

SIMULATION BASED OPTIMIZATION ENVIRONMENT AND IT'S APPLICATION TO SEMICONDUCTOR DEVICES

R. PLASUN, M. STOCKINGER, R. STRASSER, and S. SELBERHERR

Institute for Microelectronics, TU Vienna
Gusshausstr. 27-29, A-1040 Vienna, Austria
Phone +43-1-58801-3692, FAX +43-1-5059224
e-mail: plasun@iue.tuwien.ac.at

ABSTRACT

We present a versatile simulation based optimization environment. Arbitrary simulators can be integrated into this environment like in our case for Technology Computer Aided Design (TCAD) applications. The optimizer module is integrated in a simulation environment which performs job farming and load balancing and also extracts results from simulation data.

In an example we present a doping profile optimization of an n-channel metal-oxide-semiconductor field-effect transistor (MOSFET). The target function of the optimization is a function of the electrical properties of the transistor and the space of the free parameters is spanned by several coefficients of an analytically specified two-dimensional doping profile.

KEYWORDS

Simulation, Parallel and distributed simulation, Computer Aided Design, VLSI, Optimization, Computer Aided Analysis.

INTRODUCTION

The design and fabrication of smaller and faster semiconductor devices relies on the proper numerical simulation of fabrication processes and electrical characteristics. To find those process parameters which give the best device performance, optimization of complete simulation flows have to be performed. Because of the large number of simulations fault tolerant mechanisms have to be to ensure stability of the environment. In addition independent simulations must be executed in parallel to reduce the overall simulation time. To support all these requirements a versatile optimization environment was developed based on the core functions of the Vienna Integrated System for TCAD Applications (VISTA) [1]. Because of the fact that arbitrary simulators can be easily integrated into this environment it is not restricted to microelectronics applications.

OPTIMIZER

The integrated optimizer is based on the nonlinear constrained optimizer package **donlpq2** [2]. This optimizer minimizes the target function with respect to the feasible region of the given problem. In the optimization module both the target function and the nonlinear constraints can be defined as arbitrary functions of the free parameters or of the extracted simulation results. The optimizer relies on the gradients of the given target function and constraints. These gradients are not directly available, because the input data of the optimizer are extracted from simulation results and there derivatives are not available. For this reason the gradients are calculated by finite differences, so for an optimization problem with n free parameters $n + 1$ evaluations have to be performed to obtain the gradients. To speedup these calculations asynchronous event handling is performed. This makes it possible to run the simulations for the gradient evaluation in parallel.

Figure 1 shows the evaluations versus time for a complete optimization process with the parallel gradient calculation and the sequential evaluations of the model function (the file handling will be described in section Application). An adaption of the finite-difference interval to the calculated gradient is done to increase the precision of the following gradient calculations.

To track the optimization process a graphical user interface is added to the optimization module. All parameters, the target function and the responses can be shown in a diagram.

FRAMEWORK

The optimizer module is integrated into the VISTA [3] framework which is a versatile environment that consists of several modules for job farming, handling and evaluating simulation flows, and graphical user interface support. The framework consists of several modules, which are coded in VLISP [4] a object-oriented LISP interpreter with an integrated binding to the Motif widget set.

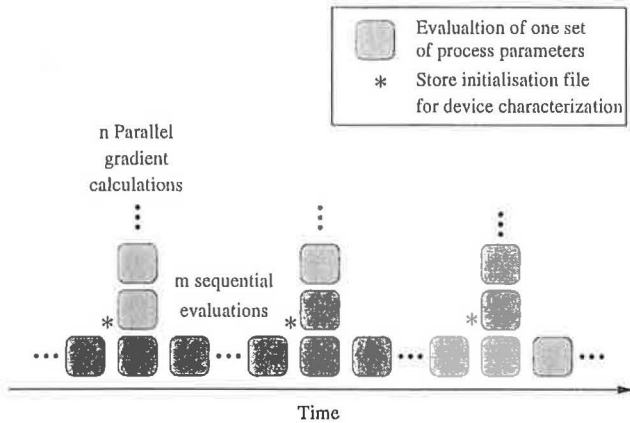


FIGURE 1: PARALLELIZED GRADIENT EVALUATION AND INITIALIZATION FILE HANDLING.

