On the Thermal Failure in Nanoscale Devices: Insight towards Heat Transport Including Critical BEOL and Design Guidelines for Robust Thermal Management & EOS/ESD Reliability

Mayank Shrivastava[1], Manish Agrawal[2], Jasmin Aghassi[3], Harald Gossner[3], Wolfgang Molzer[3], Thomas Schulz[3], V. Ramgopal Rao[2]

[1]Intel Mobile Communications, East Fishkill, USA, email: mayank.shrivastava@intel.com; shrivastva.mayank@gmail.com

[2]Center for Nanoelectronics, Department of Electrical Engineering, Indian Institute of Technology-Bombay Mumbai-400076, India, email: rrao@ee.iitb.ac.in

[3]Intel Mobile Communications, Munich, Germany, email: thomas.schulz@intel.com; harald.gossner@intel.com

Abstract— **For the first time we have reported thermal failure of FinFET devices related to fin thickness mismatch, under the normal operating condition. Pre and post failure characteristics are investigated. Furthermore, a detailed physical insight towards heat transport in a complex back-end of line (BEOL) of a logic circuit network is given for FinFET and extreme thin silicon on insulator (ETSOI) devices. Self heating behavior of both the FinFET and ETSOI devices is compared. Moreover, layout, device and technology design guidelines (based on complex 3D TCAD) are given for robust thermal management and electrical overstress / electrostatic discharge (EOS/ESD) reliability.**

Keywords- BEOL Reliability, FinFET, ETSOI, Electrothermal, ESD .

I. Introduction

A number of non-planar SOI devices such as FinFETs and nano-wire FETs are proposed as technology options for sub 22 nm node gate lengths [1][2]. However, thermal management in these nanoscale devices on SOI wafer is a big concern, which is due to (i) significantly poor thermal conductivity of ultra thin materials and (ii) tight thermal boundary conditions, i.e. cooling conditions, resulting in performance degradation and serious reliability issues. Recently ultra thin body (UTB) SOI MOSFETs were demonstrated as an option for sub 20nm gate lengths, with advantages such as lower 3D parasitics, ease of incorporating strain effects, flexibility in width selection and suppressed random dopant fluctuation [3] [4]. However, its self heating behavior needs further evaluation and clarification. So far, previous work reported in [5] modeled a thermal resistance network for FinFET devices and evaluated the impact of device dimensional parameters on self heating behavior. However, this work does not provide any guidelines for BEOL design, which is due to (i) the nature of approximations for heat generation, i.e. $Q=V_{DD}I_{ON}$ and (ii) inappropriate thermal boundary conditions. Furthermore, the thermal boundary conditions considered in the work done by [6] restrict a proper understanding of heat flow in real 3D geometries and complex BEOL. A proper understanding/characterization of the heat transport in the BEOL of a logic circuit network is therefore still missing. Currently, during the technology development phase, the device design is driven by front end of the line (FEOL) topics and core circuit performance optimization without taking into account the back end of the line (BEOL) issues in great detail regarding its impact on thermal management or EOS/ESD reliability issues. This imposes serious reliability constraints for BEOL designs. Keeping this objective in mind, for the first time we accounted for complete BEOL/FEOL of FinFET and ETSOI devices, while investigating impact of layout/device/technology parameters on the self-heating effects providing useful FEOL/BEOL design guidelines.

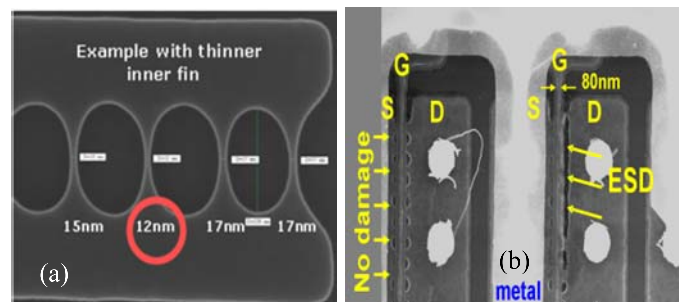

Figure 1. (a) SEM image showing ±3nm deviation in Fin thickness. This leads to weak spots which may cause early damage under thermal stress. (b) SEM image showing typical failed device from high current or thermal stress [9].

