TCAD for logic technology process development

S. Cea, S. Berrada, K. Ghosh, S. Hasan, P. Keys, R. Mehandru, B. Obradovic, V. Tiwari, C. Weber, and M. Stettler

Logic Technology Division, Intel Corporation, Hillsboro, OR, USA, email: stephen.m.cea@intel.com

***Abstract*— Process and device simulation has been invaluable for logic technology development for many technology nodes. The main goal of this work will be to review the broad and diverse simulation hierarchy that is used in industry to understand and optimize both current and future device technology options. This hierarchy spans both continuum modeling and atomistic methods / beyond continuum tools. Current process and device simulation results will be presented along possible extensions to the hierarchy to improve TCAD's ability to help technology development.**

Keywords— Process modeling, device modeling, stress, NEGF, metal modeling.

I. Introduction

Technology computer-aided design (TCAD) continues to be an integral part of modern technology development (TD). While TCAD departments use simulation to tackle issues beyond device performance[1] this talk will focus on process and device simulations which are essential components of TCAD used to help TD achieve its goals of scaling area and improving performance generation after generation. TCAD's role in the TD process include understanding current devices so that bottlenecks can be identified allowing optimization of these devices along with evaluating future device architecture and material options.

Figure 1 shows the diverse set of process and device simulators and methods needed to simulate today's advanced logic device options. The tools span a wide range of dimensions, from multiple atoms to multiple devices. In addition, the tools span a wide range of computational rigor from ab-initio to continuum and a wide range of expertise is needed. It is an exciting time to be a TCAD engineer with so many computational disciplines coming together in order to solve challenging technical problems. To help TD any combination of tools or methods may be needed, and interoperability is a valued feature. Industrial TCAD's challenges include choosing the "best" tool for the job, where the "best" will depend on the maturity of technology, accuracy needed and turnaround time requirements. In one common example continuum process simulations of the full transistor flow are used to predict the final device geometry along with active dopant and stress profiles, which are passed to a continuum device simulator to predict electrical performance. But it is important to notice that any simulator can be used to help technology for example density functional theory (DFT) simulations to look at defect formation energies, implant Monte Carlo to determine penetration depths, atomistic NEGF to determine the ballistic performance between materials or Kubo – Greenwood mobility calculators to determine how stress will affect a new material.

The main goals for this talk will be to describe the process and device simulation areas including some factors that drove the capabilities now used, some recent past / current examples and to point out some areas where improvement of simulation capability or fundamental understanding would enhance TCAD's ability to simulate current and future device options.

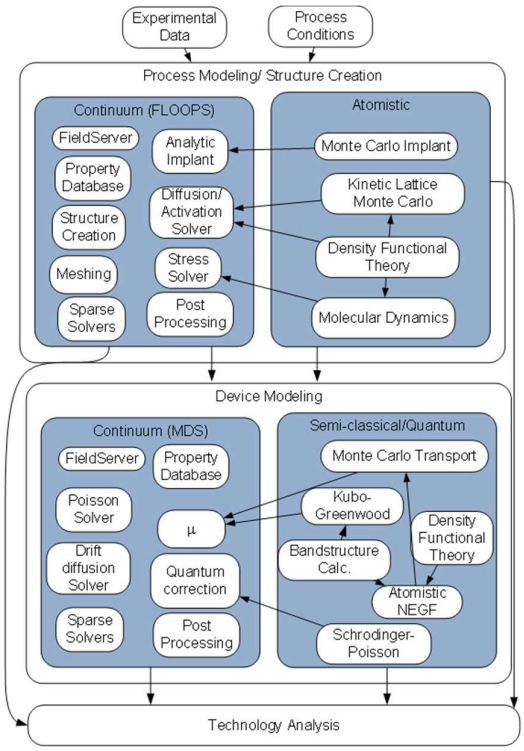

Figure 1. Overview of process and device hierarchy used including typical modeling flows and tool dependencies. Not all dependencies are depicted.

