

Design for reliability with the advanced integrated circuit (IC) technology: challenges and opportunities

Zhigang JI*, Haibao CHEN & Xiuyan LI

School of Microelectronics, Shanghai Jiaotong University, Shanghai 200240, China

Received 19 August 2019/Revised 26 August 2019/Accepted 4 September 2019/Published online 22 October 2019

Citation Ji Z G, Chen H B, Li X Y. Design for reliability with the advanced integrated circuit (IC) technology: challenges and opportunities. *Sci China Inf Sci*, 2019, 62(12): 226401, <https://doi.org/10.1007/s11432-019-2643-5>

Reliability assurance is of great importance for any commercial products, and the integrated circuit (IC) is no exception. Unlike the yield issues that are time-independent and can be screened through burn-in before taping out, the reliability issues introduce time-dependent aging in performance and eventually cause the malfunction of the ICs. Such a process can take several months or even years to happen and thus cannot be detected before entering the market. For many years, reliability was not a big concern, and design rule check during physical verification was usually sufficient. However, as technology migrates to advanced nodes, the reliability is acting as a showstopper and challenging the entire IC industry [1]. To gain a deep understanding of the emerging reliability issues, it is intuitive to explore the origin of such increasing threat.

(1) The complicated fabrication processes. This generally involves the fabrication of transistors with novel high-mobility channel materials, three-dimensional structures, ultra-thin dielectrics and substrates in front-end-of-line (FEOL) as well as the interconnect with thinner wires and the porous dielectrics in back-end-of-line (BEOL), which in turn results in higher leakage current, stronger self-heating, more noise, more electro-migration and other negative electrical aging issues. Moreover, the shrinking of the transistors into nanoscale imposes severe variability issues on top. All these reliability issues (aging and variability) pervade every aspect of IC by affecting transistor's per-

formance.

(2) The complicated operating modes. IC industry is driving into the era of dark silicon, in which only a small portion of transistors can be used at a time due to the power constraint. Computational sprinting and near-threshold computing are two common operating modes to gain higher performance within the thermal budget. Wherein, the former powers on a small amount of un-used transistors for a very short period at higher voltage/frequency, leading to aging issues. The latter turns on a larger amount of un-used transistors but at the voltage close to the threshold voltage, which is the domain sensitive to variability issues such as noise.

(3) The change in utilization workload. In the past, ICs sat idle most of the time. For example, the sleep-mode were widely used in consumer electronic products. The servers managed by each company have the average utilization efficiency to be only about 5% to 15%. However, more and more products are connected seamlessly through servers shared in the cloud in the ecosystem with the exchanging and the processing of information in an uninterrupted manner. It is anticipated that the utilization of ICs at both consumer end and the server end will be maximized, which will no doubt accelerate the IC aging and lead to more reliability problems.

Enhancing reliability through fabrication process optimization by the technologist in foundries is a straightforward solution to strengthen the

* Corresponding author (email: richard_zji@163.com)

long-term IC reliability. However, on one hand, this becomes more difficult and costly when we are approaching the physical limit. On the other hand, different types of ICs and the different application scenarios have different reliability requirements, making the conventional one-size-fits-all approach obsolete. Nowadays, the ability to take reliability metrics into consideration at the design level has been widely accepted as an effective solution: during IC design, simulation and analysis for the long-term reliability are added before circuit fabrication under the operating voltage suggested by the foundries [2, 3]. In principle, this would not only guarantee the IC reliability but also help IC designers optimizing reliability, performance, power trade-off and reduce time-to-market. A number of reliability simulation tools have been introduced by electronic design automation (EDA) tool providers. However, owing to the lack of industry-wide accepted reliability model, the practical adoption of such design-for-reliability methodology is still in its infant stage. The development of the physical-based model that can deliver accurate reliability projection in both transistor and interconnect level with high efficiency is urgently required.

