RC23834 (W0512-084) December 19, 2005
Electrical Engineering

IBM Research Report

Self-Consistent Electro-Thermal Modeling for FinFET Technologies

Satish Kumar[1], Rajiv V. Joshi[2], Ching-Te Chuang[2], Keunwoo Kim[2], Jayathi Y. Murthy[1]

[1]Department of Mechanical Engineering
Purdue University
585 Purdue Mall
West Lafayette, IN 47907

[2]IBM Research Division
Thomas J. Watson Research Center
P.O. Box 218
Yorktown Heights, NY 10598

Research Division
Almaden - Austin - Beijing - Haifa - India - T. J. Watson - Tokyo - Zurich

LIMITED DISTRIBUTION NOTICE: This report has been submitted for publication outside of IBM and will probably be copyrighted if accepted for publication. It has been issued as a Research Report for early dissemination of its contents. In view of the transfer of copyright to the outside publisher, its distribution outside of IBM prior to publication should be limited to peer communications and specific requests. After outside publication, requests should be filled only by reprints or legally obtained copies of the article (e.g., payment of royalties). Copies may be requested from IBM T. J. Watson Research Center , P. O. Box 218, Yorktown Heights, NY 10598 USA (email: reports@us.ibm.com). Some reports are available on the internet at http://domino.watson.ibm.com/library/CyberDig.nsf/home.

Self-Consistent Electro-Thermal Modeling for FinFET Technologies

Satish Kumar
Department of Mechanical Engineering, Purdue University
585 Purdue Mall, West Lafayette, IN-47907
kumar21@purdue.edu

Rajiv V. Joshi
IBM T J Watson Research Center
Yorktown Hts, NY-10598
rvjoshi @ us.ibm.com

Ching-Te Chuang
IBM T J Watson Research Center
Yorktown Hts, NY-10598
ctchuang @ us.ibm.com

Keunwoo Kim
IBM T J Watson Research Center
Yorktown Hts, NY-10598
kkim @ us.ibm.com

Jayathi Y. Murthy
Department of Mechanical Engineering, Purdue University
585 Purdue Mall, West Lafayette, IN-47907
jmurthy@ecn.purdue.edu

ABSTRACT
A self-consistent 3-D computational modeling methodology for analyzing thermal and electrical transport in nano-scale devices is developed. The methodology is applied to multi-finger double-gate FinFETs to compute the device characteristics, spatial power distribution, and spatial temperature distribution. A methodology for computing the optimum operating voltages based on the required range of operating temperature is presented. The effect of device dimensions on temperature rise and device performance for poly-gate and metal-gate devices is investigated. Results show that mid-gap metal-gate device is better than band-edge metal-gate device and poly-gate device from both heating and device performance perspective. The mid-gap metal-gate device causes 10 % less heating in the equal performance case with respect to poly-gate counterpart, and offers 1/6 lower leakage power for the equal on-current case with respect to metal-gate counterpart.

Categories and Subject Descriptors
I.6.5 [Simulation and Modeling]: Model Development

General Terms
Algorithms, Performance, Design

Keywords
FinFet, Power

1. INTRODUCTION
Scaling of conventional device in sub-nanometer CMOS technologies brings many challenges such as severe short-channel effects and self-heating effects. High sub-threshold leakage and exponential rise in gate tunneling leakage due to thin gate dielectrics could make the technology scaling unviable [1,2]. Double-gate devices such as FinFET alleviate the short channel effects due to superior gate-to-channel coupling [2-6], allowing for aggressive device scaling. In DG-FinFETs, thicker oxide can be used to reduce gate-oxide tunneling current; and their fabrication process is compatible with conventional CMOS technology [1,2].