II. Process Modeling

At Intel continuum process modeling is performed using a proprietary version of FLOOPS [2]. As shown in figure 1 it contains physical models for implantation, dopant diffusion and stress simulation. To simulate a process flow structure creation, meshing and sparse matrix solvers are needed. This section will describe some details of these models, how atomistic methods are used and areas for improvement.

One of the main outputs of process modeling is the active dopant profiles in the device at the end of the process flow. A diverse set of physical models is needed to accurately simulate dopant diffusion and activation. Models for dopant diffusion, defect diffusion, defect clustering, dopant defect clustering, dopant or defect interactions with impurities, dopant activation / deactivation and dopant segregation are all needed for process simulation. This detailed model hierarchy results in many coupled PDEs that need to be efficiently solved. FLOOPS dial-an-operator approach to simulating PDEs greatly simplifies the development and maintenance of this complex system. Figure 1 shows how atomistic methods are used to help develop models in both silicon and novel materials. Examples of how DFT, kinetic lattice monte carlo and Molecular dynamics were used to develop models for dopant diffusion in SiGe and InAs was described in Cea et. al [3].

The exploration and introduction of three dimensional devices such as Tri-Gates [4,5] made routine 3D process and

978-1-6654-0685-7/21/$31.00 ©2021 IEEE

device modeling essential for advanced TD. This drove the development of 3D solid modeling based structure creation commands, a point cloud based interface to the Delaunay mesher deLink [6] and use of both MPI and thread based parallelization to reduce turnaround time [7,3]. Today the vast majority of all process simulation or structure creation jobs are 3D and with future architectures including nanowire/nanosheets[8] and forksheets[9] this trend is expected to continue and could always benefit from improved robustness, realistic structure representation/generation and turnaround times.

Stress engineering and modeling have been an integral part of device analysis for decades and became especially important when strain was used to engineer device performance [10,11]. FLOOPS includes models for stress due to misfit strain from doping or epitaxy, intrinsic film stress, thermal mismatch strain, strains from dislocations and stress due to material growth.[11,12] An overview of stress modeling applied to both NMOS and PMOS devices over many process generations is shown in Cea et al [13], while simulation was used to explain layout effects in SiGe devices [14]. DFT and Molecular Dynamics (MD) simulations are useful both for determining mechanical properties to use in continuum simulations as well as direct simulation of strain [3]. Atomistic process simulations can also be used to generate structures for atomistic transport. Voinov et al. [15] showed MD simulations used for Cu deposition on Ta to create realistic grain boundaries. These structures were used in atomistic NEGF transport simulations to estimate grain boundary scattering.

As mentioned earlier one future architecture option is nanowire or nanosheet transistors. Whenever transistor architecture options need to be evaluated one aspect is how can stress be used to engineer improved performance. This evaluation includes how the device transport will be affected by stress along with how the device can be stressed. Simulation can help with both questions. For silicon nanosheet devices the transport is well studied including the simulating the transport in [100] confined channels which changes the mobility response to stress from the [110] confined Tri-Gate devices [16]. SiGe epitaxial S/Ds have been the main source of stress in PMOS devices for generations. One concern for this stressor is if a defect free S/D be grown off independent wires and cavity spacers. Eneman [17] used stress simulations and estimated the impact of imperfect epi by using free surfaces in different orientations relative to the current flow direction. Figure 3 shows PMOS contours for stress along the current flow direction from full flow simulations for nanowire devices with perfect epi or with different dislocation locations and strengths. The dislocations are included numerically as in [12]. Explicitly including the dislocations allows matches to data along with predicting the impact with different layouts. There is a large delta in stress in the channel for different dislocation conditions and for the most tensile stress case the saturated drive currents are 45% degraded from the perfect epi case. This exposes one area of improvement for process simulation, which is predicting epitaxial film quality especially the defects that are expected to form. Missing this capability limits the predictiveness of stress simulations. Ideally atomistic methods could be used to predict the structure and the process changes that could be used to maximize epi quality. These simulations would be difficult requiring an atomistic method able to simulate larger structures, for example MD, and it would need forcefields that can accurately predict the interactions of many atoms including semiconductors (Si, Ge) and insulators (O, N). But this is a great opportunity to move beyond continuum process simulation.