II. Device Fabrication and Experimental Results

Undoped trigate FinFET devices were realized with a metal mid-gap TiN gate and SiON dielectric (EOT=1.9nm). The fins have a target width of 15nm, length of 70nm and are of 60nm height. **Fig. 1a** shows SEM picture of the realized Fins where ±3nm deviation from the targeted fin thickness is evident. This leads to physically weak spots and may cause early failure

Authorized licensed use limited to: J.R.D. Tata Memorial Library Indian Institute of Science Bengaluru. Downloaded on October 06,2022 at 06:25:50 UTC from IEEE Xplore. Restrictions apply.

under high thermal stress. Moreover, **fig 1b** shows SEM image of a typical FinFET device failed under the ESD stress. It is evident from mitigated impact ionization at higher currents (**fig. 2**) that the device suffers from the significant self heating, which leads to device failure after a few measurements as shown in **Fig. 3** Furthermore, we also found that current density under nominal operating conditions was getting close to as it was under ESD failure conditions [8], which further validates presence of strong heating.

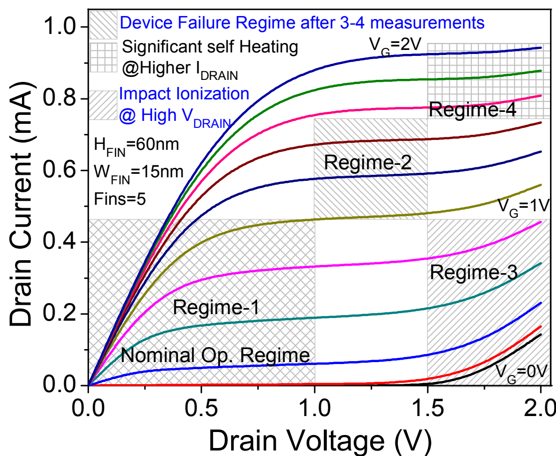

Figure 2. Measured I_D-V_D characteristics of FinFET device. Mitigated impact ionization at higher bias conditions proves the presence of significant self heating.

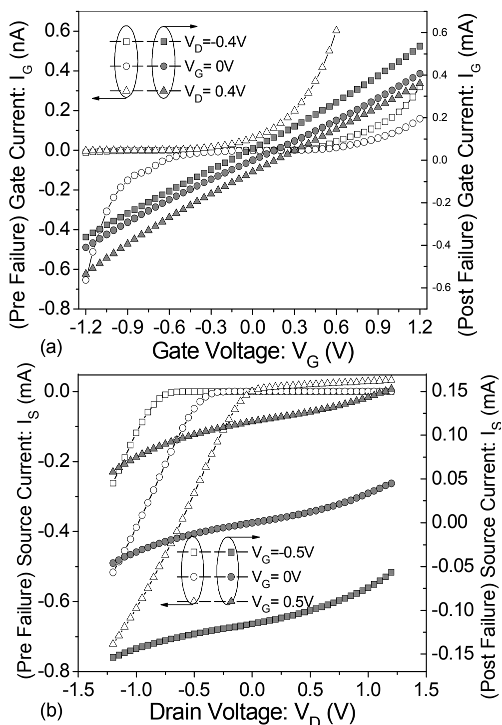

Figure 3 (a) & (b). Post and pre failure currents. Fig. shows significantly high leakage current after 2-3 measurements. Post failure characteristics shows very low resistance path from drain-to-gate and drain-to-source, which in conjunction with self heating behaviour (Fig. 2) validates an early thermal failure.

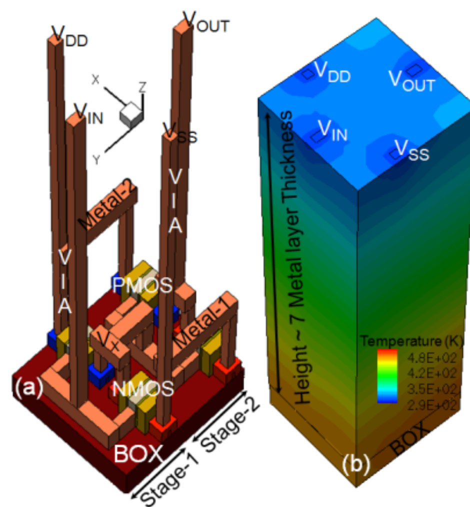

Figure 4 (a): Two stage inverter-driving-inverter with full BEOL definition. (b) Temperature distribution up-to 7 metal layers shows significance of BEOL definition.

