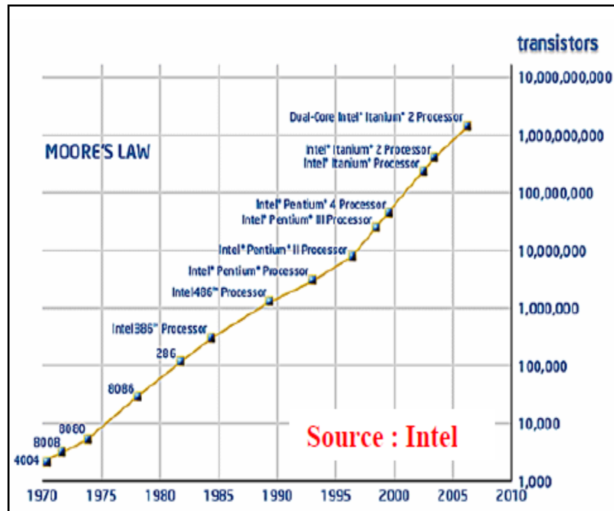

Fig. 2: Moore's Law in Microprocessors.

This extraordinaryevolution of microprocessors has been made viable by frequent downsizing of the metal oxide semiconductor (MOS) field effect characteristic. As they become smaller, they also become cheaper, utilize less power, turn out to be faster, and allowadditionaltasks per unit area of silicon.Thus, the compactness of silicon IC is able to understand, recommending excellentoperation at cut downprice per function. In reality, the perfect scaling method proposed by Dennardboost the performance and integration excellently without slightly increase in power utilizationon condition that the chip area is kept stable (Dennard *et al*. 1972). On the other hand, the actual scaling development for the 30 years from 1970 was more aggressive. Table 1 shows the improvement of ICs with examples of each trend.

Table 1: Improvement trends for ICs enabled by feature scaling.

Trend	Example
Integration Level	Components/ Chips, Moore's law
Cost	Cost Per Function
Speed	Microprocessor throughput
Power	Laptop or cellphone battery life
Compactness	Small and light weight products
Functionality	Non-volatile memory, imager

The basic technology behind this evolution is the CMOS downsizing. MOS transistor is the basic building block of the information infrastructure. Its scaling has provided path toward ever denser and faster integration. Figure 3 shows the basic structure of CMOS.

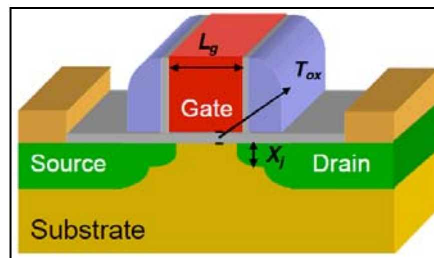

Fig. 3: The structure of CMOS.

The main theory of downsizing is to lessen all the sizes of a transistor by a factor, κ. In addition to reducing size, certain parameters must be increases or decreased. As the channel length is reduced, the performance improves.The power per switching goes down and the density improves. Figure 4 shows the CMOS downsizing and the Table 2 summerizes the CMOS parameters involved in downsizing.

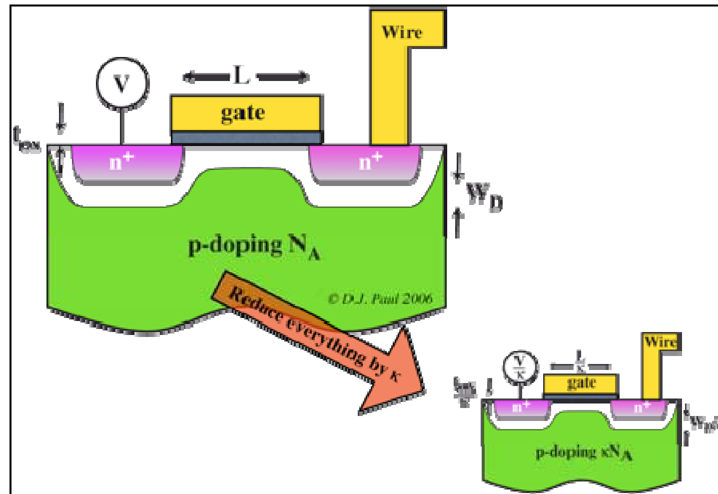

Fig. 4: The illustration of CMOS downsizing.