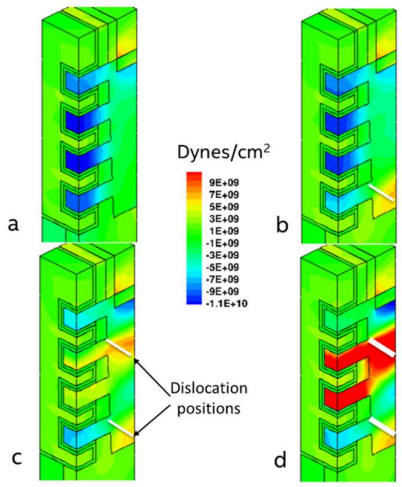

Figure 3. Stress simulation results showing the stress along the current flow direction for a) perfect epi b) one dislocation c) 2 dislocations and d) 2 dislocations with two times the strength (Burgers vector).

III. Device Simulation

Drift diffusion (DD) is the simulation method of choice for matching to experiments and process optimization due to its ease of adding phenomenological models to capture complex physics and ability to handle realistic structures [1]. Figure 1 shows how a wide variety of methods can be used to develop and calibrate enhanced DD models. Effective mass, K•P, or tight binding based Schrodinger-Poisson solvers can be used for quantum confinement models. Band structures calculated with effective mass, K•P or tight binding methods along with Kubo-Greenwood based mobility calculators are used to calibrate mobility models and any mass-based terms needed for ballistic mobility models. Monte Carlo based simulators (which are also informed by band structure calculations) are also used to calibrate high field transport parameters, for example saturation velocity or drain bias dependence on saturation velocity. As you can see there is a lot of fundamental inputs to the transport in the channel for the transistors. When used with calibrated process simulations for doping and stress profiles very good fits can be obtained [1]. While there is a lot of fundamental simulation informing the channel there is less fundamental understanding in the S/D extensions. When calibrating to data it is found that the conductivity in the S/D extensions must be degraded. Id vs Vg profiles with and without this degradation is shown in Figure 4. There are several reasons the tip might be degraded including (but not limited to) extra interface roughness, fixed charge under the spacers or dielectric confinement [18]. This uncertainty about the tip presents an opportunity for more fundamental modeling of the transport under the spacer. Understanding if this source of degradation can be removed represents a significant opportunity to improve device performance.

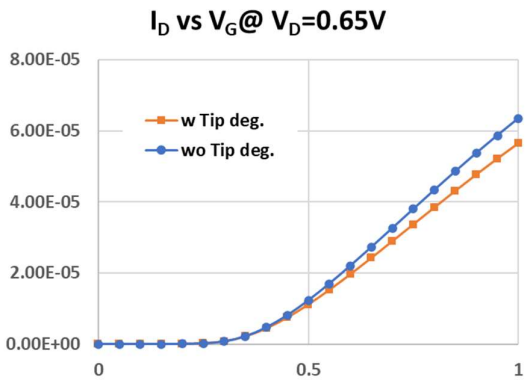

Figure 4. Simulated Id-Vg at Vd =0.65 with tip degradation included and with it turned off.

Figure 1 also shows that all device simulators in the hierarchy can be used individually and together to inform technology development. Each tool has strengths and weaknesses and in an industrial TCAD department the best tool(s) for the job should be used.