Although FinFETs seem promising for future generation electronics, self-heating effects could impose severe limitations on the commercialization of these devices. High heat dissipation is consequence of high packaging density, low thermal conductivity of surrounding gate-dielectrics (SiO_2) and reduced thermal conductivity of thin-Si film due to phonon boundary scattering and phonon confinement effects [2]. High temperature rise could have harsh effects on the device performance, which compels to analyze the effect of temperature rise on device performance for accurate predictions of electrical characteristics of these devices. This necessitates the development of a coupled electro-thermal transport model for these devices. Device electrical characteristics computed in 2-D computational domain are reasonably accurate if calibrated properly, while heat-diffusion problem is 3-D in nature as heat spreading occurs in entire substrate, metal interconnects and adjoining devices. Computing device characteristics in 3-D domain (as required for thermal transport) is computationally very expensive and almost infeasible with exiting commercial solvers (Medici, Taurus etc.) for multi-finger devices, while 2-D domain may lead to erroneous results [7].

In this paper, a computational technique/methodology is presented to couple the 2-D electrical transport model to 3-D thermal transport model in a self-consistent manner for accurate predictions of both electrical and thermal characteristics. To

account for reduced thermal conductivity of Si thin film (due to phonon boundary scattering and phonon confinement), experimentally extracted thermal conductivity of confined thin Si layer is used [2]. The developed computational method is applied to compare the poly-gate devices to metal-gate devices. The effect of non-uniform power distribution inside the channel of the device is explored. Finally, band-edge and mid-gap metal-gate devices are analyzed, and a methodology for predicting optimum operating voltage is presented.

2. COUPLED ELECTRO-THERMAL MODELING TECHNIQUES

Coupled electro-thermal model is developed using the commercial device and thermal simulation packages to analyze the effect of heat dissipation on the electrical characteristics. The device simulations are performed using Taurus [8], while thermal simulations are performed using Fluent (Fourier's law used) [9]. A two-dimensional cross-section of the device is considered for the electrical simulation (Fig. 1). The domain of thermal simulation consists of array of 6 fingers, five layer of metallic interconnects, back oxide and the Si-substrate (Fig. 2). Each finger has 10 fins and the total number of fins in the device under consideration is 60 [2]. Fig. 3 shows the flow chart of the methodology used to couple the electrical transport to thermal transport. At first, electrical simulations are performed to compute the device characteristics and spatial power distribution at various biased voltages across the device. This spatial power in 2-D electrical grid is mapped to 3-D thermal grid, which is then used in the thermal simulations to compute the spatial temperature distribution inside the device. Numerical results show that spatial temperature variation inside the device-channel is not very high and an average temperature of the device can be computed (see Sec. 3.1), which can be further used as representative device temperature for electrical simulations. Thermal simulations are performed for different heat sources as input, which corresponds to different biased voltage across the device. A polynomial relation is obtained between input heat source and average device temperature. The device simulation and thermal simulation is performed in iterative loop at each bias voltage till convergence using the computed polynomial function (right side loop in Fig. 3). Convergence is achieved when the average device temperature computed from the thermal simulations is same as the temperature at which the device simulation is performed to compute the input power or heat source for the thermal simulation (Fig. 3).

Fin dimensions corresponding to 45 nm technology node is used for most of simulations presented here unless mentioned separately (channel length L_C ~ 25 nm, fin width W ~ 9 nm and fin height H ~ 40 nm). Experimentally extracted thermal conductivity of thin-Si film used in simulations is 40 W/m^{o}K (1/4th of bulk value) [2]. For the simulations, adiabatic boundary conditions have been assumed on the four sides of the computational domain. These boundary conditions reflect the fact that no heat practically leaves through those walls and no temperature gradient is expected on those sides. For the top boundary, a convective boundary condition has been assumed to model a solder bump with a convective coefficient same as that of air. For the bottom surface, an isothermal boundary condition of 300 ^{o}K has been imposed.

