Edited by H. Ryssel, P. Pichler - September 1995

Simulating Deep Sub-Micron Technologies: An Industrial Perspective

P. Packan

Intel Corporation,
5200 NE Elam Young Pkwy,
Hillsboro, Oregon, 97124-6497

Abstract

Meeting the performance goals necessary to be competitive in the semiconductor industry will force novel process and device designs to be evaluated and optimized. Process and device simulators can be a valuable tool in the evaluation and optimization process. As device dimensions approach the 0.10 μm regime, device and process simulators will be pushed to new levels. In device simulations, non-local hot electron effects and mobility modeling are crucial for predictive simulations. The shallow, highly doped junctions required to improve short channel effects forces accurate diffusion models for extremely low energy implants as well as predictive modeling of extended defect interactions. Current areas of application for device and process simulation tools including development, optimization and manufacturing are discussed.

1. Introduction

It is becoming increasingly difficult to maintain the current pace of transistor performance improvement. Not only are these improvements expected, but they are expected with a concomitant decrease in power consumption. As transistor gate lengths of MOS devices approach the 0.10 μm regime, short channel effects are extremely difficult to control. In order to meet these challenges, aggressive scaling options will be required. It is not obvious which, if any, of the current scaling options will provide the best means of meeting the performance goals of the next generation devices. Because of these uncertainties, many possible solutions need to be evaluated. However, due to time and material constraints, it is not always possible to evaluate all options. A wide range of inputs are used to focus resources in the most promising directions. Increasingly, process and device simulation tools are being used to assist in this process. These simulators are being used in a variety of ways from technology evaluations to split lot design and optimization. However, all of these applications share one thing in common: they are critically dependent on the accuracy of the simulation tools.

Development of a new technology goes through a number of distinctive phases. In the initial phase, there are a number of process and device design options from which to choose. Once this is done, the second phase, optimization, is entered. In this phase, the process flow and device characteristics are refined and optimized to attain the best device performance. Finally, the technology is ready for manufacturing. In this

phase, impacts of manufacturing changes on performance and yield are evaluated. In addition, simulation tools can be used to quickly determine causes of process excursions. In each of these phases, process and device simulation tools can play a key role in decreasing the time to develop and maintain a new device process.

2. Architecture Development

During the initial stages of technology development, architectural options for both the process and device must be evaluated. Methods to suppress short channel effects (retrograde wells, halo implants, etc.), isolation schemes, metalization, etc., must all be addressed. As expected, the starting point for this evaluation is the previous generation technology. It may be that a clear direction for scaling the previous technology exists. Decreasing implant energies, reducing thermal cycles and printing smaller gate dimensions may be the simplest and most direct path to the next generation technology. However, as device dimensions continue to scale, this approach may not be feasible. Effects that were second and third order in older generations may now dominate transistor performance. Transient enhanced diffusion, implant channeling, salicide resistance, velocity saturation, etc. may play such a large part in determining transistor performance that conventional scaling may not be possible. In this case, new architectural options must be evaluated.

During this early phase of technology development, simulation tools can have their largest impact. At this time, absolute accuracy is not as critical as the ability to accurately assess the relative trade-offs between different device and process options. This requires a more global calibration, even though local accuracy may be sacrificed. Since there may be a wide range of device and process options, models must be physically based to account as much as possible for unforeseen interactions. These simulations can then be used to identify the most promising options for experiments in silicon.

The push in the semiconductor industry for MOS devices is toward shallow, highly doped S/D extensions (tips) with deeper S/D's for contacting the device [1],[2]. The reason for this is two-fold. First, shallow junctions decrease short channel effects by decreasing charge sharing effects. This allows the gate lengths to be decreased without undo source-drain leakage. Second, highly doped, abrupt profiles are needed to decrease the external resistance or at least to offset any resistance increase due to the shallower junctions.