Atomistic NEGF simulations do an excellent job of capturing quantum effects, providing the upper limits of device performance in the ballistic limit, and can be used to model tunneling currents. Stettler et al [1] demonstrated how atomistic NEGF is the best tool to simulate leakage in devices where band-to-band induced barrier lowering is dominant using NEMO5 [18]. In indirect semiconductors scattering needs to be included to accurately capture this effect. One drawback of atomistic NEGF is that it is very CPU intensive. Using the low rank approximation (LRA) approach described in [20], results in a much smaller basis set that uses the minimum set of wave functions needed to accurately reproduce the band structure in the E-k regions of interest. Figure 5 shows the speed up achieved by using LRA vs full rank. The speedups are most on nanowire geometries which is fortunate since these simulations are the most CPU intensive.

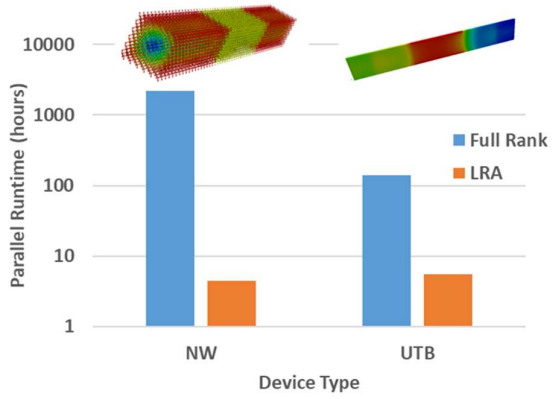

Figure 5. Total parallel CPU time in hours for full rank and LRA simulations of both nanowire (~530x speedup) and ultrathin body (~25x speedup) simulations.

Semi-classical simulations like Monte Carlo are well suited for device exploration because of the more natural addition of scattering mechanisms. This allows for simulations of quasi-ballistic transport which is important for evaluating novel materials versus silicon since each material will have different carrier scattering and amounts of ballisticity. When evaluating device options typically all levels of the hierarchy are needed and the more their areas of applicability and overlap can be extended provides value. Adding scattering to NEGF, better capturing band structure effects in Monte Carlo and extending both to simulate more realistic structures will enable better device evaluations.

Since atomistic, semi-classical and DD simulations often need to be run on the same structures to enable calibration we have extended our process simulator to also generate atomic representations of a structures. Figure 6 shows the flow for generating both atom and element meshes. There are several benefits of this approach. First, it is easy to guarantee that all device simulators are running on the same geometry. Second, there is more precise control of non-semiconductor material thicknesses than typically available in atomistic simulators, it is easier to add more complex doping profiles and generate more realistic geometries. In atomistic NEGF simulations more realistic doping and geometry of materials surrounding the semiconductor are important. Figure 7 shows a very idealized Ge UTB structure with a 4.4 nm thick body and a more realistic 4.4nm Ge UTB that includes a larger gate height. Ballistic simulations were performed using atomistic NEGF with the LRA approximation. Id vs gate bias results are shown in Figure 8. In the on state the tall gate device has better control of the tip or source drain extension region which results in higher current. Figure 7 also shows the potential profile differences between the devices.

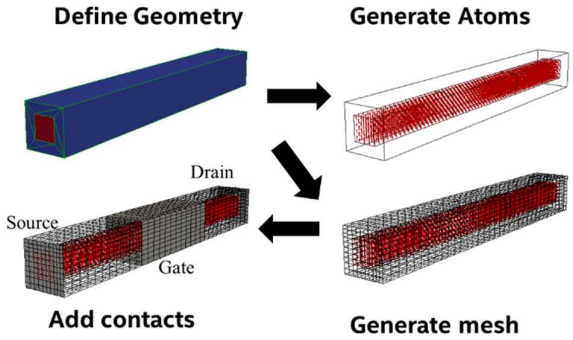

Figure 6. Extending process simulation/structure generation to include generation of both atomic and FEM meshes. Once meshes are generated the contacts and doping are added.