3. RESULTS AND DISCUSSION

3.1 Non-uniform Power Distribution

Effect of non-uniform power distribution (inside Si-fin) on temperature distribution is analyzed for 45 nm technology. A typical power distribution in a lateral cross-section of fin is shown in Fig. 4. Power distribution is computed using $P = \vec{J} \bullet \vec{E}$, where J and E are current and electric field vector. High power dissipation inside the fin close to both front and back gate-oxides is observed as expected due to high current in the inversion layer of the devices. Peak power locations are close to the drain side of the channel (Fig. 4), which could lead to hot spots in the device. Spacing and structure for grid-cells in 2-D electrical simulation and 3-D thermal simulation are different. Using a cubic interpolation scheme, power distribution in 2-D electrical grid is converted to power distribution in 3-D thermal grid. In each lateral layer of thermal grid, similar power distribution is assumed for thermal simulations. Spatial temperature distributions in a fin-cross-section are shown in Fig. 4 and Fig. 5 for uniform and non-uniform power distribution, respectively. The uniform power distribution is computed taking average of power in all grid-cells. For uniform power distribution symmetric temperature profile is observed (Fig. 5), while for non-uniform power distribution peak temperature close to drain is observed as expected due to high power dissipation close to drain electrodes (Fig. 6). Even though the temperature distribution is not symmetric in a lateral cross-section of fin, the variation in temperature is less than 2 ^{o}K, which suggests that heat is equilibrated in the device when steady state is reached.

To understand the physics behind heat transport, simulations of different device dimensions and boundary conditions are performed. Temperature variation in a lateral cross-section of fin increases with increasing fin dimension. The above numerical observations suggests that low volume of nanoscale device and thick oxide between substrate and fin region helps in lateral spreading of heat, which is reflected in low temperature variation in a lateral plane of device. Transport in a device with single fin and an isolated thermal boundary (except at bottom where fixed temperature of 300 ^{o}K is applied) is also simulated. It is observed that for this device the temperature difference in lateral cross-section increases to 18 ^{o}K. In some previous numerical simulation of thermal transport in electronic devices, single fin with insulated boundaries is considered for the modeling [10]. The present numerical simulation show that this type of boundary condition may give reasonable results for large size fins ($\sim \mu m$), but is completely inappropriate for single-fin nanoscale-devices and one must consider array of fins with bigger computational domain to capture the real temperature distribution. Commercial solvers (~ Medici, Taurus) are very expensive to solve both device and thermal simulations in array of devices and there is need to solve electrical (2-D) and thermal (3-D) transport using different computational domain and couple them in efficient manner to reduce the computational burden.

3.2 Metal Gate vs Poly-Si Gate Devices

Higher thermal conductivity of metals with respect to poly-Si makes them better gate materials for effective spreading of heat inside the device, which could alleviate the temperature rise. To compare the performance and temperature rise of the device, n^+-poly-Si gate device is compared against the metal-gate (K~190

W/m^{o}K) device. Metal-gate device with work-function (ϕ) of 4.18 eV has similar I_{off} compared with the poly-gate device (which has less I_{on} because of the poly-silicon depletion, See Fig. 7). I_{on} similar to poly-gate device can be achieved with metal-gate device of ϕ ~ 4.33 eV. In this case, I_{off} for the metal-gate device reduces significantly (Fig. 7). Thermal simulations for these devices show that temperature rise in poly-gate device is 20 ^{o}K higher than the metal-gate device of ϕ ~ 4.33 eV. Metal-gate device is observed to be better than poly-gate device from both device performance and effective heat removal perspectives.

3.3 Coupled vs Non-Coupled Solver

The effect of heat-dissipation on device performance is analyzed using developed coupled electro-thermal modeling methodology. I_{DS}-V_{GS} characteristics and average temperature rise in the device computed from coupled and un-coupled solver are compared for poly-gate device and metal-gate device of ϕ ~ 4.18 eV and 4.33 eV in Fig. 8 and 9. Here coupled solver follows the computational technique described in Sec. 2 and Fig. 3, while uncoupled solution corresponds to case when the device and thermal simulation is performed just once. For metal gate device, 10-15% higher I_{on} and 20-25 ^{o}K higher average temperature (T_{av}) is observed compared with the uncoupled solver, while for the poly-gate device 6% higher I_{on} and 11 ^{o}K higher average temperature (T_{av}) is observed. The device performance is supposed to deteriorate with increasing temperature, but the opposite is observed here due to opposing effects of mobility μ and V_t. It is observed that with coupled solver the mobility decreases (due to higher temperature), but V_t of device also decreases. The increase in I_{on} due to reduced V_t dominates over the decrease in I_{on} due to reduced μ, resulting in net increase of I_{on} of the device.