Perhaps the most straightforward way to decrease the junction depths is to scale the implant energies. This is a methodology adopted by many companies and explains the current push to implant energies below 5 keV for both n and p-type dopants [3]. It is well known that TED is strongly dependent on the dopant species and implant dose [4]-[6]. In addition, at these extremely low energies, there is also a strong energy dependence that can differ significantly for different dopant species. The diffusion in most processes currently used in industry are dominated by TED. Therefore, simulation of trade-offs in junction depth, lateral underdiffusion, resistance and ultimately drive currents depend on the accuracy of these TED models.

Figure 1 shows simulations of diffused boron profiles for 5e14 boron tip implants of 5, 10, 20 and 40 keV. These tip profiles were implanted through a screen oxide and underwent an anneal of 900 C for 10 min. The depths of the profiles at a concentration of 5e17 are shown in table 1. Also shown is the percent change in junction depth as the energy is scaled. It is seen that scaling the implant energy alone will not continue to lead to a constant scaling of the junction depth. It is becoming increasingly difficult

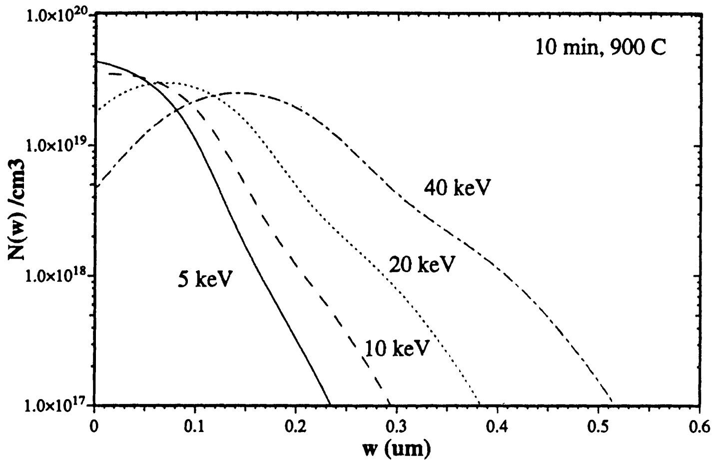

Figure 1: Simulated boron concentration profiles for 5e14 boron implanted at energies from 5 to 40 keV through a 10 nm oxide and annealed at 900°C for 10 min.

arsenic energy	Junction Depth (μm)		Scaling Percentage	
	900°C, 10 min	1000°C, 30 sec	900°C, 10 min	1000°C. 30 sec
5	0.17	0.16	19%	20%
10	0.24	0.22	28%	27%
20	0.32	0.30	28%	30%
40	0.45	0.43	—	—

Table 1: Effects of energy scaling on junction depth. Scaling the junction depth by decreasing the implant energy is becoming more difficult as the implant energies are reduced. TED and extended defects play a large role for this trend.

to scale tip junction depths merely by decreasing the energy. At these low energies, the improvement in the "as implanted" profile are being offset by TED effects.

A complicated interaction of implant and TED effects determines final junction depths. By changing the anneal conditions as well as the implant conditions, these scaling relationships can be altered. Table 1 also shows junction scaling factors if the anneal cycle is changed to 1000 C, 30 sec. For this case, the scaling factors vary even more. It will become increasingly difficult to scale junction depths and alternate doping techniques may need to be addressed.

These results become even more complicated when the damage from the source-drain implants is included. The interactions between point defects, extended defects and dopant atoms must all be taken into account for accurate, predictive process simulation. Even seemingly insignificant changes in thermal cycles can have large consequences on final device characteristics. Through the routine use of process and device simulations, these effects can be monitored to assure that any changes will not adversely affect device performance.

