Design–Reliability Flow and Advanced Models Address IC-Reliability Issues

Mohamed Selim, Eric Jeandeau, Cyril Descleves
Mentor Graphics

***Abstract*—**

Reliability effects are real threats with advanced process nodes. This paper describes the industrial challenges associated with accurate modeling of aging problems. Joint design–reliability flows can mitigate their effects, and allow designers to take those effects into account as early as possible.

I. INTRODUCTION

Advanced short-geometry CMOS processes are subject to aging that causes major reliability issues, degrading the performance of integrated circuits over time. Degradation effects causing aging are hot carrier injection (HCI) and negative bias temperature instability (NBTI), in addition to positive bias temperature instability (PBTI) and time dependent dielectric breakdown (TDDB). Below 90 nm, consideration of these effects is becoming mandatory for design flows targeting quality and reliability. This paper describes the state-of-the-art simulation flow that can help designers address these issues and to create more reliable designs.

These particular reliability effects modify the fundamental behavior of the transistors, such as threshold voltage (Vth) and the mobility factor [1]. No applications that make full usage of the process performance are really safe. These changes will affect timing delays, drive currents, leakage, linearity, and every possible specification that may appear in IC design, be it for automotive, biomedical, military-aerospace, wireless communications, or video. Basically, all industry sectors are potentially affected.

Device failure phases are described as:

a) Infant Mortality (early rate)

b) Normal Operating Life (constant random failures, intrinsic rate)

c) Wear Out

It would be obvious that for new technology nodes the wear out failure curve shifts to the left, which means earlier device degradation as shown in **Figure 1** [2]. The actual stress effect depends on the design itself. The manufacturing process can be refined so that these effects are globally minimized. However degradation is tightly linked to the design itself, the architecture, the chosen device geometries and, most importantly, the actual stimuli applied to the circuit during operation will strongly determine the magnitude and speed of degradation. Fortunately, tools are now available that help designers to understand exactly when and how much devices are stressed and damaged, and how this aging impacts the device or circuit performance after a certain time, allowing designers to take corrective actions if necessary. The necessity of linking classical design flows with the capability to predict reliability has been recognized by the industry [3, 4, 5].

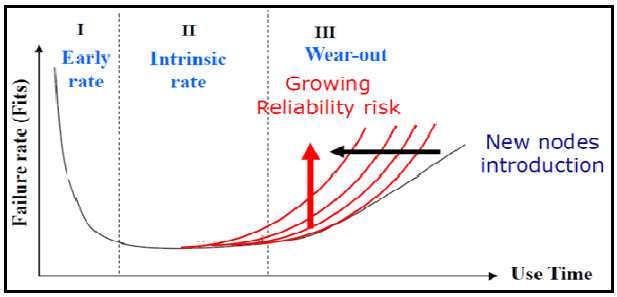

Figure 1 Reliability bathtub and the effect of new node introduction.

II. RELIABILITY SIMULATION

Each device experiences a different stress that depends on its exact individual bias conditions, in addition to global conditions such as temperature. This stress is computed individually and integrated over a particular time window [6]. This time window is chosen to be compatible with the CPU time constraints for circuit simulation. By extrapolation, an estimate of the stress seen by each device during a much longer periodic operation (maybe weeks, months, or years) is computed.

This stress quantity is then used to compute degraded values of the model parameters (such as the threshold, mobility, etc.). Using these degraded model parameters, a new simulation is run which represents what would happen after *N* years of operation. In this aged simulation, each device uses its own individual degraded model with updated Vth, mobility etc., because the stress is device-specific and

Workshop on Early Reliability Modeling for Aging and Variability in Silicon Systems – March 18th 2016 – Co-Located with DATE 2016 - Dresden, Germany

so is the update. Fresh and aged simulation results can be overlapped and compared.

Models can either take into account DC stress or dynamic stress, including the recovery effect [4]. Including the recovery effect leads to less pessimistic estimations of the degradation compared to a static DC stress approach hence avoiding over-design to compensate the degradation. Although rather sophisticated, this entire process is now completely integrated in regular SPICE flows (**Figure 2**). From the designer's perspective, performing an aging analysis of a cell or circuit is just a new command in the simulator, and just as simple to use.