3.4 Optimum Operating Voltages

The coupled simulations show that the peak temperature can go above 500 ^{o}K for the array of device under consideration. Although better cooling techniques could be applied or spacing of fins could be altered to reduce the temperature rise, there is a need to predict an optimum operating voltage for the specific device design and operating temperature range under consideration. This optimum condition is influenced by the objectives like high I_{on} in saturation regime and low peak temperature in a specified range to prevent burnout of device. For the present device, I_{DS}-V_{DS} characteristics is computed for V_{GS} varying in the range of 0.4-1.4 V (Fig. 10) and corresponding peak temperature is shown in Fig. 11. The isotherm curves for temperature of 350 ^{o}K and 400 ^{o}K are also drawn on Fig. 10. If the desired range of operating temperature is 350-400 ^{o}K, the optimum operating voltage (V_{GS} = V_{DS} ~ 0.8 V) can be found in the shaded region between two isotherm curves based on the criterion of maximum I_{on} and requirement of operating in saturation regime (Fig. 10). The same methodology could be applied to compute the optimum operating voltage for any device knowing the required temperature range of operation.

3.5 Effect of Device Dimensions

The effect of device dimensions is analyzed by varying the channel length L_C in the range of 20-30 nm (Fig. 12) and fin height H in the range of 20-50 nm (Fig. 13). Decreasing L_C increases I_{on} (Fig. 12), but leakage power (short channel effects) and peak temperature also increase for both metal-gate device (ϕ ~ 4.33 eV) and n^+-poly-gate device. Increasing H increases the effective surface area for the current flow and I_{on} increases correspondingly (Fig. 13). The temperature rises linearly with increasing H. This is because the power dissipation (and hence heat generation) is proportional to the square of I_{on} , whereas the surface area for heat dissipation is only linearly proportional to H. Thus, for the design of FinFETs, the choice of fin height H is not only constrained by the electrical device width requirement and mechanical stability, but also by the thermal consideration, which must be considered to prevent thermal burnout of device.

3.6 Mid-gap vs Band-edge Metal Gate Devices

The coupled solver is used to compare the performance of mid-gap and band-edge metal-gate devices. For mid-gap metal-gate device undoped body is used, while for the band-edge metal gate device doped body is used. The I_{off} and I_{on} current of these two different devices are compared for L_C varying in the range of 20-30 nm in Fig. 14 and 15, respectively. Using mid-gap metal- gate (ϕ ~ 4.70 eV), I_{on} comparable to band-edge metal-gate (ϕ ~ 4.18 eV) can be achieved, while I_{off} is much lower for mid-gap metal-gate (low leakage power). Surface scattering and change in V_t with varying channel length seems to be two competing factors to affect device performance. At short L_C (~20nm) mid-gap metal-gate device have lower I_{on} compared with band-edge metal-gate device, but this trend reverses at long L_C (~30nm) (Fig. 15). Highly doped band-edge metal-gate device has higher surface scattering compared with undoped mid-gap metal-gate device due to higher normal electric field. Relatively low I_{on} for band-edge metal-gate device at long L_C (~30 nm) seems to be due to higher surface scattering. The trend reverses at short L_C (~ 20nm) as the reduction of V_t with reducing channel length plays a more influential role compared with the surface scattering.

4. CONCLUSIONS

Coupled electro-thermal modeling methodology has been developed using commercially-available, physics-based electrical simulator Taurus and thermal transport solver Fluent to analyze the effect of self-heating on device performance. Significant changes in electrical and thermal characteristics of devices using coupled solver are observed. For multi-fin FinFET, nanoscale effect of non-uniform spatial power distribution on temperature profile is not very significant. Temperature rise in devices limits the operating voltage of the device and optimum operating point could be derived based on the desired operating temperature range. Mid-gap metal-gate device is found to offer better electrical and thermal characteristics compared with band-edge metal-gate device or n^+ poly-Si gate device. In this work, thermal analysis is performed in the framework of classical Fourier-conduction equations. To get better accuracy, sub-continuum models for the thermal transport could be applied using the similar methodology. One of the interesting extensions of the present work will be to use the methodology at the chip level after computing the equivalent thermal resistance of different components of the chip.