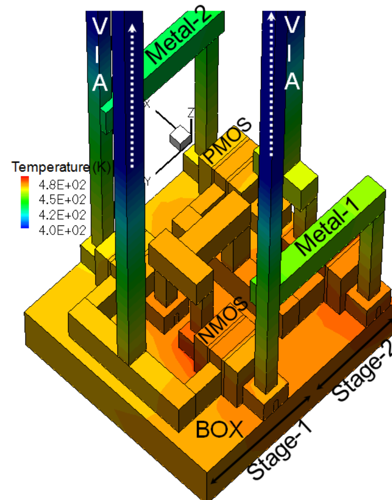

Figure 5. Temperature distribution along the interconnect metals. Figure shows max. Contribution of heat flux through metal interconnects.

III. Simulation Framework and Design Guidelines

In order to fully capture the heat flow through various interconnect lines and inter layer dielectric (ILD) regions, a two stage inverter-driving-inverter with full BEOL definition is realized for TCAD simulations (**Fig. 4a**). Devices are realized as given in our previous work [4] and their layout is defined by using a predictive technology model [7]. The simulation approach used in this work to extract quasi-static temperature (or the worst case temperature), which should be the cumulative rise in peak temperature after 1000s of pulses, is discussed in our work elsewhere [9]. Thermal properties of various regions are calibrated as per their exact dimensions and region specific material used on silicon. **Fig. 4b and 5 and 6b** show the temperature distribution across BEOL (thickness equivalent up-to 7 metal layers), along the interconnect metals and in the active device region. Fig. 6a shows maximum contribution of heat flux through metal interconnects instead of Si substrate. This is attributed to (i) very high thermal resistance contributed by the BOX and Silicon substrate region,

Authorized licensed use limited to: J.R.D. Tata Memorial Library Indian Institute of Science Bengaluru. Downloaded on October 06,2022 at 06:25:50 UTC from IEEE Xplore. Restrictions apply.

compared to metal interconnects and (ii) significantly lower thermal conductivity of Silicon region at higher temperatures. This eventually leads to most of the heat sink into the overlying back end metallization instead of Silicon substrate. Furthermore, temperature distribution across the interconnect metals and active (Si) region along different planes show: (i) NMOS has a higher temperature rise as compared to PMOS, which is due to NMOS having a relatively higher drive current (ii) Devices close to I/O pads have lower heating as compared to others, which is due to their better thermal coupling. (iii) Maximum heat flux is through the interconnect (metal) lines as compared to BOX and ILD layers. Significance of proper BEOL definition for electrothermal modeling is evident from **Figs. 4, 5 & 6. Fig. 7** shows calibration of TCAD models for 3D drift diffusion transport (considering quantum corrections) with our measurement data.

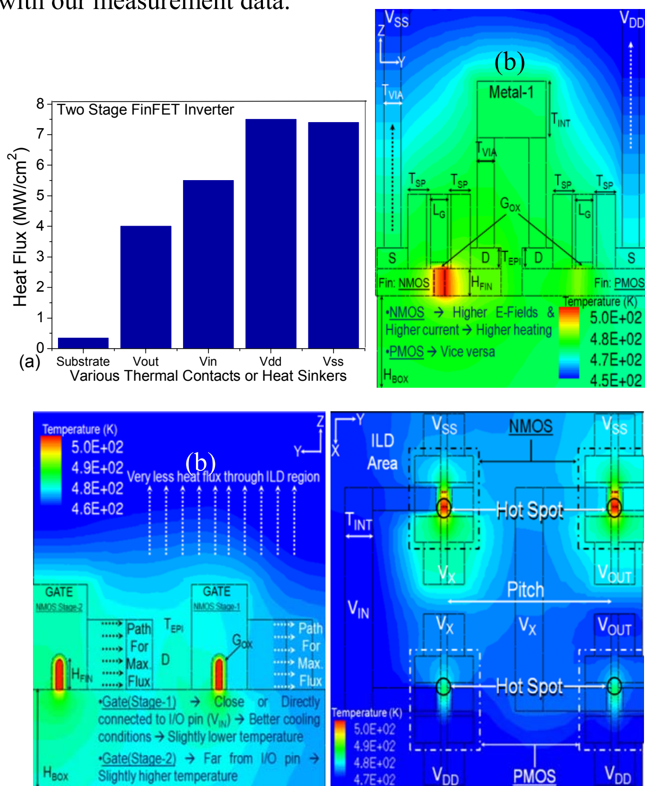

Figure 6. (a) Heat Flux per unit area from various heat sinkers (or thermal contacts). (b) Temperature distribution across the interconnect metals and active (Si) region along different planes. Fig. shows: (i) NMOS has higher temperature rise as compared to PMOS, (ii) Devices close to I/O pads have lower heating as compared to others, which is due to better cooling condition, (iii) Max. Heat flux is through interconnect metals as compared to BOX and ILD layers.