Table 2: The CMOS parameters involved in downsizing.

Parameter	Constant F downsizing	Generalized Downsizing
Device Dimension (L_g, W, t_{ox})	$1/\kappa$	$1/\kappa$
Doping Concentration	κ	$\alpha\kappa$
Voltage (V)	$1/\kappa$	α/κ
Electric Field ($\vec{F}$)	1	α
Carrier velocity (v)	1	1
Depletion Layer Width (W_D)	$1/\kappa$	$1/\kappa$
Capacitance, ($C = \varepsilon A/t_{ox}$)	$1/\kappa$	$1/\kappa$
Current (I)	$1/\kappa$	α/κ
Circuit Delay Time ($\tau \sim CV/I$	$1/\kappa$	$1/\kappa$
Power Dissipation per Circuit ($P \sim IV$)	$1/\kappa^2$	α/κ
Power-delay product per circuit ($P\tau$)	$1/\kappa^3$	α/κ
Circuit Density ($\propto 1/A$)	κ^2	κ^2
Power Density (P/A)	1	α^2

Discussion:

Evolution In Cmos Downsizing:

The evolution of CMOS downsizing has been achieved by the downsizing of their components such as MOSFETs.Ever since their beginning of early 1970's, there were many downsizing limits from the 0.8 µm generation. In the middle of 1980's, 1 µm was believed as the feasible limit due to the predicted problem in restraint of the short-channel effects and also due to the predicted problem in optical lithography(Iwai, 2004). However, in late1990's, 0.25 µm was advised as the limit due to the increment of the estimated source/drain resistance and predicted direct-tunneling leakage within the gate oxides and dopant fluctuation in the channel (Iwai, 2009). Additionally, in the beginning of the year 2000's, 100 nm was assumed to be the end point due to many predicted problems in minimizing the physical factors of MOSFETs. Luckily, each expected limit has been verified wrongly by the successful evidence of commercial products using a smaller MOSFET size compared to the expected by the limits.

At the moment, the 22nm node is the CMOS procedure phase resulting 32 nm. Based on the Figure 5, in the CMOS generations from 90 nm to 22 nm shrinking, the transistor reduced the performance. However, the

boosters were added (strain polysilicon in the 90 nm and 65 nm nodes and strain +High-*k* metal gate in the 45 nm and 32 nm nodes) to keep on to steer the transistor roadmap ahead.

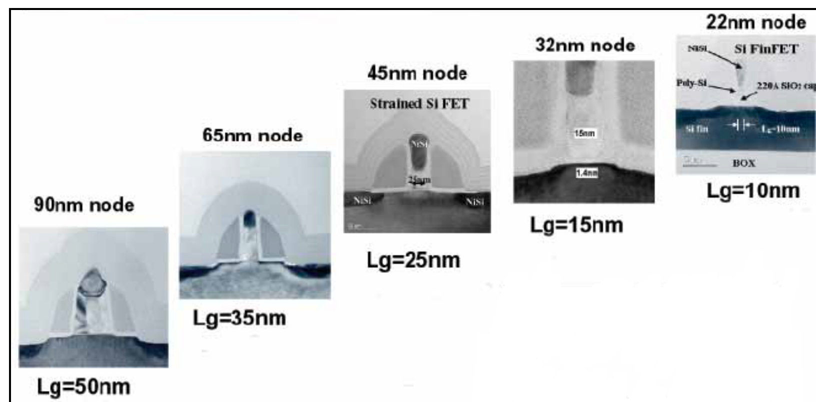

Fig. 5: Transistor Continuous Scaling.

The future replacement to 22 nm gate length will be 7 nm gate length, predicted by International Technology Roadmap for Semiconductor (ITRS) for 10 years up to 2018(Iwai, 2009). Gate oxide thickness must be two guidelines of magnitude smaller than that of the gate length (0.7 nm). At the moment, the 1.2 nm thick oxynitride film being utilized in manufacturing and it is predicted the silicon dioxide equivalent thicknesswill be decreased by 0.5 nm in ten years later(Iwai, 2004). Even though it is proved that a MOS transistor with 0.8 nm oxynitride gate insulator is still operating, several critical issues are expected when recognizing an LSI with such a thin gate insulator. It is not easy to expect the limit of the downsizing.There are several possible limits have been recommended for the downsizing as shown in Table 3.