In conjunction with lowering implant energies, thermal cycles are being decreased to limit TED effects. It is commonly accepted that short time, high temperature anneals help to minimize TED [4]. Almost every major semiconductor company uses RTA processing in some capacity. Many companies are using very low temperature processing with strategically placed RTA anneals to help minimize TED. In these flows, the majority of dopant diffusion is controlled by these few RTA steps. Since enhanced diffusion effects due to TED are not localized, many process steps will

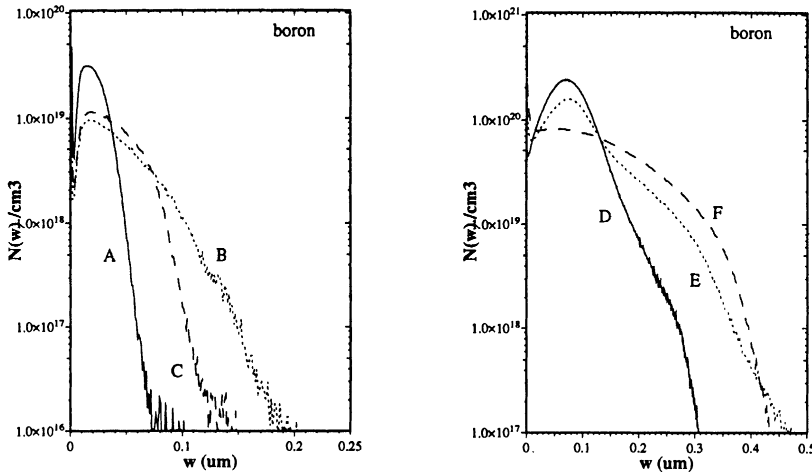

Figure 2: Boron profiles for a 1e14, 20 keV BF2 implant: A) as implanted, B) 800°C, 120 min anneal and C) 1000°C, 10 sec + 800°C, 120 min anneal. A silicon pre-amorphizing implant was used. Also shown are boron profiles for a 2e15, 20 keV boron implant: D) as implanted, E) 900°C, 10 min anneal and F) 1000°C, 30 sec + 900°C, 10 min. Although the addition of an RTA cycle will often help to minimize TED effects, this is not always the case. Extended defect-dopant interactions can strongly influence diffusion.

interact to define the final dopant profiles. Tip implants will affect channel profiles. S/D implants will affect tip profiles. It is critical to understand these interactions when developing an integrated process.

Figure 2 shows the effect of adding an RTA step immediately after ion-implantation. For low concentration implants (curves A,B,C), the addition of an RTA step can decrease dopant diffusion. For this experiment, a silicon pre-amorphization 1e15, 60 keV implant was used, followed by a 1e14, 20 keV BF2 implant. In one case, an 800°C, 120 min furnace anneal was performed. In the second case, a 1000°C, 10 sec anneal was performed prior to the 800°C furnace anneal. In this case, the RTA anneal reduced the total amount of diffusion. However, figure 2 also shows (curves D,E,F) that the addition of an RTA step immediately following an implant will not always lead to a reduction in junction depths. In this figure, the effect of inserting a 1000°C, 30 sec RTA step prior to a 900 C, 10 min anneal is shown. The boron was implanted at 20 keV to a dose of 2e15. For this case, the addition of an RTA step does not lead to a shallower junction. Extended defects and clustering effects interact to change both the TED dependency and mobile dopant concentrations. It is extremely difficult to simulate these effects using current models and can be done locally at best. For these high concentration profiles, modeling active concentrations is necessary, but difficult. As thermal cycles decrease, modeling the kinetics of activation and deactivation will be become even more important since the active concentrations determine device performance. This will mean that extended defect-dopant interactions will become even more critical as S/D extensions continue to increase in concentration and decrease in depth.