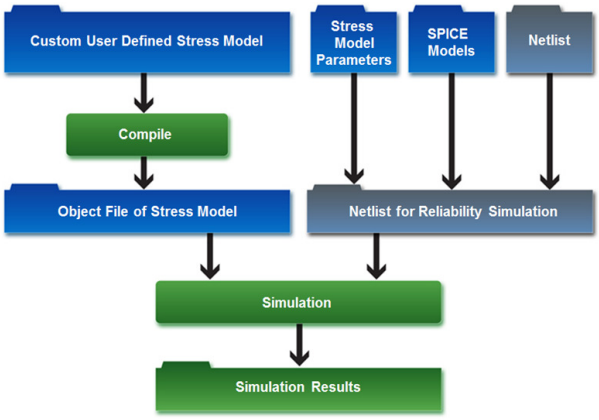

Figure 2: A joint design and reliability simulation flow

III. AGING AND RELIABILITY SIMULATION IN ELDO

The computation of the stress and the update of the electrical model parameters is done with user-defined equations. These two sets of equations are called a stress model and an update model. They can have virtually any complexity, from the simplest two-lines-of-code model to a pages-and-pages model. Examples of realistic models can be shown in [7, 8, 9]. The models to be used are described using a C API to communicate with the simulator kernel, allowing maximum computing efficiency. As explained, a typical aging analysis requires two simulations runs. The first run uses the nominal (fresh) models to produce the waveforms, but it is also in charge of computing the stress quantities. The second run uses the individually degraded models, and the CPU cost is exactly the same as a nominal run. Thus, the total aging analysis is at least 2x more costly than a nominal simulation. Depending on the complexity of the stress model, the total combined CPU time is usually comprised between 2x and 2.5x the time of a nominal simulation. Therefore, the cost in terms of CPU time is very modest, given the importance of the provided insight—especially compared to Monte Carlo analysis or simple corner case analysis where the ratio is easily in the range of 100X.

To further complicate an already difficult subject, aging appears to be a statistical process unto itself. If you consider two identical devices drawn side-by-side, they will have slightly different (fresh) threshold voltages, drive current, and leakage currents. The distribution of these parameters is monitored by the foundries, allowing statistical simulations by designers. But, unfortunately, the same is true for their aging characteristics. Two identical devices with identical bias conditions do not age exactly in the same way. The Vth shift, for example, is a statistical variable with its own mean, variance, etc. This spread of the aged transistor parameters also will be reflected onto the measured performances such as a propagation time (**see Figure 3**) and distortion level. However, the required measurements needed to capture the statistical nature of the aging process are much more complicated, lengthy, and costly than regular process monitoring.

Solid information about the statistical properties of the aging process is rarely available from the foundries however Eldo easily support statistical analysis on top of the aging flow. For example, it is possible to assign a certain variance to any of the parameters that define the aging models and run a Monte Carlo simulation to get a statistical view of the aging trajectories.

Reliability models are implemented in the Eldo UDRM (User Defined Reliability Model) interface [10]. The modeled damage could follow one or more of several damage mechanisms, which show gradual degradation in device performance. Other mechanisms, which cause sudden and complete damage of the device, (like the abrupt oxide breakdown for example) are not targeted in this scope.

Reliability changes can be analyzed for any degradation effect: HCI, NBTI, PBTI and TDDB.

In a Long-Term Reliability Simulation Scheme, there are two possible methods: Two simulations scheme & Repetitive scheme

In the Two simulations scheme the steps are as follows:

1. A transient simulation with the "fresh" device is done.
2. At the end of the fresh simulation, the amount of damage each device has been subjected to, caused by the stress applied on the device, is calculated.
3. The transistor models are updated accordingly using the equations specified in the user defined reliability functions.
4. A new transient simulation is run with the aged device.

In the Repetitive scheme, the long period Tage is divided into smaller time intervals Ti (where Tage=ΣTi). The same steps as the two simulations scheme are followed, except that steps 2 to 4 are repeated NBRUN times, where NBRUN is the number of time intervals. The calculation of the stress is updated at the end of every time interval. This process is repeated until t=Tage. This approach can account for the gradual changing bias conditions as a result of device degradation. It is obvious that if the number of time divisions was chosen to be equal to one, the Repetitive scheme will be identical to the two simulations scheme. The more general case, the Repetitive scheme, is implemented in Eldo to ensure a high level of accuracy, and account for the

gradual changing bias conditions as a result of device degradation which could significantly affect the results. This flow is described in **Figure 4**.

Up to recently, aging mechanism has mainly focused on active devices. However, new challenges have appeared in industrial and automotive power applications where resistors and capacitor degradations are becoming critical issues. For example the need for accounting accurate resistors instability modeling has become mandatory in latest high voltage processes. Eldo has been addressing those new challenges while targeting any kind of design regions and not only blocks which contain transistors. Furthermore, as Eldo can run electro-thermal simulation using an electrical netlist connected to a thermal netlist (built using RC network), one can mix those 2 capability to apply aging on the thermal RC components which are part of the simulated thermal network. Without this capability, an electro-thermal simulation considers that the RC thermal network does NOT degrade over time. Eldo can help to remove this limitation in applying degradation models for both the electrical and thermal netlists..