Table 3: Predicted limitations for downsizing.

Year	Expected Limit Size
1971	10 µm
1975	3 µm
1982	1.5 µm
1985	1 µm
1989	800 nm
1994	600 nm
1995	350 nm
1998	250 nm
1999	180 nm
2000	130 nm
2002	90 nm
2006	65 nm
2008	45 nm
2010	32 nm
2011	22 nm
Future(approximate 2013)	16 nm
Future(approximate 2015)	11 nm
Future (approximate 2018)	7 nm

Limitations in CMOS Downsizing:

Future Scaling trends have been predicted by ITRSfor the ten years up to 2018 while the physical gate length is predicted to be 7nm (Iwai and Wong, 2006). There are some challenges for the integration in the CMOS downsizing which are performances challenges, power consumption challenges, economical challenges and technological challenges.

Performance Limitations:

The major problems for performance degradation in the ultra-large scale circuits is the interconnect delay due to the increment in the resistance and the capacitance (Kuhn, 2009) values of narrow and dense interconnection metal lines.For instance, while the dimension of copper wire is cut to less than 100nm width, the resistivity of the conductor will raise surface scattering effect, as shown in Figure 6 (Iwai, 2007). In accordance with scaling theory, the drain current per unit gate width ought tokeep onstable. Nevertheless,

animportantdecrease of the drain current value per unit gate width for sub-100 gate length MOSFETs was reported recently (Iwai, 2007). This fact is due to the non-optimized MOSFET structure and procedure.

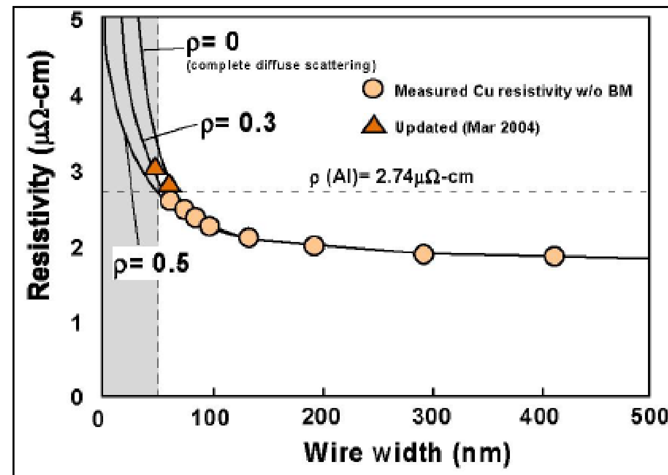

Fig. 6: Increase of resistivity because of surface scattering effect.

At this time, inadequateimprovement of MOSFET structure and procedure is considered to be the cause for this and the beginning of new technologies.For example,raised source drain, plasma doping combined with flash or laser strengthening, NiSisalicide, mobility improvement by strained Si channel, silicon on insulator (SOI) and three-dimensional structure, high dielectric constant (high-*k*) gate insulator, and low dielectric constant (low-*k*) interlayer for interconnects are likely to definitelyanswer the problems and increase the performance (Kuhn, 2011). Low-*k* materials apply for back end procedureundergo from high mechanical and thermal stress throughout packaging stage. The high-*k* materials may possiblyreduce the current leakage problem when the dielectric is made to turn out to benarrower to sustain physical scaling (Kuhn, 2011). In spite of this, the trend of these materials to transform their properties in high temperature is among the challengesrequiring to be explainedseparately of adding additionalindustrial process.

Besides that, the enhancement of new circuit and system design architecture is also increases the performances of the integrated circuits. Parallel processing and optimized interconnect with the support of steeringdevices are two best samples. As explainedafterward, system-on-chip(SOC)approach by embedding the DRAM in the logic unit enhances the transfer speed between logic and memory units (Haron and Hamdioui, 2010). It is predicted that on the whole, performance of electronic system might be increasedmore at least down to the production of 20 or 10nm lengths as a result of both the bettertool technology and new system structures. On the other hand, for sub-10nm gate length transistors, the drain current level should be better.

