IRPS Workshop: Self-heating a Reliability Issue?

Moderators: M. A. Alam (Purdue),C. Prasad (Intel), A. Kerber (Global Foundries)

1. Background.

The physics is clear: In a 1961 paper, Rolf Landauer explained that "information is physical," and thereby dashed any hope of calculating without spending energy. And when the power is dissipated deep within a stack consisting of front-end, back-end, and packaging materials, self-heating is inevitable. Therefore, no-one doubts that self-heating exists in every electronic device, the only question is: Is self-heating significant enough to matter?

Historically, it did matter. In fact, self-heating catalyzed the three revolutions in electronics based on vacuum tubes, bipolar transistors, and MOSFETs. The transitions occurred only when self-heating made scaling or integration untenable. In an oft-repeated story, CRAY-1 was delayed not because of the electronics, but due to a persistent Freon leak. Without the Freon cooling, the computer would not work. After years of relatively benign increase, self-heating is beginning to force cellphone designers to adopt adaptive voltage control to keep heating at bay. Cooling has emerged an important design consideration for Virtual/Augmented reality headsets. Microsoft in proposing to submerge data-centers in ocean waters to keep them cool. Some universities are justifying their supercomputers based on the hope that it will reduce the dormitory heating bill in the winter!

Therefore, no-one doubts that self-heating is important for modern computer technologies. Instead, the discussion is focused on a narrower problem: The role of excess self-heating when transistors evolve from planar to the bulk-FinFET to gate-all-around NW transistors. Some argue that the thermal resistance associated with backend and the packaging is already so large that the additional thermal resistance due to topology change does not matter. Others say that things are already hot and the transistor topology change is significant enough to cause problems. The discussion is complicated by the fact that different groups may use the same transistor technology, but the test-structures, operating conditions, applications, and use profiles differ so substantially that it is difficult to reach a consensus.

With this background, approximately 70 participants in the IRPS self-heating platform exchanged their views, shared their opinions, and responded to a google survey. The panelists framed the discussion with a few questions: (1) Is self-heating real and can it observed? (2) So what if it heats? Does it matter to you? (3) If it matters, how do you deal with it? (4) Are there important second order effects, such as persistence or duty-cycle effects that our current models cannot capture? The discussion was animated and highly informative, see below.

2. Discussion during the Workshop

2.1 Does Self-heating matter?

Majority of the survey responded (~80%) felt that self-heating is important, *but only for certain applications.* The discussion showed that people find self-heating important in two different contexts: system performance and reliability qualification.

From a system perspective, self-heating is an important issue for cell-phones, virtual reality handsets, and data-centers. One of the participants explained how Joule heating in PCB boards may raise the temperature of the ICs. A redesign of the PCB interconnects may offer significant improvement. Others emphasized that hot-spots in multi-chips must be managed with integrated sensors. The culprit is dense packing of blocks and an inability to adopt active cooling. Here, power management integrated circuit need to play an important role. Without it, participants from the car-industry recounted how self-heating has resulted in dielectric cracking.

The second issue of reliability qualification, the discussion focused on HCI degradation (apparently very high) and the challenge of projecting to use conditions. Here, temperature increase as high as 70C has been reported. Some participants were concerned that even if a product is qualified for a given test condition, the self-heating situation may become worse when systems are stacked in data-center. In other words, actual system use conditions may either improve and acerbate reliability challenges.

2.2 **If Self-heating is important, what are the key contributors?**

When the panelists asked about the key causes of self-heating, most workshop participants identified an unfortunate confluence of multiple factors: confined transistor geometry, interfacial resistance between different materials, dense packaging and difficulty of integrating heat sinks. One of the conference papers noted that "On advanced technology nodes, increases in power density and phonon confinement, non-planar architectures and different material systems can exacerbate local self-heating due to active power dissipation, which can affect device performance and reliability in various ways." Most participants agreed with the assessment and implicated the thermal conductivity and the geometry of the channel being the most important contributor to self-heating. If true, device engineers will be held primarily responsible for self-heating and the integration of lower-thermal conductivity materials, such as III-V and Ge, may be an issue.