The VISTA framework is able to distribute the computation tasks to a heterogeneous cluster of workstations. In this job farming mechanism a dynamic load balancing method provides efficient usage of the available workstations to decrease the overall computation time. For this reason the framework monitors the workload of all the registered workstations by periodic load polling on each machine. With individual load limits, performance, and number of CPUs for each host a weight is calculated. The framework uses this ranking for the distribution of jobs. This ensures efficient resource utilization and it also avoids overloading the workstations.

In the Figure 2 a screen shot of the graphical user interface of the job farming and load balancing mechanism is shown. In the upper part of the panel the available hosts, the individual limits, and the current and predicted loads are shown. The lower part shows a list of the queue where the active and waiting tasks are marked with different colors.

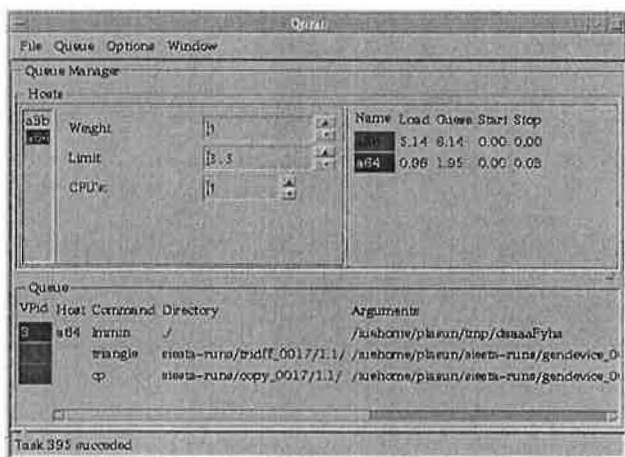


FIGURE 2: USER INTERFACE OF THE SIMULATION ENVIRONMENT.

The binding between the external optimizer program and the framework is done by an agent which is even able

to interact if the framework and the optimizer are running on different workstations.

The model function requested by the optimizer is in general defined as an arbitrary flow of simulation steps. The framework schedules the pending simulations, executes the steps when their input data becomes available, and performs result extraction from the simulator output. The evaluation of the target function and the constraints is also carried out by the framework.

In order to perform complex optimizations several thousand simulations have to be executed. However, during execution it is likely that an error occurs because of the large number of steps. There are many possible reasons for such errors like heavy loaded network, full temp-partitions, or convergence problems of the simulator. In the VISTA framework an error handling mechanism supports re-execution of failed tasks. These failed tasks are scheduled for re-computation and if the used simulator provides different solvers a more stable but in general slower algorithm is selected. This mechanism is very efficient and necessary for the stability of the optimization environment.

APPLICATION

To simulate the manufacturing process of a modern Very Large Scale Integration (VLSI) technology several simulators have been developed during the past decade. For example the channel and substrate doping profile of sub-micron transistors define the device performance. A general proof which doping profile is the best for some desired electrical characteristics cannot be given. Optimization of doping profiles offers a new chance to build better devices in the next century. For our application we use a specialized device generator, a tool for creating MOS devices with analytical doping profiles. These doping profiles are generated by superposition of arbitrary Gaussian function and two-dimensional splines.

The electrical parameters of the devices, like voltage over current curves, are extracted using MINIMOS-NT [5] a general purpose device simulator. For the gradient evaluations which are performed in parallel to speed up the overall computation time a very powerful feature of MINIMOS-NT is used. An initialization file can be specified to improve the convergence behavior, so the simulator needs only a few iterations to find a solution. In Figure 1 the structure of the optimization process with the used method for the gradient calculations and initialization file handling is visualized.

With these two tools we build a simplified transistor model with the parameters of the analytical doping profile on input and the electrical characteristic as output quantities.

OPTIMIZATION RESULTS

As an example an n-MOS device with $0.25\ \mu\text{m}$ geometry gate length and $5\ \text{nm}$ gate oxide thickness is optimized for $1.5\ \text{V}$ supply voltage. The goal of the optimization is to find a doping profile for this specified device structure to achieve certain performance improvements. This doping profile is generated by a device generator. The computation time required for the generation of the doping profiles is short, compared to a complete manufacturing process simulation.