5. ACKNOWLEDGMENTS

The authors would like to thank the IBM-management for supporting this work. K. Kim and C. T. Chuang were partially supported by the DAPAR contract NBCH30390004 for this work.

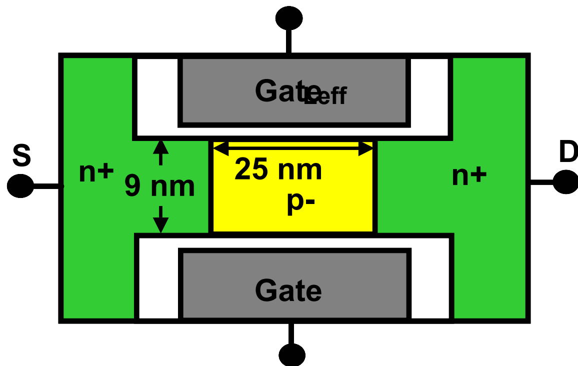

Figure 1. Schematic cross-section of computational domain for electrical simulation of double-gate FinFET device.

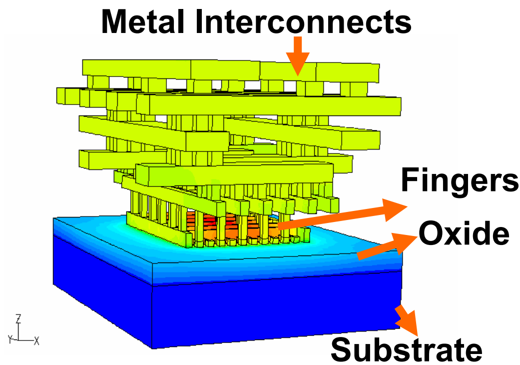

Figure 2. 3-D thermal domain for thermal simulation of multi-fin FinFET device.

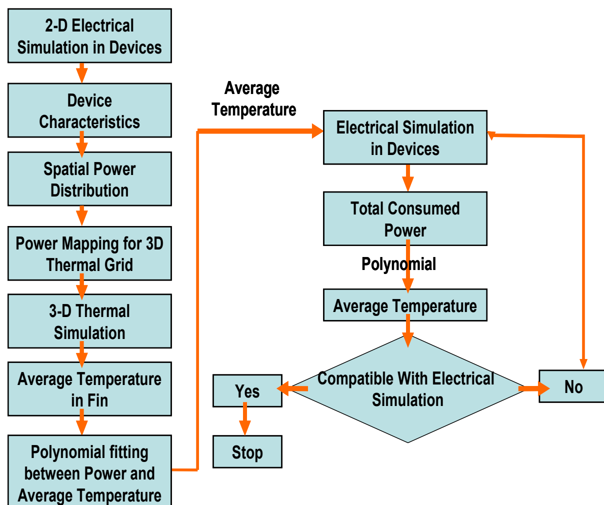

Figure 3. Flow-chart for coupled electro-thermal modeling.

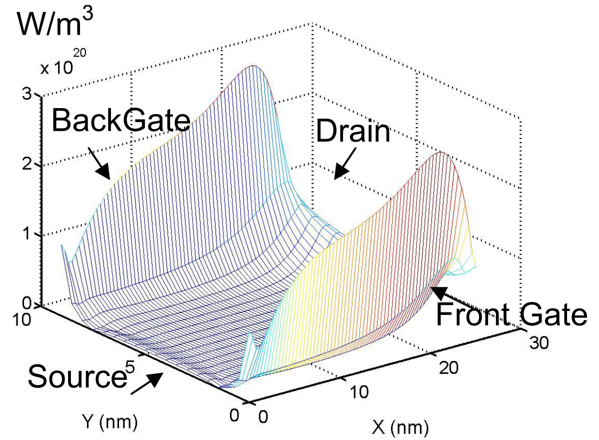

Figure 4. Typical power distribution in fin cross-section obtained from electrical simulation.