IV. JOINT DESIGN–RELIABILITY FLOW

Using a joint design-reliability flow allows the designer to predict the behavior of the circuit versus "wall-clock" time. Important metrics can be traced versus time and verified against the specifications. For example, **Figure 5** shows how the operating frequency of a CMOS oscillator degrades over time. The absolute period and the relative degradation (in %) of the frequency are shown in the upper and lower plots, respectively. The *x*-axis is the time in years in logarithmic scale. The frequency is degraded by nearly 5% after only one year of operation. The rest of the circuit may be able to accommodate such degradation, or not.

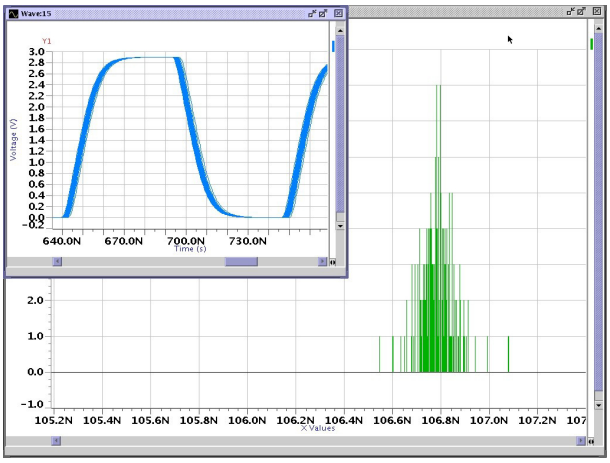

Figure 3: The spread of aged transistor parameters are reflected onto measured performances such as a propagation time.

Predicting the degradation of a key metric (here it is an oscillation frequency, but it could be power consumption, a distortion level, or a settling time) is interesting. However, it is a compound result that depends on many variables. The next thing a designer wants to know is which device is primarily responsible for the degradation of the observed metric. Joint design-reliability flows allow identification of devices that are subject to the largest degradation. The information is typically presented in tabular format with sorting criteria. For example, a table may show the relative degradation of the drive current, the linear current, and also the trans-conductance (gm), Vth, or generally any quantity of interest. The designer can choose the sorting criterion. In the example shown in **Figure 6**, the results are sorted by decreasing "delta-Vth," where the devices that have their threshold voltage degraded the most severely are presented first. This allows the designer to immediately identify the areas in the circuit that require extra attention.

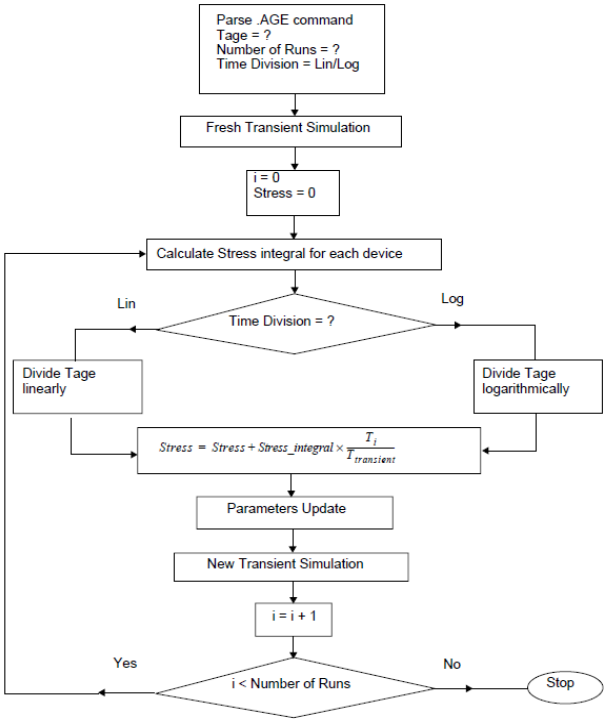

Figure 4 Reliability simulation repetitive scheme flowchart.

IP protection is one of the main benefits of this flow. For many companies or foundries, the details of the equations and models used to predict degradation are not considered public information. Rather, they consider it to be sensitive proprietary information that they are not keen to disclose in any way. For this purpose, Eldo has developed encryption mechanisms that allow full protection of the information. They can run the simulation using only binary non-human–readable model files. As well, security encryption keys can be used to restrict access and control the execution of the models by the simulators. Once encrypted and protected, only licensed partners, customers or sub-contractors, such as design houses, can make use of the protected libraries.