Material Limitations:

For the scaling of CMOS transistors, new materials are proposed as shown in Figure 7(Haron and Hamdioui, 2010). It is because Silicon (Si), Silicon dioxide (SiO_2),Aluminium (Al), Copper (Cu) and Salicides are restricted by their physic potentials such as relative dielectric constant (ϵ), carrier mobility (μ), carrier saturation velocity (v_s), breakdown field strength (E_c) and conductivity. Since, these materials achieve their physical limit, a device cannot sustain with their operation performance. Even though Cu is less susceptible to electromigration than Al, the substance is more susceptible to open defect when used as interconnect wires. The ultra-thin gates SiO_2 were appeared to be more reliable due to the better consistency for the thinner film and furthermore due to a small amount of trapped chargers caused by tunneling. The new materials such as high-*k* gatedielectrics and new structure such as three dimensional MOSFETs should be carefully introduced to avoid fatal problem in reliability and yield issues for integration (Martin, 2011).

Power Utilizationlimitations:

Power utilization for high performance logic integrated circuit is the limiting aspect of the logic CMOS. Lowering the supply voltage is the bestefficientapproach to reduce the dynamic power utilization(Iwai, 2009). If the current tendency in the chip frequency and the number of the transistors of the chip persist, power utilization of a high performance microprocessor would turn out to be 10kW within several years and heat generation at the

silicon chip surface would become 1000W/cm^2. It is equal point of the surface of a rocket nozzle (Iwai, 2007). The important increase in power compactness is the result of the deficient supply voltage decrease and the exponential increase in the transistor density. Low voltage technology (even though is complicated) and suitable control of the increase in chip density and chip dimensionat the same time with certain new cooling technologies possibly willin some measureanswer this problem. In addition, a few innovative system power management processes, for example variable clock frequency and variable voltage supply, will also assist in this matter. Alternatively, the gate leakage current can be decreased by using a materially thicker high-K dielectrics and the sub-threshold leakage current can be covered up by using a three dimensional(3D) structure, such as fin-FET(Iwai and Wong, 2006) as shown in Figure 8. Those quotas would be proposed for low stand-by power gadgets, even earlier than the high performance logic units recognized and the manufacturing cost becomes rational(Iwai and Wong, 2006).

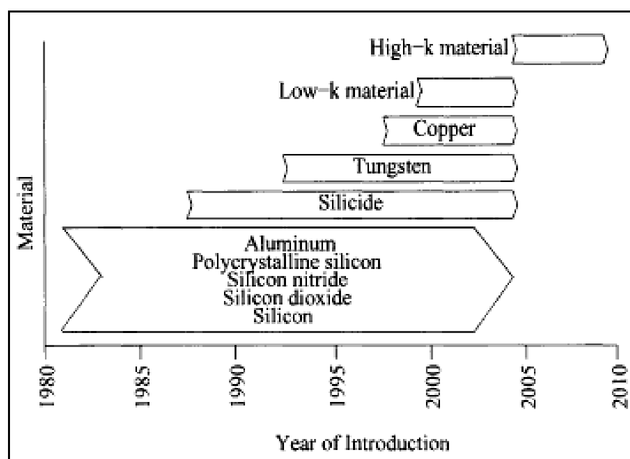

Fig. 7: Introduction of new material.

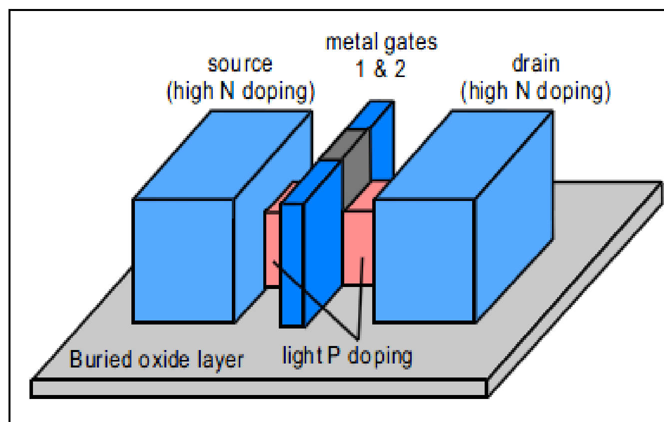

Fig. 8: Multiple gate or Fin-FET structure should be used to have a better control of the short-channel effects.