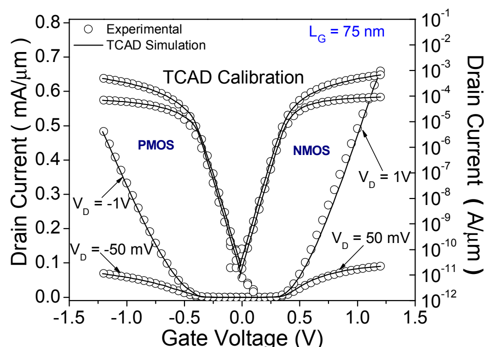

Figure 7. Calibration of TCAD models for 3D drift diffusion transport (considering quantum corrections) with experimental data.

A. *Impact of Device Scaling*

Fig. 8 shows impact of (a) SOI/fin thickness and (b) scaling of active width of ETSOI device & fin width of FinFET device on self-heating effects. Moreover, **fig. 9** shows self heating effect due to ETSOI device. It is also evident from **Figs. 8 & 9** that ETSOI device has a more relaxed self heating compared to FinFET device. Rise in lattice temperature depends on (i) volume of power source (active region), (ii) volume of heat sink in the surrounding and (iii) thermal boundary condition in the exteriors- defines the quasi-static temperature rise. Relaxed lattice temperature can be attributed to a smaller power density in a given active Si volume compared to FinFET device, while keeping the BEOL definition same for both the devices (**Fig. 9**). Furthermore, fin height & fin width scaling in FinFET devices and SOI thickness & active area scaling in ETSOI devices reduce the temperature rise due to simultaneous drop in drain current. Moreover, channel length scaling (data not shown here) in both devices leads to a temperature rise, which is due to the slightly higher drive currents and slightly higher thermal resistance, i.e. lower thermal coupling with I/O pads. This gives an indication that technology scaling should not affect the thermal performance significantly if fin width and fin height are scaled simultaneously with channel length scaling.

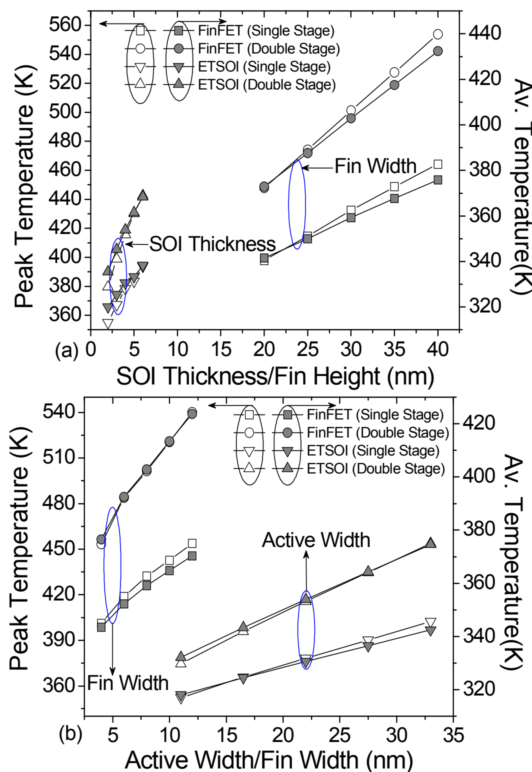

Figure 8. Impact of (a) SOI thickness and Fin thickness and (b) active width of ETSOI device & Fin Width of FinFET device scaling on self heating effects. Fig. shows: (i) ETSOI device has smaller temperature rise as compared to FinFET device, which is due to smaller power density in a given volume compared to FinFET device, (ii) Fin height scaling in FinFET device and active area scaling in ETSOI device relaxes the self heating effects and (iii) SOI thickness scaling in ETSOI and Fin Width scaling in FinFET device reduces the temperature rise due to simultaneous drop in drain current.