Velocity saturation and velocity overshoot are extremely important in very small MOS devices. NMOS devices, in particular, show a very strong velocity saturation effect in the sub-micron regime. Accurate evaluation of the improvement in device performance for different device options requires predictive simulation of these non-local hot carrier effects. If these effects are not taken into account, incorrect evaluations of de-

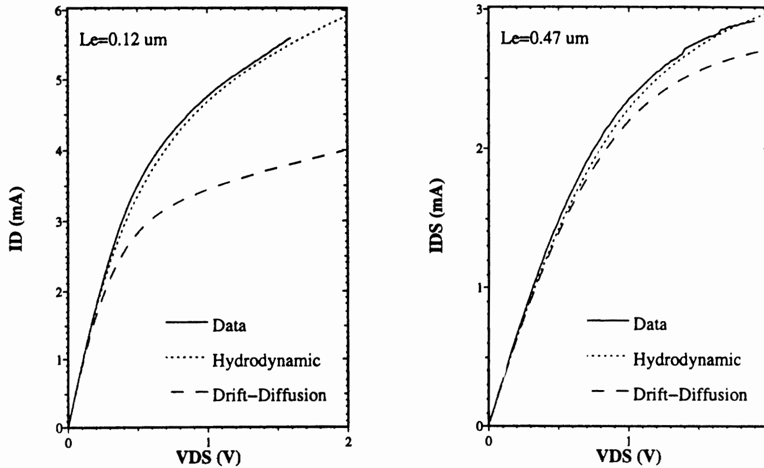

Figure 3: IV characteristics for a thick gate SOI structure. To accurately simulate these devices, non-local hot electron effects must be modeled. Using a hydrodynamic model with an energy relaxation time of 0.5 ps results in excellent matches.

vice improvement will be made. Hydrodynamic, Monte Carlo and other deterministic solution methods for solving the Boltzmann transport equation such as the scattering matrix approach and spherical harmonic expansion can capture non-local effects such as velocity overshoot [7]-[9]. The accuracy and ability to globally calibrate each of these methods is still debated.

Figure 3 shows data and simulated IV device characteristics for thick gate SOI structure developed by Assaderaghi, et. al. to evaluate velocity saturation and overshoot effects [10]. Two different size devices are shown. Drift diffusion simulations alone cannot be used to model both devices. Hydrodynamic simulations with an energy relaxation time of 0.5 ps match the data quite well. It should be noted that the value of the relaxation time is sensitive to the mobility model used. These results show that drift diffusion models can grossly underpredict saturation currents. When device architectural options such as halo implants and retrograde wells which strongly effect electric fields are evaluated, accurate simulation of these non-local effects are critical.

3. Architecture Optimization

Once an architecture has been determined, it must be optimized. During this portion of technology development, the simulation tools can be used to save time and silicon by effectively targeting split conditions for experiments. The simulations can also be used to better understand underlying physical mechanisms responsible for different phenomena. The requirements on accuracy of the simulations increase during the optimization phase. This may force locally calibrated models and may even necessitate the addition of new models to comprehend specific effects that may have been considered second order during the initial definition stage.

The optimization of one aspect of a trench isolation process can be used to illustrate these points. It has been reported that narrow channel effects for trench isolated devices can be very different than LOCOS isolated devices. For the LOCOS structure, gate edge lifting increases the gate oxide thickness at the trench edge. This leads to an increase in threshold voltages for narrow channel devices. However, the opposite trend has been observed for trench isolated devices, namely, a parasitic transistor

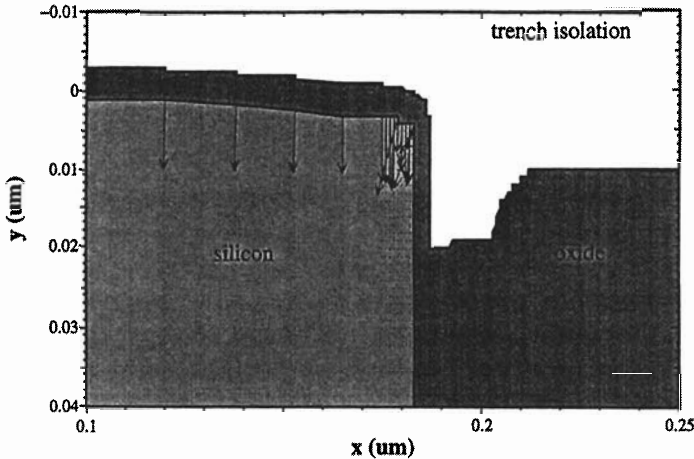

Figure 4: Electric field lines near the trench edge corner showing the importance of the corner geometry. High electric fields at the corner create a parasitic edge device.