Economical Limitations:

The increasing cost in semiconductor industry is basically contributed by the cost of production and testing that growing exponentially with the time as the CMOS size is downsized. A new wafer foundry as shown in Figure 9 could cost approximately 25 million dollar today, and will rise by one-fold in 2010, as forecasted by The National Institute of Standards and Technology (NIST).The increment of the cost is also mainly caused by the equipment cost, clean room facilities, and lithography process difficulty (Wider and Neppi, 1992). The reduced size of circuit is sensitive to hard and soft faults. These defective-prone circuits require to be verified comprehensively with the purpose quality assurance. Moreover, complicated test technique will acquire extra testing steps and time as a result increasing the test cost.

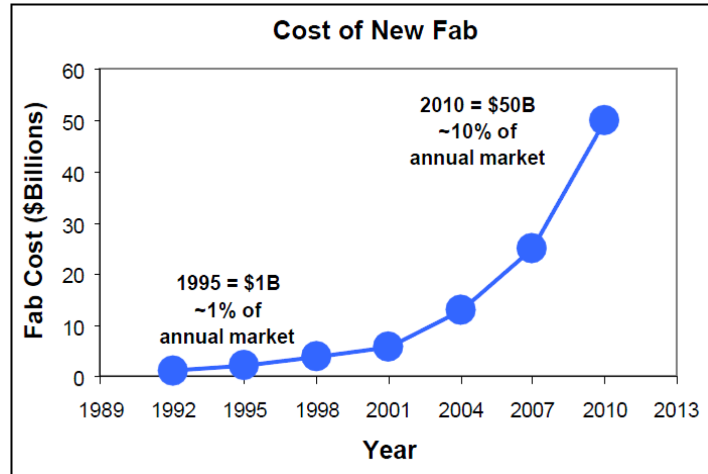

Fig. 9: Wafer foundry cost.

The cost minimum changes advance toward the lower right from year-to-year, as shown inFigure10. On the other hands, with continuous time and technology advancement, it is useful in order to increase the amount of components on an IC chip (Schwierz *et al*. 2010). This study is a basic significance since in semiconductor industry, cost is universal problem. On condition that the cost per component can be reduced by a realistic aspect when increasing the IC difficulty, the amount of devices per chip will be increase from one circuit production to the following production.

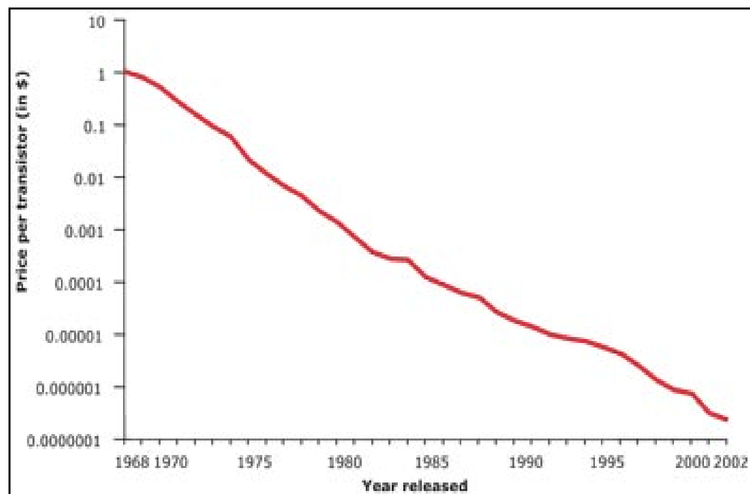

Fig. 10: Average transistor prices by year.

Technological Limitations:

CMOS transistors are essentially pattern on wafer by means of lithography and masks. It aims that the lithography technology is one of the main drives behind the transistor downsizing. On the other hand, the lithography procedures cannot manage with the decreasingcharacteristic of CMOS transistors. Lithography techniques such as proximity X-ray steppers and ion beam are limited by complexities in managing mask-wafer space and consistent exposure of photoresists on wafer. Additionalcrisis is the incapability of polishing procedure to sustain the consistent thickness of wafer and reliable mask according to(Gupta *et al*. 2003). Referring toSkotnicki *et al*., (2005), patterning lesser feature than wavelength of light needs trade-off amongdifficulty, pricey masks and potential design limit. Figure 11 shows the development of mask from the 180 nm technology to the recent technology.