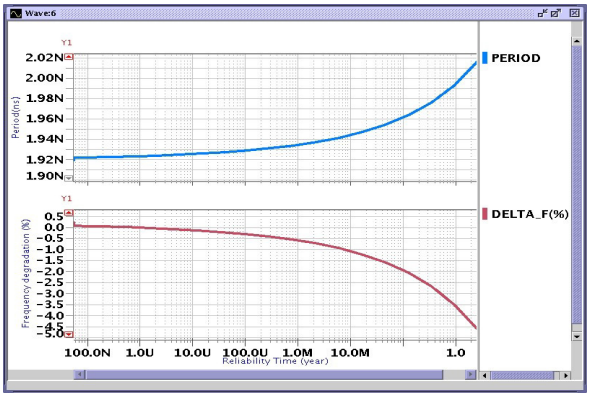

Figure 5: The operating frequency of a CMOS oscillator degrades over time.

V. REAL DESIGN AGING USING ELDO UDRM

A complete use model of the flow starting from stress and update model development, implementation in Eldo UDRM interface and performing aging analysis on a design and showing how the stress effects was crucial for the design is described in [7]. They worked on a zero crossing detector comparator (ZCDC) which is an essential part of Smart Power ICs driving inductive load. Its function is to avoid the inversion of output current. The design included 2 MOS devices who operated on different stress, the switching of both transistor is monitored for the operation of the ZCDC hence accurately modeling Vth of both transistors and the shift occurring in both M1 and M2 behavior due to stress is mandatory to have accurate simulation results.

```
DELTA measurement summary :
---------------------------

DIDSAT      DELTA(ID8(*),VGS=3,VDS=3,VBS=0)
DIDLIN      DELTA(ID8(*),VGS=3,VDS=0.1,VBS=0)
DGM         DELTA(GM(*),VGS=1,VDS=3,VBS=0)
DVT         DELTA(VT(*),VGS=3,VDS=0.1,VBS=0)

      | DIDSAT                | DIDLIN                | DGM                   | DVT                   |
------|-----------------------|-----------------------|-----------------------|-----------------------|
X5.M1 | -3.9511U    (3.1%)    | -416.9088N   (2.9%)   | -1.7672U    (6.0%)    |  25.3177M    (4.9%)   |
X6.M1 | -2.5325U    (2.0%)    | -288.5018N   (2.0%)   | -1.0701U    (3.6%)    |  13.0694M    (2.5%)   |
X7.M1 | -1.3389U    (1.1%)    | -165.9644N   (1.2%)   | -528.4427N   (1.8%)   |  4.9516M     (1.0%)   |
X8.M1 | -1.2010U    (1.0%)    | -150.5232N   (1.1%)   | -469.5469N   (1.6%)   |  4.2034M     (0.8%)   |
X4.M1 | -1.1973U    (1.0%)    | -150.0971N   (1.1%)   | -467.9514N   (1.6%)   |  4.1836M     (0.8%)   |
X21.M1| -1.1964U    (1.0%)    | -150.0016N   (1.1%)   | -467.5942N   (1.6%)   |  4.1792M     (0.8%)   |
X11.M1| -1.1865U    (0.9%)    | -148.8792N   (1.0%)   | -463.4004N   (1.6%)   |  4.1274M     (0.8%)   |
X15.M1| -1.1820U    (0.9%)    | -148.3658N   (1.0%)   | -461.4859N   (1.6%)   |  4.1038M     (0.8%)   |
X14.M1| -1.1817U    (0.9%)    | -148.3297N   (1.0%)   | -461.3511N   (1.6%)   |  4.1021M     (0.8%)   |
X17.M1| -1.1812U    (0.9%)    | -148.2718N   (1.0%)   | -461.1356N   (1.6%)   |  4.0994M     (0.8%)   |
X12.M1| -1.1783U    (0.9%)    | -147.9448N   (1.0%)   | -459.9175N   (1.6%)   |  4.0845M     (0.8%)   |
X9.M1 | -1.1772U    (0.9%)    | -147.8230N   (1.0%)   | -459.4640N   (1.6%)   |  4.0789M     (0.8%)   |
X2.M1 | -1.1760U    (0.9%)    | -147.6825N   (1.0%)   | -458.9412N   (1.6%)   |  4.0725M     (0.8%)   |
X10.M1| -1.1754U    (0.9%)    | -147.6240N   (1.0%)   | -458.7236N   (1.6%)   |  4.0698M     (0.8%)   |
X13.M1| -1.1676U    (0.9%)    | -146.7310N   (1.0%)   | -455.4048N   (1.5%)   |  4.0291M     (0.8%)   |
X19.M1| -1.1674U    (0.9%)    | -146.7046N   (1.0%)   | -455.3067N   (1.5%)   |  4.0279M     (0.8%)   |
X3.M1 | -1.0730U    (0.9%)    | -135.8913N   (1.0%)   | -415.6760N   (1.4%)   |  3.5515M     (0.7%)   |
```

Figure 6: The results are sorted by decreasing "delta-Vth." The devices that have their threshold voltage degraded the most severely are presented first.