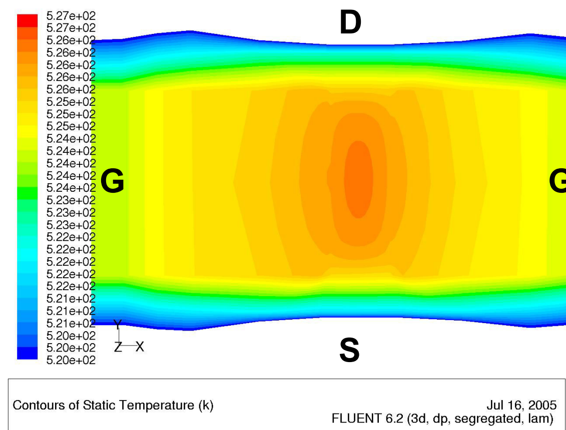

Figure 5. Temperature profile in fin cross-section for uniform heat source. D, S and G correspond to Source, Drain and Gate.

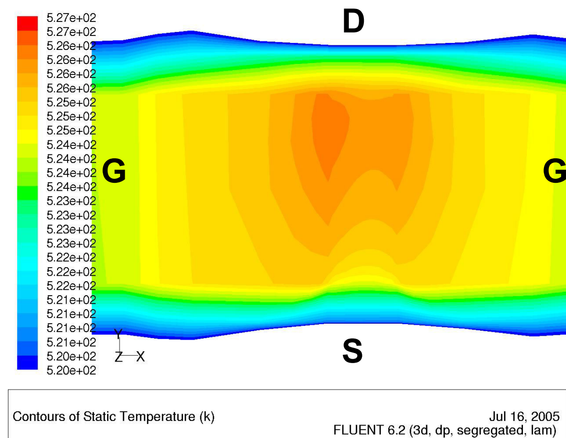

Figure 6. Temperature profile in fin cross-section for non-uniform heat source. D, S and G correspond to Source, Drain and Gate.

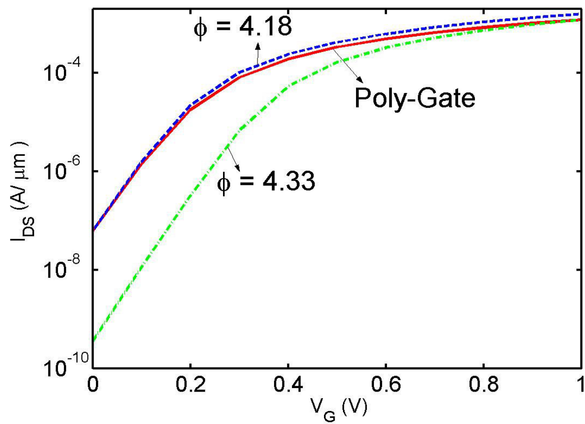

Figure 7. I_{DS}-V_{GS} characteristics at V_{DS} = 1 V for metal-gate and poly-gate devices.

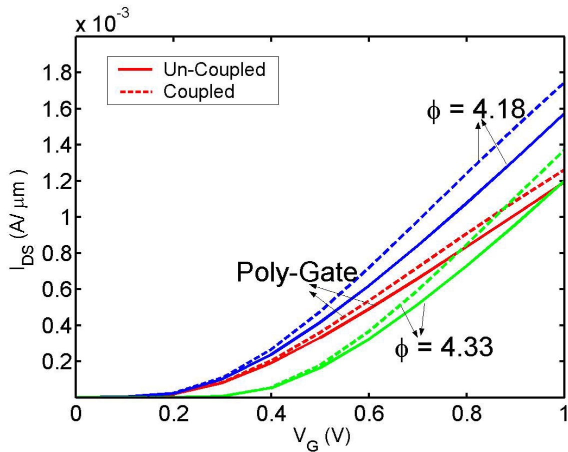

Figure 8. Comparison of I_{DS}-V_{GS} characteristics using coupled and non-coupled solver at V_{DS} = 1 V.