Others however suggested that the additional resistance due to channel geometry/materials, however large, contributes only a small fraction of the total thermal resistance. While it may appear to be very important in isolated test structures, the differential increase of self-heating due to transistor geometry would be negligibly small in the system context. After all, if the community knows how to design with SOI circuits without self-heating being a show-stopper, it cannot be a design issue for bulk FinFET. Despite this difference in opinion for Si-based bulk FinFET, almost everyone agreed that surround-gate transistors (i.e. NW-transistors) based on Ge and III-V channel material could be worse, but one should wait for more data before we jump to conclusions.

2.3 How should one measure self-heating?

Almost half the survey respondents reported using a sophisticated combination of self-heating characterization techniques such as AC conductance (or other RF methods), temperature dependence resistance techniques, pulsed I-V (or other ultra-fast technique). If a group had to choose one technique, they often resorted to resistance-sensors integrated with the gate or placed near a hot-spot. The discussion showed that while there are some confusion regarding the accuracy of the techniques, the key concern is the use of representative test condition) that will quantify self-heating close to actual use condition.

3) How should one project the test-results to use conditions?

The projection to use conditions was a topic of lively animated discussion. How does self-heating scale at high-frequency? Does the self-heating depend on duty-cycle? How do obtain a AC/DC ratio? What is the relevance of the time-scale difference between electrical vs. thermal response (5 ps vs. 10ns)? In other words, do we need to worry about persistence of the self-heating long after the transistor is turned off? Why even use a DC tester -- shouldn't one use a ring-oscillators? How do you correct for HCI lifetime at the use condition -- wouldn't it involve nonlinear projection based on self-heating? If SH reduces lifetime so dramatically, can burn-in erode reliability-margin considerably? Should we worry about core transistors, or only I/Os? Would self-heating affect on-state TDDB and do we need a new reliability model for TDDB lifetime? Some of these questions have been debated for decades (e.g. AC/DC ratio, use of DC vs. RG measurement), while others are new (e.g. nonlinear projection). Various participants from industry and academia explained what they do, but it became clear by the end that finding a generally acceptable and physically justified projection model should be one of the most important goals for the community.

4) What is the best way of modeling self-heating?

Regarding the modeling of self-heating and extrapolation to use condition, there were several surprises. More than half the respondents reported using highly sophisticated, self-consistent "correct-by-design" (proactive) techniques. Others treat self-heating as a run-time variability issue and allocate a guard-band to ensure design robustness. Not everyone uses modeling or guard-band based approaches however: some feel that their applications is not susceptible to self-heating effect and a self-heating-aware design be unnecessarily pessimistic.

To model the effect of self-heating at the design phase as well as reliability qualification, most favored a distributed thermal compact model, with a few suggesting that a single element RC model is actually sufficient. A few participants however cautioned against using these simplified models, because the nonlinearity of the thermal conductivity as a function of temperature and film thicknesses may introduce significant errors. This is particularly true for high-k dielectrics. Despite the computational cost, therefore, the community should invest in solution of the phonon Boltzmann equation, supported by 3D finite element modeling.

5) What is the most important reliability issue in a self-heated IC?

When asked about the reliability mechanism most affected by self-heating, NBTI, HCI, and EM floated to the top of the list. An animated discussion followed: Some suggested using EM as the "canary" for self-heating: Given its strong temperature sensitivity ($E_A \sim$ 1eV), qualification of EM should ensure that the rest of reliability issues pose no problem. Others objected: It is not clear if EM is the most important reliability issue at the use conditions, therefore focusing exclusively on EM-qualification as a proxy for all reliability issues may be dangerous.

5) At what level of the hierarchy, is the self-heating problem addressed most effectively?

The BTI/HCI/EM reliability must be modeled and qualified at the device/cell-template level. Based on initial test results, there may be an opportunity to redesign/optimize the devices to meet the targets for reliability and performance. Once the device design is in place, SOC design must manage hot-spots and usage-specific aging. The circuit level compensation has shown to be highly effective in this regard. Finally, at the system level, where one may stack CPUs in close proximity, one need to actively compensate for aging and heating by throttling down cores as needed. In particular, the challenge mutual heating, where a relatively modest circuitry suffer because highly-active neighboring circuit, can only be handled for system level analysis. The workshop participants felt that self-heating is best managed when these hierarchical mitigation strategies are pursued in tandem.