Authorized licensed use limited to: J.R.D. Tata Memorial Library Indian Institute of Science Bengaluru. Downloaded on October 06,2022 at 06:25:50 UTC from IEEE Xplore. Restrictions apply.

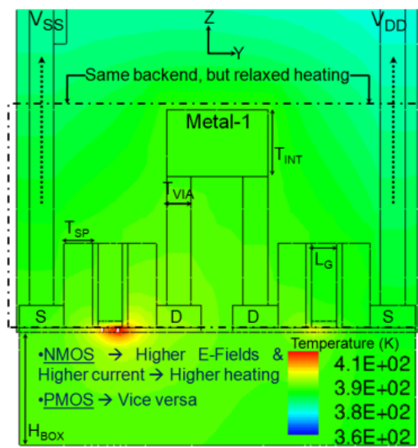

Figure 9. Self heating in ETSOI devices. It shows relaxed self heating in ETSOI device, which is due to lower volume of power source while keeping the BEOL same.

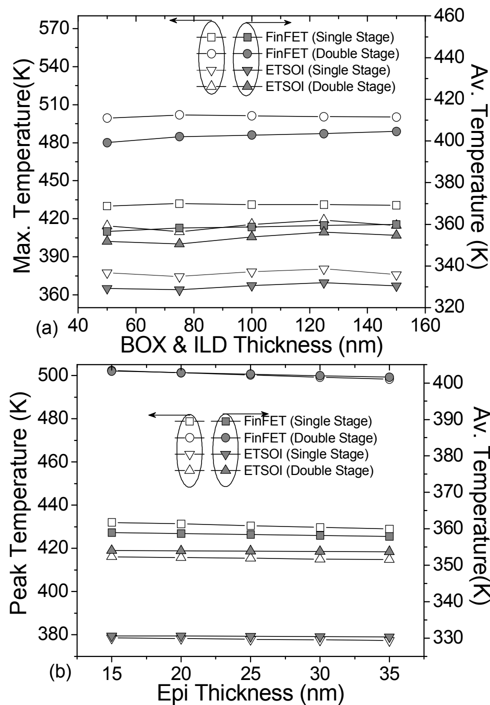

Figure 10. Impact of (a) Buried Oxide, Inter Layer Dielectric thickness (at level M1) and (b) Epitaxial raised S/D thickness on the self heating behaviour. It shows that both the technology parameters have no impact on the self heating in FinFET and ETSOI devices.

B. Impact of Technology Parameters

Fig. 10 show that (a) Buried oxide (BOX); (b) Inter-Layer Dielectric (ILD at Metal-1); and (c) Epitaxial (raised S/D); thickness has no impact on the self heating in both FinFET and ETSOI devices unlike to [6]. BOX thickness has a negligible impact due to the significantly lower thermal coupling between power source (active Si) and heat sink (external boundary) through BOX regions as shown in Fig. 6a. A lower ILD thickness does not impact much on the interconnect distribution at the Metal-1 level, i.e. total thermal resistance from metal lines, which attributes to no change in the self heating behavior while changing ILD thickness. Note that increasing (decreasing) the thickness of epitaxial region increases (decreases) the volume of heat sink but at the same time it also increases (decreases) thermal resistance and eventually balance the self heating behavior.

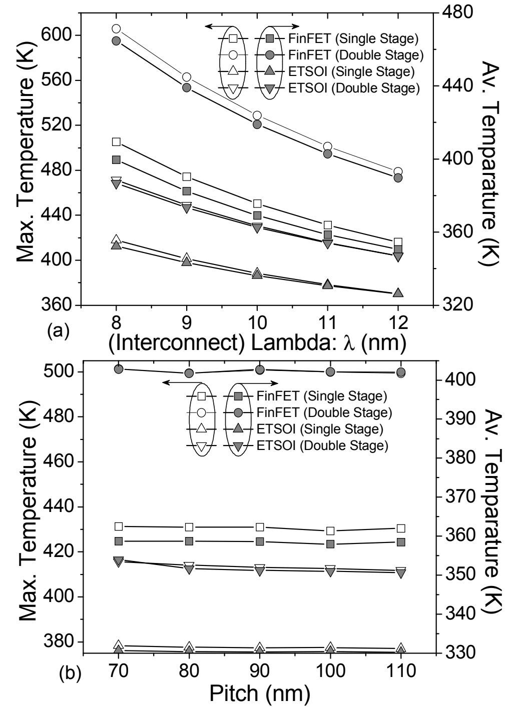

Figure 11. Impact of (a) Interconnect scaling and (b) pitch scaling on the self heating behaviour. It is evident from figure that active area pitch can be scaled without affecting self heating behaviour, whereas interconnect height and interconnect thickness impacts significantly on the self heating behaviour.