Transistor reliability — the aging effect. The use of high-k dielectrics (HK) is one key element for future advanced technology. Both the generation process and the trapping/detrapping process during the negative/positive bias temperature instability (N/P BTI) and hot carrier degradation (HCD) will be enhanced by the higher electric field and the self-heating effect. The exploration into the following directions could be timely and important.

(1) Understanding different types of traps. The root for the BTI or HCD degradation is the traps within the dielectrics and at the interface [4, 5]. These traps can either be pre-existing after the fabrication or be generated from the aging. Separating these traps, understanding their properties and the temporal/voltage dependence can be critical. Most of the existing models, such as two-stage model, reaction-diffusion (R-D) model and composite model, were established by either assuming a single type of trap or separating different types of traps through fitting with certain empirical models. This is not sufficient for advanced nodes in which the reliability physics can vary due to the change of fabrication processes, the integration of new materials and the shift to new device structures. New characterization method for the three-dimensional structure is needed to separate each type of traps as well as their energy level, vertical/lateral spatial location. Based on these reliable

data, the proper understanding and modelling of traps can be possible.

(2) Modelling for mobility degradation. Although it is widely accepted that mobility degradation is important, there is a lack of testing data to set up accurate mobility degradation model due to difficulties in its measurements. The state-of-the-art circuit simulator still uses the empirical models developed in the 1990s, which induces uncertainties. Many questions remain to be answered. For example, how the different types of traps at different locations can affect mobility? What are the processes dependence on mobility degradation? How can mobility degradation be affected by device geometry? All these understanding can be critical for building accurate reliability models.

(3) Modelling for the interaction of different aging mechanisms. As the common practice, models were typically developed using the tests finely tuned to maximize the effect of one aging mechanism, which is over-simplified. In practice, multiple mechanisms can occur either simultaneously or sequentially, and they interact with each other. For example, BTI and HCD can occur sequentially in transistors in one ring oscillator. The total degradation is much smaller than adding the degradations induced by each mechanism alone. This can be understood by thinking about the traps: although occurring at different energy/spatial locations, both mechanisms are related to similar types of traps. Therefore, the increase of trap by one degradation mode will suppress the degradation by the following stress mode. Without considering such coupling effect, the degradation can be overestimated, causing an unnecessary compromise in IC design. With the decreasing number of traps in transistors in advanced technologies, such coupling effect will become stronger. The bottom-up methodology, i.e., modelling from the traps, could be one feasible solution to naturally take into account.

(4) The involvement of the self-heating effect. In advanced nodes, the higher and narrower fins in tri-gate or FinFET transistors are required to meet the higher driving current requirement. Such isolation with three-dimensional structure results in higher thermal resistance to the bulk that leads to higher local temperatures. This is the so-called local self-heating effect, which deteriorates the transistor reliability. Therefore, it must be understood and modelled. The key challenge for the study in our mind is how the proper projection can be made from the accelerated condition used in the test to the operating condition that is usually at high frequency, with different duty cycle. In addition, owing to the lack of proper characterization

method, it is still not clear about the relevance of the time-scale difference between electrical and thermal response.