M1 drain voltage switches from VDD to 0 following a 1 MHz square waveform while M2 is stressed in DC conditions. M1 is stressed for half a period while in the other half it is off and can partially recover the degradation. Therefore after reliability simulation, M2 Vth will become more negative than M1 one and comparator will switch at a positive drain voltage given by Vth shifts difference in the input stage MOS. This effect, which can't be predicted without the modelling of NBTI recovery, can generate unacceptable reverse currents causing the failure of circuit operation as observed in a real case. This has been illustrated in [7] using aging simulations and giving results in **Figure 7** which should be used by the designer to mitigate the aging risk on his design's behavior.

VI. CONCLUSIONS

Reliability effects are real threats with advanced process nodes, joint design–reliability flows, however, can mitigate their effects and allow designers to take those effects into account as early as possible.

Eldo provides a fully customizable and robust aging simulation interface that allows IP protection where details of the equations and models used to predict degradation can be encrypted to secure the IP.

This solution can be used with any type of analysis: AC, DC, Transient, RF, statistics, sensitivity and mixed signal simulations.

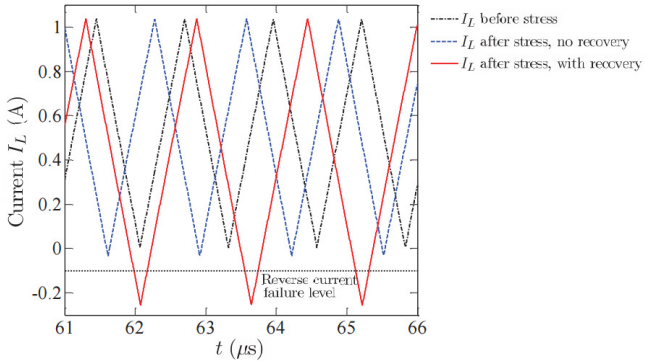

Figure 7. 6 Simulated current waveforms before (black) and after (blue and red) reliability trial (1000h, 175°C). Due to mismatch in Vth degradation of M1 and M2 an offset in ZCDC is generated resulting in a too large reverse current (red curve). The failure condition can't be predicted without recovery modeling (blue curve). [7]

REFERENCES

[1] Xiaojun Li, Jin Qin, and Joseph B. Bernstein, "Compact Modeling of MOSFET Wearout Mechanisms for Circuit-Reliability Simulation," IEEE Transactions on Device and Materials Reliability, Vol. 8, No. 1, March 2008..

[2] Alain Braviax, IIRW tutorial 2010

[3] Wenping Wang, Vijay Reddy, Anand T. Krishnan, Rakesh Vattikonda, Srikanth Krishnan, and Yu Cao, ―An Integrated Modeling Paradigm of Circuit Reliability for 65nm CMOS Technology,‖ IEEE 2007 Custom Integrated Circuits Conference (CICC).

[4] S. Bhardwaj, W. Wang, R. Vattikonda, Y. Cao, and S. Vrudhula, ―Predictive Modeling of the NBTI Effect for Reliable Design,‖ Proceedings of the IEEE Custom Integrated Circuits Conference, pp. 189–192, September 2006.

[5] Mridul Agarwal, Varsha Balakrishnan, Anshuman Bhuyan, Kyunglok Kim, Bipul C. Paul, Wenping Wang, Bo Yang, Yu Cao, and Subhasish Mitra, ―Optimized Circuit Failure Prediction for Aging: Practicality and Promise,‖ IEEE 2008 International Test Conference.

[6] Mikido Sode, et al. ―Reliability Simulation Environment Tackles LSI Design,‖ Chip Design, June 2007.

[7] Alagi, Filippo, et al. "Compact model for parametric instability under arbitrary stress waveform." 2014 44th European Solid State Device Research Conference (ESSDERC). 2014.

[8] Huard, Vincent, et al. "CMOS device design-in reliability approach in advanced nodes." IEEE intl' reliability physics symposium. 2009.

[9] Alagi, Filippo, Roberto Stella, and Emanuele Viganó. "Aging model for a 40 V Nch MOS, based on an innovative approach." IETE Journal of Research 58.3 (2012): 191-196.

[10] Eldo UDRM manual.