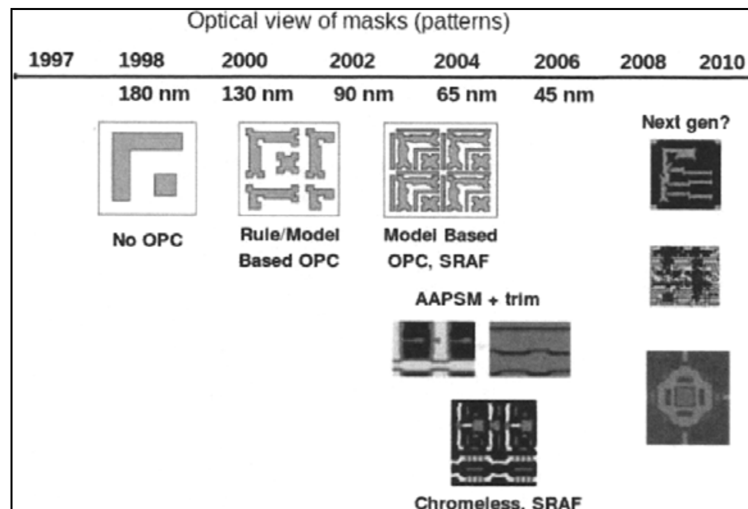

Fig. 11: Evolution of CMOS mask pattern.

Conclusion:

The evolution of CMOS beyond 22nm and toward 7nm faced many challenges and opportunities in design. The CMOS downsizing limitations are addressed in term of performance, material, power utilization, economical and technological limitations.The performance of the CMOS is estimated to increase together with the new material and technology development. Besides, high-*k* dielectric materials are being explored and could reduce the current leakage problem. It also reduces the power utilization with the help of variable clock frequency and variable voltage supply. Behind the transistor downsizing, the lithography technology is one of the key important processes. But then, it cannot handle with the CMOS downsizing due to the difficulty of the procedure that can be too costly. The technology advancement will focus on the manufacturing items concerning in producing the low-priced chip. Hence the progress of manufacturing will arise and technology revolution from the recent approach clean room, wafers and apparatus will carry on.

References

Akter, M., M.B.I. Reaz, F. Mohd-Yasin and F. Choong, 2008. Hardware implementations of an image compressor for mobile communications. Journal of Communications Technology and Electronics, 53(8): 899-910.

Akter, M., M.B.I. Reaz, Mohd-Yasin, F. and F. Choong, 2008. A modified-set partitioning in hierarchical trees algorithm for real-time image compression. Journal of Communications Technology and Electronics, 53 (6): 642-650.

Dennard, R.H., F.H. Gaensslen, L. Kuhn and H.N. Yu, 1972. Design of Micron MOS Switching Devices. In International Electron Devices Meeting, pp: 168-170.

Gupta, P. and A.B. Kahng, 2003. Manufacturing-aware physical design. In Proceeding of IEEE IACM International Conference on Computer-Aided Design, pp: 681-687.

Haron, N.Z. and S. Hamdioui, 2010. Why CMOS Scaling Coming to an END. In the Proceedings of 3rd International Design and Test Workshop (IDT), pp: 98-103.

Iwai, H. and H. Wong, 2006. Nano-CMOS Technology for Next Fifteen Years. In IEEE Region 10 Conference TENCON, pp:1-4.

Iwai, H. and S. Ohmi, 2002. Silicon Integrated Circuit Technology from Past to Future. Microelectronics Reliability, 42(4-5): 465-491.

Iwai, H., 2003. CMOS Scaling toward sub 10nm regime. In the Proceedings of 11th IEEE International Symposium on Electron Devices for Microwave and Optoelectronic Applications, pp: 30-34.

Iwai, H., 2007. Future of Nano CMOS Technology. In the Proceeding of 2007 International Workshop on Electron Devices and Semiconductor Technology (EDST), pp: 57-61..

Iwai, H., 2007. Future of Silicon Integrated Circuit Technology. In the Proceeding of 2nd International Conference on Industrial and Information Systems (ICIIS 2007), pp: 571-576..

Iwai, H., 2008. CMOS technology after reaching the scale limit.In the proceedings of the IEEE-EDS 8th International Workshop on Juction Technology, pp: 1-2.

Iwai, H., 2009. Roadmap for 22nm and beyond. Microelectronic Engineering, 86(7-9): 1520-1528.