with a decreased threshold voltage which can lead to a kink in the subthreshold IV characteristics [11],[12]. The origin of these effects were evaluated using process and device simulations. Figure 4 shows the electric field in the corner region of the device. As has been reported, the amount of curvature in this corner region strongly affects the electric field. This higher electric field will permit this edge region to invert at a lower voltage. Another important effect is seen in figure 5. This figure shows that dopant dose loss from the silicon substrate can occur during annealing of implanted dopants even though the surface is capped with an oxide. Figure 5 shows this effect for phosphorus. Dose loss is also seen for other dopants. The exact mechanism for this phenomenon is still under debate, but it is generally believed that a dopant rich interfacial layer forms between the silicon and the oxide. The dopant atoms in this layer are electrically inactive and can be removed by etching the oxide. At the trench corner, two dimensional effects increase the dose loss. This effect can be as important as the electric field effect. In addition to these effects, oxide edge thinning can play a role. Possible solutions to the narrow device Vt decrease include sidewall implants to increase the edge dopant concentrations and careful modification of sidewall oxidation conditions to change the edge geometry. The relative importance of the electric field crowding and dopant loss can be isolated and the appropriate solution determined through these types of simulations.

4. Manufacturing

After the device and process have been optimized in a development environment, it must be transferred to manufacturing. Simulations can aid in determining if the process conditions chosen are precariously close to a process window cliff. Response surface modeling can be used to evaluate the impact of process skews on device performance. These simulations can help target any problem areas where particularly tight control specifications may be needed. This requires extremely accurate simulations since the effect of very small changes are being evaluated.

The simulations can also be used to evaluate the impact of process changes that could make the process more manufacturable. Changes which affect throughput such as stabilization times and temperatures, ramp rates, etc. can be evaluated. However, these changes can adversely affect device performance and must be monitored.

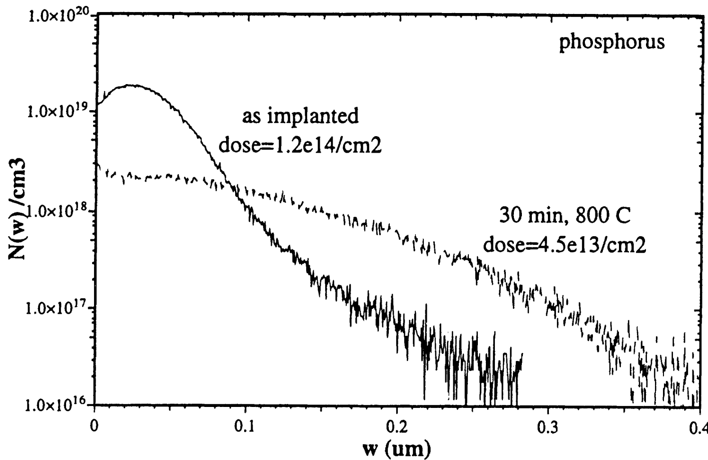

Figure 5: Phosphorus doping profiles showing extensive loss of dose during anneals in which TED effects dominate. This effect leads to a depletion of dopant at the trench edge.

Matching results from fabrication facility to fabrication facility can also benefit from simulations. Equipment and environmental differences must be compensated for and simulations can be used to identify possible solutions.

Knowledge base systems can be constructed to help to determine the cause for excursions in electrical characteristics. Electrical characteristics can be compared through the knowledge base systems and related back to particular process steps. Quick determination of the causes of process excursions is extremely valuable and because of this is drawing increasing attention as a priority area for simulation tools.

Dislocation loops caused by the interaction of point and extended defects are affected by implant and anneal conditions and can be strong yield limiters in a manufacturing environment. Relatively small changes in process conditions can strongly affect these extended defects and cause heavy leakage and even catastrophic electrical shorts. It is not possible to model these types of interactions in any global way at this time. However, the ability to model such effects would be extremely useful to manufacturing groups so that potential yield limiting process windows can be avoided.