C. Impact of Layout Parameters

Fig. 11a shows that interconnect height and thickness (~λ) significantly impact the self heating behavior. Increasing (decreasing) the interconnect dimensions improves (degrades) the thermal performance due to lower (higher) thermal resistance. Moreover, **fig. 11b** shows that active area pitch can be scaled without affecting the self heating behavior. This is attributed to a balance between (i) rise in temperature due to higher coupling between two active regions and (ii) fall in temperature due to improved coupling between heat sink with active region.

D. Impact of Materials Thermal Properties

Fig. 12 shows that (a) interconnect and (b) interlayer dielectric thermal conductivity have a significant impact on the self heating in both FinFET and ETSOI devices. Increasing interconnect material's thermal conductivity reduces the thermal resistance and improves the cooling conditions, which eventually leads to a significantly relaxed temperature rise. Increasing the ILD thermal conductivity improves the thermal coupling of ILD layer with metal interconnects and eventually relaxes the self heating effects. Moreover, it also shows a significantly relaxed lattice temperature when interconnect thermal conductivity was equivalent to carbon nano tube

Authorized licensed use limited to: J.R.D. Tata Memorial Library Indian Institute of Science Bengaluru. Downloaded on October 06,2022 at 06:25:50 UTC from IEEE Xplore. Restrictions apply.

(CNT) like material (~35W/K-cm). On the other hand, (a) gate oxide and (b) BOX thermal conductivity has almost a negligible impact on self heating effects. This is attributed to (i) very small physical thickness of gate-oxide material and (ii) due to a very weak thermal coupling between the heat sink and active region through BOX region, respectively.

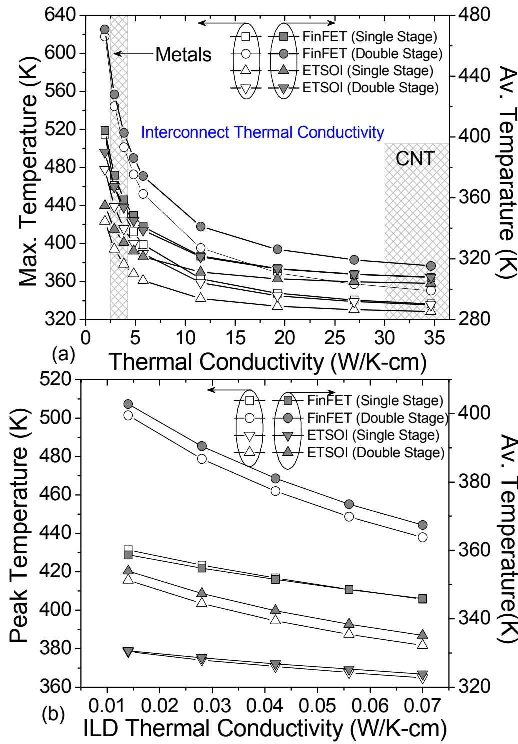

Figure 12. Impact of various material parameters (a) interconnect and (b) interlayer dielectric thermal conductivity on self heating behaviour. It is evident from figure that both the parameters have significant impact on self heating in FinFET and ETSOI devices.

E. Impact of Power Supply Scaling

Fig. 13 shows significantly relaxed lattice temperature with the supply voltage scaling. It is evident from the figure that power supply scaling is also one of the key requirements to mitigate the self heating effects in nanoscale devices.

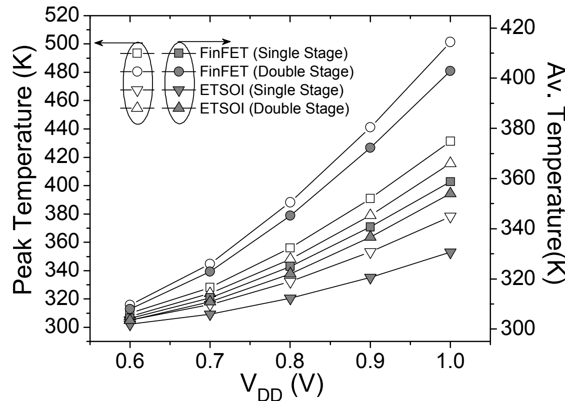

Figure 13. Impact of power supply scaling. Peak temperature (i.e. Joule Heating) proportional to I instead of I^2, which is unlike to standard theories for bulk CMOS devices.