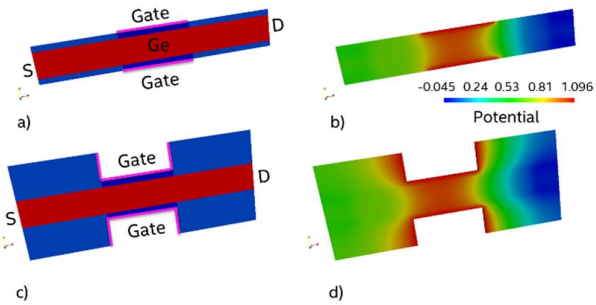

Figure 7. Ge UTB structures and potential profiles for an ideal case (a,b) and more realistic case (c,d) with gate contact on the side of the spacers.

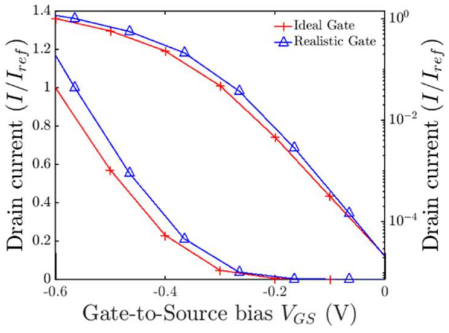

Figure 8. Id vs gate bias results for idealized vs realistic structures from Figure 7 relative to current for ideal gate at V_{gs}=-0.6, which show the impact of more realistic structures on NEGF results.

IV. Metal Resistance Simulation

As dimensions scale both line and via resistance increase and simulation can be used to both understand and evaluate different options. As mentioned earlier, atomistic structure creation and DFTB based NEGF simulations of transport were used to evaluate grain boundary scattering [15]. Figure 9 shows a system for predicting via resistances. DFT simulations are used to generate an interface with lowest energy state, while DFT+NEGF simulations are used to calculate the current for the structure. Given the contact area a resistivity can be extracted. The bulk resistivity can be taken from experiment or be calculated [15]. Via resistance for both conventional metal systems vs two novel metal options is shown in Figure 10. In these results interface resistivity was determined from simulations while bulk resistivity vs size was taken from experiment.

V. Conclusion

An overview of the diverse set of simulation methods and tools used in industrial process and device TCAD was presented. Predictive epi S/D quality simulations, fundamental modeling of tip conductivity and the ability to simulate more realistic structures with MC and atomistic device simulation were identified as areas where improvement would help TCAD more accurately evaluate future device options.

References

[1] M. Stettler et al, "State-of-the-art TCAD: 25 years ago and today," IEEE Internaional Electron Device Meeting (IEDM), 2019, pp. 39.1.1-39.1.4.

[2] M.E. Law and S.M. Cea, "Continuum based modeling of silicon integrated circuit processing: An object oriented approach," Comp. Mat. Sci. 12, 289-308, (1998).

[3] S. M. Cea et al., "Process modeling for advanced device technologies". J Comput Electron 13, 18–32 (2014).

[4] J. Kavalieros et al., "Tri-Gate Transistor Architecture with High-k Gate Dielectrics, Metal Gates and Strain Engineering," 2006 Symposium on VLSI Technology, 2006. Digest of Technical Papers., 2006, pp. 50-51.

[5] C. Auth et al., "A 22nm high performance and low-power CMOS technology featuring fully-depleted tri-gate transistors, self-aligned contacts and high density MIM capacitors," 2012 Symposium on VLSI Technology (VLSIT), 2012, pp. 131-132

[6] P. Fleischmann and S. Selberherr, "Enhanced advancing front Delaunay meshing in TCAD," International Conference on Simulation of Semiconductor Processes and Devices, 2002, pp. 99-102

[7] S, M. Cea et al. "Challenges in 3D Process Simulation for Advanced Technology Understanding". In: Grasser T., Selberherr S. (eds) Simulation of Semiconductor Processes and Devices 2007. Springer, Vienna.