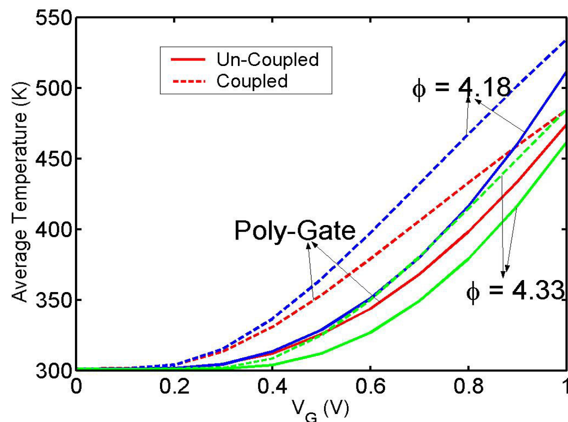

Figure 9. Comparison of average temperature using coupled and non-coupled solver at V_{DS} = 1 V.

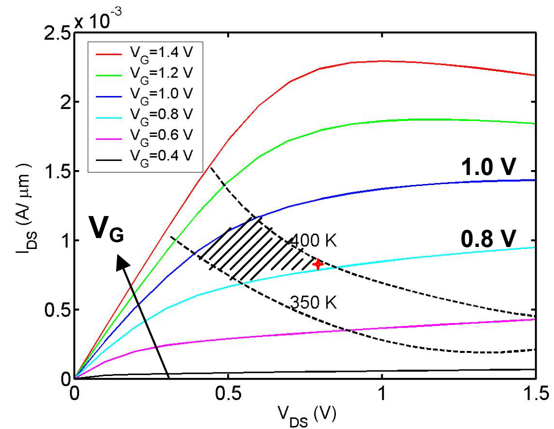

Figure 10. I_{DS}-V_{DS} characteristics at different V_{GS}. The optimal operating point (V_{GS} = V_{DS} ~ 0.8 V) is shown by asterisk in the shaded region ($\phi \sim$ 4.33 eV).

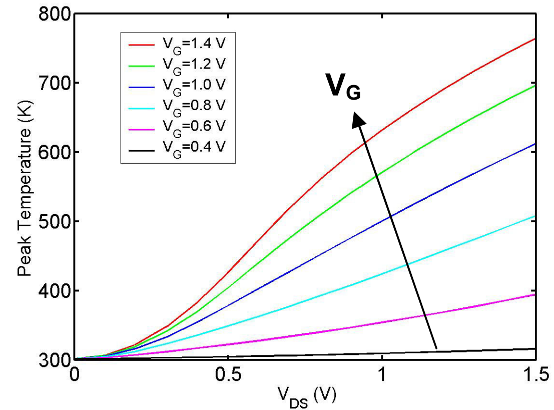

Figure 11. Peak temperature dependence on V_{DS} at different V_{GS} for metal-gate device ($\phi \sim$ 4.33 eV).

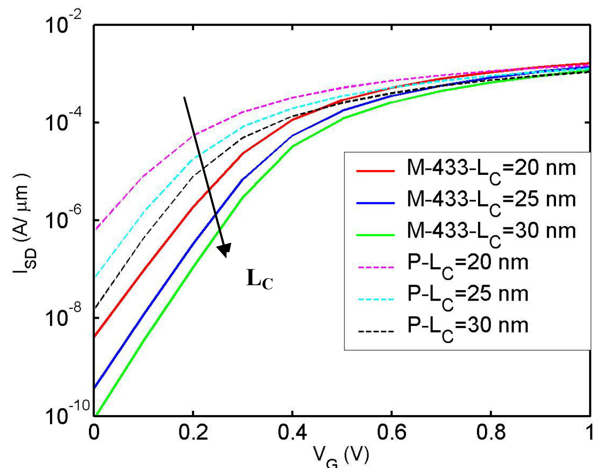

Figure 12. Comparison of I_{DS}-V_{GS} characteristics for metal-gate (M-L_C, ϕ= 4.33 eV) and poly-gate (P-L_C) device at V_{DS} = 1 V for various channel length.