3. Conclusions and recommendations

Towards the end of the workshop, the discussion boiled down to the following conclusions:

a) The physics and general importance of self-heating are well understood. For silicon-based logic and memory design, many groups have sophisticated modeling and characterization capabilities in place. The key challenge is to quantify self-heating for application-specific use conditions and define its implications for NBTI, HCI, and EM.
b) Self-heating correlates and conflates various degradation modes, especially in high-frequency testers. The community needs to develop tools to deconvolve these effects before they can be projected to operating conditions.
c) The traditional HCI lifetime projection model must be updated. The community urgently needs a new lifetime projection methodology that accounts for frequency, voltage, temperature, and self-heating effects self-consistently and on equal footing.
d) Demonstrating self-heating impact on digital circuit degradation, such as ring-oscillators, would be beneficial in determining discrete device degradation to circuit

aging correlations related to BTI/HCI and in identifying potential gaps due to self-heating.

e) The community needs to characterize/report self-heating in new channel materials (Ge and III-V) and new systems (e.g. Google Home, Amazon Echo) to reach a consensus regarding the self-heating challenges of these materials and systems.
f) A holistic, system-level perspective to managing self-heating is essential for variety of systems. The reliability community must appreciate the system perspective to avoid any unanticipated failures of an otherwise conservative design.

Perhaps we will see progress towards these goals in the next year's IRPS.

References.

[1] Rolf Landauer, "Irreversibility and heat generation in the computing process" IBM Journal of Research and Development, 5 (3): 183–191, 1961. Also see Landauer, "The physical nature of information." *Physics letters A* 217.4-5 (1996): 188-193.

[2] S. Mittl and F. Guarín. "Self-heating and its implications on hot carrier reliability evaluations." *Reliability Physics Symposium (IRPS), 2015 IEEE International*. IEEE, 2015.

[3] M. A. Wahab, S. Shin, & M. A. Alam, "3D Modeling of Spatio-temporal Heat-transport in III-V Gate-all-around Transistors Allows Accurate Estimation and Optimization of Nanowire Temperature," *IEEE Transactions on Electron Devices*, *62*(11), 3595-3604, 2015.

[4] S. Rikhi, "Leading at the edge of Moore's law with Intel Custom Foundry", Intel Developer Forum, 2014

[5] Haldun Küflüoğlu, Min Chen, Shan Lu, Ashwin Rabindranath, Reza Kakoee, Shih-Hsin Hu, "Thermally-aware Sensor Allocation for Real-time Monitoring and Mitigation of FEOL Aging in System-on-Chip (SoC) Applications," Proc. of IRPS, 2017.

[6] "Characterization of Self-Heating and Current-Carrying Capacity of Intercalation Doped Graphene-Nanoribbon Interconnects," J. Jiang, J. Kang, and K. Banerjee, Proc. of IRPS, 2017.

[7] V. Sukharev, A. Kteyan, J.-H. Choy, S. Chatterjee and F. Najm, "Theoretical Predictions of EM-induced Degradation Test-Structures and On-Chip Power Grids with Analytical and Numerical Analysis," Proc. of IRPS 2017.

[8] K. Lee et al., "Effect of Joule Heating on Electromigration in Dual-Damascene Copper Low-K Interconnects," IRPS Proceedings, 2017.

[9] C. Prasad, S. Ramey, and L. Jiang, "Self-Heating in Advanced CMOS Technologies," IRPS Proceeding, 2017.

[10] E. Bury et al. "Statistical Assessment of the full VG/VD Degradation Space Using Dedicated Device Arrays," IRPS Proceeding, 2017.

[11] N. Lajo, E. Armagan, F. Chain, and S. Padiyar, "Test Method for Moisture-Driven Failure: Effect of VLSI Device Self-Heating and Local Temperature Rise," IRPS Proceeding, 2017.

[12] M. Alam, "The Emerging Challenge of Self-Heating in Surround Gate Devices: A Material, Device, and System Perspective," IRPS Tutorial, 2017