[8] H. Mertens et al., "Vertically stacked gate-all-around Si nanowire transistors: Key Process Optimizations and Ring Oscillator Demonstration," 2017 IEEE International Electron Devices Meeting (IEDM), 2017, pp. 37.4.1-37.4.4.

[9] P. Weckx et al., "Stacked nanosheet fork architecture for SRAM design and device co-optimization toward 3nm," 2017 IEEE International Electron Devices Meeting (IEDM), 2017, pp. 20.5.1-20.5.4

[10] T. Ghani et al., "A 90nm high volume manufacturing logic technology featuring novel 45nm gate length strained silicon CMOS transistors," *IEEE International Electron Devices Meeting 2003*, 2003, pp. 11.6.1-11.6.3,

[11] S. M. Cea et al., "Strain Modeling in Advanced MOSFET Devices". ECS Transactions, 3(7), pp. 429-438 (2006)

[12] C. E. Weber, S. M. Cea, H. Deshpande, O. Golonzka and M. Y. Liu, "Modeling of NMOS performance gains from edge dislocation stress," 2011 International Electron Devices Meeting, 2011, pp. 34.4.1-34.4.4.

[13] S. Cea et al., "Strain Modeling in Advanced MOSFET Devices," ECS Meeting Abstracts MA2020-02(24): pp. 1741-1741.

[14] M. A. Stettler et al., "Industrial TCAD: Modeling Atoms to Chips," in IEEE Transactions on Electron Devices, doi: 10.1109/TED.2021.3076976.

[15] B. Voinov et al., "Material Modeling in Semiconductor Process Applications," Journal of Microelectronic Manufacturing, vol. 3 iss. 4, (2020).

[16] P. Packan et al., "High performance Hi-K + metal gate strain enhanced transistors on (110) silicon," 2008 IEEE International Electron Devices Meeting, 2008, pp. 1-4,

[17] G. Eneman, et al., "Stress Simulations of Fins, Wires and Nanosheets", in *ECS Trans.*, Vol. 98, no. 5, pp. 253-265, 2020.

[18] Björk et al., "Donor deactivation in silicon nanostructures", Nature Nanotechnology, 2009, 4, 103-107

[19] G. Klimeck, F. Oyafuso, T.Boykin, et al., "Development of a nanoelectronic 3-D (NEMO 3-D) simulator for multimillion atom simulations and application to alloyed quantum dots," CMES (2002)

[20] G. Mil'nikov, N. Mori, Y. Kamakura, "Equivalent transport models in atomistic quantum wires," Phys. Rev. B (2012).

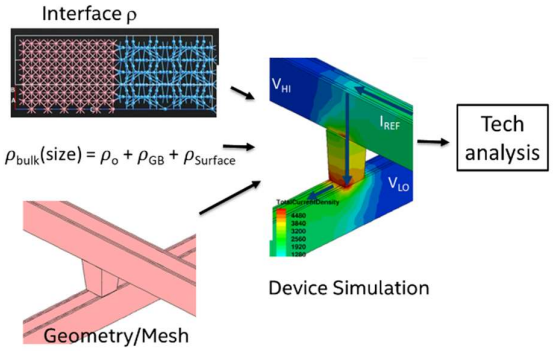

Figure 9. Simulation infrastructure used to estimate Via resistance including interface resistivity, bulk resistivity as a function of size and geometry.

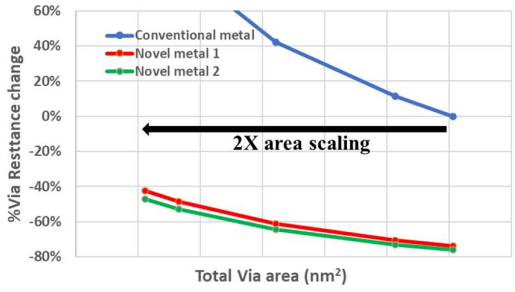

Figure 10. Simulation of via resistances vs via area for novel metal systems compared to conventional metal.