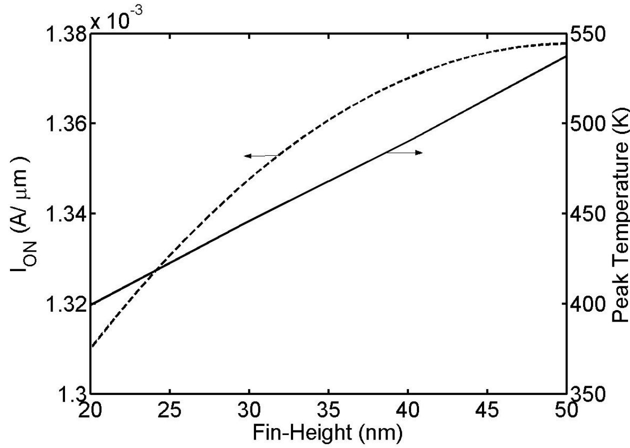

Figure 13. Effect of fin height on I_{ON} for metal-gate device (ϕ = 4.33 eV) at V_{DS} = 1 V.

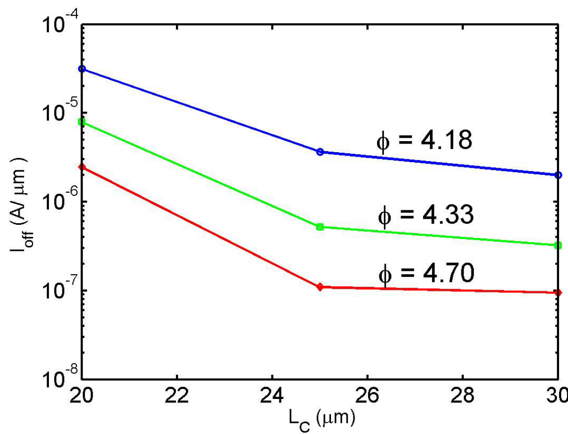

Figure 14. Comparison of I_{OFF} dependence on channel length L_C for metal-gate devices with different workfunctions (ϕ = 4.18, 4.33 and 4.70 eV) at V_{DS} = 1 V.

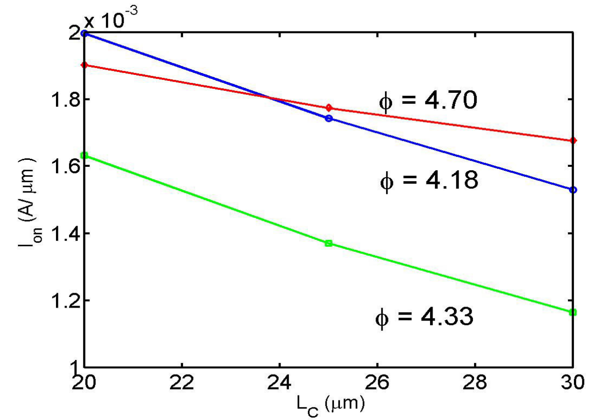

Figure 15. Comparison of I_{ON} dependence on channel length L_C for metal-gate devices with different workfunctions (ϕ = 4.18, 4.33 and 4.70 eV) at V_{DS} = 1 V.

6. REFERENCES

[1] Keunwoo, K. et al., *Leakage Power Analysis of 25-nm Double-Gate CMOS Devices and Circuits*, IEEE Trans. On Elec. Dev., 52, 5 (May 2005), 980-986.

[2] Joshi, R. V., Pascual-Gutiérrez, J. A., and Chuang, C. T. 3-D *Thermal Modeling of FinFet*, ESSDERC 2005.

[3] Wong, H. S. P., Chan, K. K, and Taw, Y., *Self-aligned Double-Gate MOSFET with a 25 nm Thick Silicon Channel*, IEDM 1997, 427.430.

[4] Nowak, E. et al., *Scaling Beyond the 65 nm Node with FinFet-DGCMOS*, CICC 2003, 125-129.

[5] Colinge, C. A., Colinge, J. P., Park, J. T. Proceedings of the IEEE International SOI Conference, 2001, 115-116.

[6] Joshi, R. V. et al., European Device Research Conf., Sept 2004.

[7] Alessandro, V. D., and Rinaldi, N., *A. Critical Review of Thermal Models*, Elec. State Dev, 46, 2002, 487-496.

[8] http://www.synopsys.com/products/tcad/taurus_medici_ds.html

[9] http://www.fluent.com

[10] Roig, J. *Modeling of non-uniform heat generation in LDMOS transistors*, Elect. State Dev., 49, 2005, 77-84.