IV. Conclusion

We reported thermal failure of FinFET devices, even under nominal operating conditions. A new simulation framework and importance of thermal boundary condition is discussed for accurate electrothermal modeling and for gaining a better physical insight towards self heating behavior in nano-scale SOI devices. Our results demonstrate that in order to build 3D thermal resistance network for FinFET and ETSOI like devices, proper BEOL definition is as important as the FEOL definition. It was found that scaling of all the layout/technology parameters does not have the same impact on the self heating behavior, i.e. only few parameters in BEOL have the maximum impact and need to be considered while building electrothermal models. Based on this study a design guideline is extracted and summarized in **table-I**. Furthermore, either (i) power supply scaling or (ii) use of ETSOI devices in conjunction with CNT like high thermal conductivity materials for interconnects are the key requirements in order to alleviate the thermal issues in these nano-scale CMOS technologies.

TABLE-I: Influence of Technology and design parameters on the self heating behavior. (Markers: - ↑→Rising, •→No significant change and ↓→Falling)

Parameters	Impact
Metal Thermal Cond. (K_M)	$\uparrow K_M \rightarrow \downarrow\downarrow T_{MAX}$
Power Supply (V_{DD})	$\downarrow V_{DD} \rightarrow \downarrow\downarrow T_{MAX}$
Interconnect Lamda (λ)	$\uparrow \lambda \rightarrow \downarrow T_{MAX}$
Fin Height (H_{FIN})	$\downarrow H_{FIN} \rightarrow \downarrow T_{MAX}$
Active Si Width (W_A)	$\downarrow W_A \rightarrow \downarrow T_{MAX}$
BOX Thickness (T_{BOX})	$\downarrow\uparrow T_{BOX} \rightarrow \bullet\ T_{MAX}$
Epi Thickness (T_{EPI})	$\downarrow\uparrow T_{EPI} \rightarrow \bullet\ T_{MAX}$
ILD Thickness (T_{ILD})	$\downarrow\uparrow T_{ILD} \rightarrow \bullet\ T_{MAX}$
Spacer Width (W_{SP})	$\uparrow W_{SP} \rightarrow \downarrow T_{MAX}$
SOI Thickness (T_{Si})	$\downarrow T_{Si} \rightarrow \downarrow T_{MAX}$
G_{OX} Thermal Cond. (K_{GOX})	$\downarrow\uparrow K_{GOX} \rightarrow \bullet\ T_{MAX}$
BOX Thermal Cond. (K_{BOX})	$\downarrow\uparrow K_{BOX} \rightarrow \bullet\ T_{MAX}$
Fin Width (W_{FIN})	$\downarrow W_{FIN} \rightarrow \downarrow T_{MAX}$
ILD Thermal Cond. (K_{ILD})	$\uparrow K_{ILD} \rightarrow \downarrow T_{MAX}$
Gate Length (L_G)	$\downarrow L_G \rightarrow \uparrow T_{MAX}$

Acknowledgment

Authors would like to acknowledge Prof. Dinesh Kumar Sharma (EE Department, IIT Bombay) for some insightful discussions.

References

[1] H. Kawasaki *et. al.*, IEDM-2008.

[2] Y. Jiang, *et. al.*, IEDM 2008.

[3] O. Weber *et. al.*, IEDM, 2008.

[4] Mayank Shrivastava, *et. al.*, IEDM, 2009.

[5] B. Swahn, *et. al.*, Transactions on VLSI System, July 2008.

[6] S. Kolluri, *et. al.*, IEDM, 2007.

[7] Predictive Technology Models [Online: www.eas.asu.edu/~ptm].

[8] Harald Gossner, *et. al.*, IEDM-2006.

[9] Mayank Shrivastava, *et. al.,* T-ED, June 2010.

Authorized licensed use limited to: J.R.D. Tata Memorial Library Indian Institute of Science Bengaluru. Downloaded on October 06,2022 at 06:25:50 UTC from IEEE Xplore. Restrictions apply.