Iwai, H., 2009. Technology Roadmap for 22nm and Beyond. In the proceedings of 2nd International Workshop on Electron Devives and Semiconductor Technology (IEDST'09), pp:1-4.
Iwai, H., 2012. Materials and structures for future nano CMOS. In the IEEE Nanotechnology Materials and Devices Conference (NMDC), pp: 14-18.
Kin, K. and G.S. Park, 2011. Landscape for Semiconductor Analysis: Issues and Challenges. In proceeding of the 18th IEE International Symposium on the Physical and Failure Analysis of Integrated Circuits (IPFA), pp: 1-9.
Kuhn, K.J., 2009. Moore's Law past 32nm: Future Challenges in Device Scaling. In the 13th International Workshop on Computational Electronics (IWCE'09), pp: 1-6.
Kuhn, K.J., 2011. CMOS Scaling for the 22nm Node and Beyond: Device Physics and Technology. In the Proceeding of International Symposium on VLSI Technology, Systems and Applications (VLSI-TSA), pp: 1-2.
Martin, M.F., 2011. High-k/Metal Gate Innovations Enabling Continued CMOS Scaling. In the Proceedings of ESCCIRC 2011, pp: 50-58.
Marufuzzaman, Mohd., M.B.I. Reaz, M.S. Rahman and M.A. Mohd. Ali, 2010. Hardware prototyping of an intelligent current dq PI controller for FOC PMSM drive. In Proceedings of 6th International Conference on Electrical and Computer Engineering, pp: 86-88.
Mogaki, S., M. Kamada, T. Yonekura, S. Okamoto, Y. Ohtaki and M.B.I. Reaz, 2007. Time-stamp service makes real-time gaming cheat-free. Proceedings of the 6th ACM SIGCOMM Workshop on Network and System Support for Games, NetGames '07, pp: 135-138.
Mohd-Yasin, F., A.L. Tan and M.B.I. Reaz, 2004. The FPGA prototyping of Iris recognition for biometric identification employing neural network. Proceedings of the International Conference on Microelectronics, ICM., pp: 458-461.
Moore, G.E., 1965. Cramming More Components onto Integrate Circuits. Electronics Magazine, 38(8).
Reaz, M.B.I. and L.S. Wei, 2004. Adaptive Linear Neural Network Filter for Fetal ECG Extraction. Proceedings of the International Conference on Intelligent Sensing and Information Processing, pp: 321-324.
Reaz, M.B.I., F. Mohd-Yasin, S.L. Tan, H.Y. Tan and M.I. Ibrahimy, 2005. Partial encryption of compressed images employing FPGA. In Proceedings of IEEE International Symposium on Circuits and Systems, pp: 2385-2388.
Reaz, M.B.I., F. Choong and F. Mohd-Yasin, 2006. VHDL modeling for classification of power quality disturbance employing wavelet transform, artificial neural network and fuzzy logic. Simulation, 82(12): 867-881.
Reaz, M.B.I., F. Choong, M.S. Sulaiman and F. Mohd-Yasin, 2007. Prototyping of wavelet transform, artificial neural network and fuzzy logic for power quality disturbance classifier. Electric Power Components and Systems, 35 (1): 1-17.
Reaz, M.B.I., M.I. Ibrahimy, F. Mohd-Yasin, C.S. Wei and M. Kamada, 2007. Single core hardware module to implement encryption in TECB mode.Informacije MIDEM, 37(3): 165-171.
Reaz, M.B.I., M.T. Islam, M.S. Sulaiman, M.A.M. Ali, H. Sarwar and S. Rafique, 2003. FPGA realization of multipurpose FIR filter. In Proceedings of Parallel and Distributed Computing, Applications and Technologies, pp: 912-915.
Schwierz, F., H. Wong and J.J. Liou, 2010. Nanometer CMOS. Pan Standford.
Skotnicki, T., Hutchby, J.A., T. King, H.-S.P. Wong, F. Boeuf, 2005. The End of CMOS Scaling: Towards the Introduction of New Materials and Structural Changes to Improve MOSFET Performance. IEEE Circuits and Devices Magazine, 21(1):16-26.
Wider, A.W. and F. Neppi, 1992. CMOS Technology Trends and Economics. IEEE Micro, 12(4):10-19.