Transistor reliability — the time-dependent variability effect. In advanced nodes, the smallest transistors are in nano-scale, and severe variability issues can occur. Although extreme ultraviolet lithography (EUV) can suppress the time-independent process variations, such as line edge roughness, the limited number of electrical-active traps in the dielectric introduce intolerable time-dependent variability (TDV). The trapping/detrapping process induces stochastic and abrupt changes in device properties, such as V_{th} and driving current. This, in turn, affects the ICs, such as the timing in the digital circuit leading to serious jitter issues and the temporal mismatch in the analog circuit leading to the accuracy loss. Owing to the channel percolation effect, the impact of a single trap is much larger than the classical charge sheet approximation. The V_{th} -shift, ΔV_{th} , caused by a single trap, has been observed to be over 70 mV on the 22 nm node. What makes things worse is that such impact will further increase proportionally with downscaling. For low-power applications in which the overdrive leaves less headroom to resist ΔV_{th} , a single trapped charge can effectively turn off the transistor. However, current SPICE models do not consider time-dependent variability, which leaves the simulation with uncertain error margins in the verification and signed off process. The availability of the proper simulation tool for the statistical variability projection will determine whether we can successfully continue the scaling in power, performance and area. The variability modelling is at the center of this task and has attracted great attention from both academia and industry in recent years [6, 7]. The trap-centric approach has been widely used in the community, in which most of the efforts have been put on the random telegraph noise (RTN) with the focus on the capture/emission properties of each individual trap. This involves laborious characterization and data analysis. Qualitative analysis with single-trap TDV has been demonstrated. However, the quantitative analysis on multi-trap TDV will be required for real-case project. Wherein, the accuracy and efficiency enhancement should be addressed properly.

(1) Extracting fundamental properties for the individual traps. Accurate modelling for the trapping/detrapping phenomenon requires fundamental properties for a trap including capture cross-section (CCS), spatial location, energy level, and thermal activation energy. Without them, it is impossible to develop quantitative models. However,

there are scarce data on them, and early works often contradict each other. Taking CCS as an example, some groups suggested that CCS has a continuous distribution, while others found only two discrete values.

(2) Proper characterization methodology. This usually involves the stress-and-sense procedure. For stress, many TDV studies tend to follow the aging methodology: raising gate bias, V_g , higher than the operating condition to accelerate the test. The higher V_g , however, can trigger the traps at a higher energy level that will be electrically inactive under real use-bias, so that the TDV can be over-estimated by including these “wrong” traps. For sensing, since threshold voltage shift, ΔV_{th} , plays a key role in the simulation flow, most of the research typically measured drain current fluctuation at stress level and then divided it by transconductance. However, such conversion can over-estimate ΔV_{th} by over 200%. This is due to the neglect of the gate bias dependence of a trap’s impact. Although some efforts have been made to understand this phenomenon, the experimental characterization and physical-based modelling are urgently needed.

(3) Understanding the coupling effects. Similar to the aging, the TDV and other local statistical variability introduced by the discreteness of charge and matter are also strongly coupled. Their interaction is currently lacking of research. The dominant sources of variability in advanced technology nodes differ, and the differentiation between bulk and FinFETs in terms of TDV is not well studied experimentally and understood in terms of modelling and simulations.

(4) High-sigma variability simulation. Modern ICs such as SoC is using an increasing number of transistors. Therefore, it is important to perform statistical variability simulation with high-sigma (i.e., 6 sigma), in which the reliability issue at the tail can be estimated. This, however, is time demanding or even impossible for the state-of-the-art Monto-Carlo approach. A new statistical approach is urgently needed, which requires collaboration between the technologies and the mathematicians.

Interconnect reliability — EM and TDDDB. The continuous scaling also requires the reduction of the wire width in the interconnect. Therefore, even with the relatively small current flow, such a small cross-sectional area can cause high current density. This can introduce the momentum transfer from electrons to the conductor atoms, leading to the conductor atoms migration and eventually creating voids in the conductor. Such electromigration process can be divided into (1) void nu-

cleation phase under the co-influence of electron current driving force and inverse stress before the formation of voids and (2) void growth phase under the influence of electron current driving force after the formation of voids. Reliable mathematical models still need to be established in the stage after the formation of voids. In addition, an effective mathematical model for the process after void formation will be helpful to analyze the physical mechanism for the failure in electromigration. The understanding can be used to optimize design research, and solve the programs of material selection, reliability evaluation and life prediction that are produced by the current design and evaluation procedures relying heavily on the high-cost and long-term experimental work. Traditional analytical methods of electromigration, such as Black equation and Blech theory, may lead to over-design. In order to ensure the timing accuracy of the chip, more guard bands are needed, resulting in inefficiency and huge losses in the area, performance, power consumption and reliability budget. These conservative design rules are not suitable for future technology scale. Therefore, in the design process, a balance must be found to ensure circuit performance and not seriously affect the reliability of electromigration. In order to achieve this balance, further study of the dominant factors of the physical dynamics of migration and the failure process is necessary.