5. Summary

Process and device simulations are used in a wide variety of ways within industry. Each application has specific requirements on accuracy, reproducibility and predictivity. As the cost of development continues to rise at an accelerated rate, the benefit of predictive process and device simulations increases. The trend in device design is toward shallow highly doped S/D extensions to decrease both resistance and short channel effects. To attain these extremely shallow highly doped junctions, implant energies are decreasing to the sub-5 keV range as doses increase to the 1e15 range. This puts added importance on low energy TED and extended defect interactions. Solubility and clustering kinetics determine the electrical profiles which are necessary for accurate device simulations. As the device dimensions continue to decrease, non-local hot carrier effects become very important for novel device evaluation.

There are still many interactions in the process and device simulations which are not well modeled. Extended defect formation and dopant interactions are locally modeled

with empirical models at best. Hydrodynamic and Monte Carlo device simulations focusing on non-local effects are not well calibrated and thus are not readily applied to problems of interest. However, as device dimensions continue to decrease, non-local effects must be accurately addressed to predictively assess and evaluate new technology options. Simulations currently provide an advantage in developing a technology generation. However, the promise of future impact is much larger if more predictive simulations become a reality.

References

[1] Y. Taur et al.,"High Performance 0.1μm CMOS Devices with 1.5 V Power Supply,"*IEDM Tech. Dig.*, pp. 127-130, 1993.

[2] T. Hori,"A 0.1μm CMOS Technology with Tilt-Implanted Punchthrough Stopper (TIPS),"*IEDM Tech. Dig.*, pp. 75-78, 1994.

[3] A. Hori et al.,"A 0.05μm CMOS with Ultra Shallow Source/Drain Junctions Fabricated by 5keV Ion Implantation and Rapid Thermal Annealing," *IEDM Tech. Dig.*, pp.485-488, 1994.

[4] P. Packan and J. Plummer,"Transient enhanced diffusion of low-concentration B in Si du to ^{29}Si implantation damage,"*Appl. Phys. Lett.*, vol. 56, pp. 1787-1789, 1990.

[5] S. Solmi et al.,"High-concentration diffusion in silicon: Simulation of the precipitation phenomena,"*J. Appl. Phys.*, vol. 68, no. 7, pp.3250-3258, 1990.

[6] P. Griffin et al.,"Species, Dose and Energy Dependence of Implant Induced Transient Enhanced Diffusion,"*IEDM Tech. Dig.*, pp. 295-298, 1993.

[7] C. Jacoboni and L. Reggiani,"The Monte Carlo method for the solution of charge transport in semiconductors with application to covalent materials," *Rev. Mod. Phys.*, vol. 55, pp645-705, 1983.

[8] N. Goldsman et al.,"A physics-based analytical/numerical solution to the Boltzman transport equation for use in device simulation," *Solid-State Elec.*, vol. 34, no. 4, 1991.

[9] A. Das and M. Lundstrom,"A scattering matrix approach to device simulation,"*Solid-State Elec.*, vol. 33, no. 10, pp. 1299-1307, 1990.

[10] F. Assaderaghi et al.,"Saturation Velocity and Velocity Overshoot of Inversion Layer Electrons and Holes,"*IEDM Tech. Dig.*, pp. 479-482, 1994.

[11] K. Ohe et al.,"Narrow Width Effects of Shallow Trench-Isolated CMOS with n+-Polysilicon Gate,"*IEEE Trans. Electron Devices*, Vol. 36, pp. 1110-1116, 1989.

[12] K. Ishimaru et al.,"Trench Isolation Technology with 1μm Depth n- and p-wells for a Full-CMOS SRAM Cell with a 0.4μm n+/p+ Spacing," *Symp. on VLSI Tech.*, pp. 97-98, 1994.