For advanced technologies, the reduction of the interconnect coupling capacitance becomes critical for high-speed transmission within ICs. Therefore, the use of low-k, ultralow-k, or even porous dielectrics is expected¹⁾. However, these dielectrics can be fragile both electrically and mechanically, making them susceptible to the time-dependent dielectric breakdown (TDDB), which was only the concern in the transistor level. In fact, the international technology roadmap for semiconductor (ITRS) has identified both TDDB and electromigration as the key challenge for future interconnect. The research in the future should incorporate process variations into the TDDB assessment approach statistically, such as the interconnect line-to-line spacing variation, the interconnect edge roughness and the dielectric variation. Moreover, current studies assume that the interconnect work upon the constant temperature, which is impossible in the practical IC operating conditions. The practical temperature is load-dependent and influenced by the environment condition, which should also be considered.

Conclusion. The speed, computational power, and enhanced functionality of ICs based on the advanced technology promise to transform both our work and leisure environments. To guarantee the IC reliability, the design for reliability approach is critical, which requires accurate modelling for both the transistor and interconnect reliability. This involves the proper characterization and a deep understanding of the physical properties of traps. From the technologist's point of view, we briefly addressed the challenges and opportunities in the existing research activities. Finally, we anticipated that the advance in machine learning could accelerate this process since the reliability research generated a significant amount of data, which can be used for the training of predictive recursive neural network for the reliability projection and the convolutional neural network for the reliability analysis.

Acknowledgements This work was supported by National Natural Science Foundation of China (Grant No. 61604095) and Shanghai Natural Science Fund (Grant No. 19ZR1475300).

References

- 1 Auth C, Aliyarukunju A, Asoro M, et al. A 10nm high performance and low-power CMOS technology are featuring 3rd generation FinFET transistors, self-aligned quad patterning, contact over active gate and cobalt local interconnects. In: Proceedings of IEEE International Electron Devices Meeting (IEDM), San Francisco, 2017
- 2 Huang R, Jiang X B, Guo S F, et al. Variability-and reliability-aware design for 16/14nm and beyond technology. In: Proceedings of IEEE International Electron Devices Meeting (IEDM), San Francisco, 2017
- 3 Ji Z, Zhang J F, Lin L, et al. A test-proven as-grown-generation (A-G) model for predicting NBTI under use-bias. In: Proceedings of Symposium on VLSI Technology (VLSI Technology), Kyoto, 2015
- 4 Yu Z Q, Zhang J Y, Wang R S, et al. New insights into the hot carrier degradation (HCD) in FinFET: new observations, unified compact model, and impacts on circuit reliability. In: Proceedings of IEEE International Electron Devices Meeting (IEDM), San Francisco, 2017
- 5 Ji Z, Hatta S F W M, Zhang J F, et al. Negative bias temperature instability lifetime prediction: problems and solutions. In: Proceedings of IEEE International Electron Devices Meeting (IEDM), Washington, 2013
- 6 Wang R S, Guo S F, Zhang Z, et al. Too noisy at the bottom? Random telegraph noise (RTN) in advanced logic devices and circuits. In: Proceedings of IEEE International Electron Devices Meeting (IEDM), San Francisco, 2018
- 7 Ren P P, Wang R S, Ji Z G, et al. New insights into the design for end-of-life variability of NBTI in scaled high-k/metal-gate technology for the nano-reliability era. In: Proceedings of IEEE International Electron Devices Meeting, San Francisco, 2014

1) ITRS interconnect. 2015. <http://www.itrs2.net/itrs-reports.html>.