The profile of the device consists of Gauss-functions and two-dimensional splines. The source/drain regions consist of Gaussian-functions of donors is shown in Figure 3.

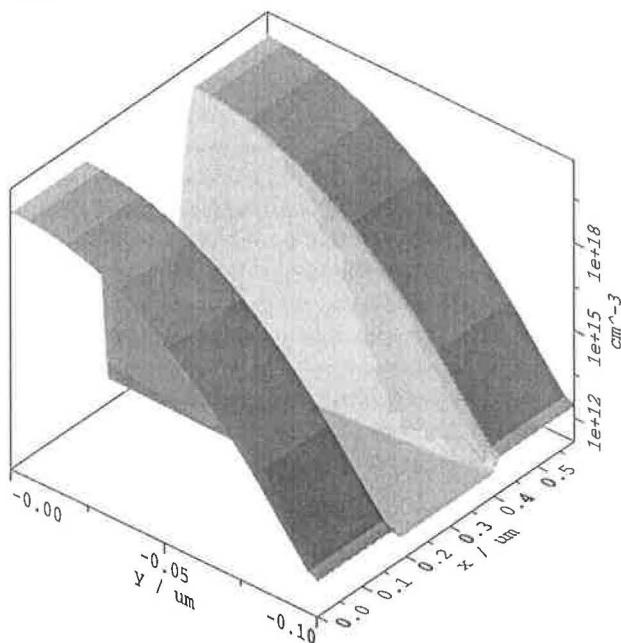


FIGURE 3: DONOR DOPING PROFILE OF THE MOS TRANSISTOR.

The channel profile created by two-dimensional splines where Figure 4 shows the 62 discretization points of the channel doping as the cross-sections of the drawn mesh. The acceptor doping profile generated on this mesh of the initial device is shown in Figure 5.

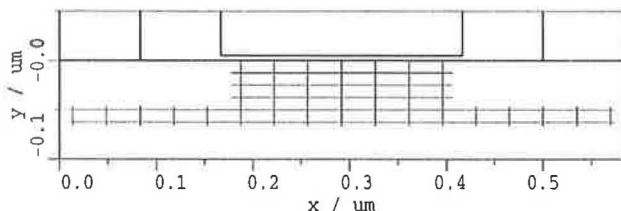


FIGURE 4: DISCRETIZATION OF THE SPLINE POINTS IN THE CHANNEL REGION OF THE N-MOS TRANSISTOR.

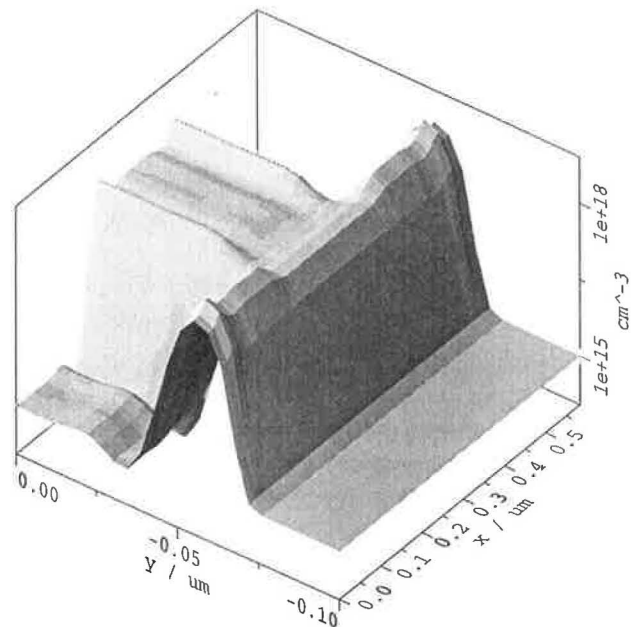


FIGURE 5: INITIAL ANALYTICAL ACCEPTOR DOPING PROFILE OF THE N-MOS TRANSISTOR.

The optimization target is defined for achieving maximum drive-current. Without a constraint, this optimization would result in a decrease of the threshold voltage, and the leakage-current would be drastically increased. Therefore, the leakage-current is entered as a constraint for the optimization process and kept below constant value.

To extract the drive- and leakage-current two device simulation steps have to be performed. The VISTA framework distributes the simulation jobs over a cluster of workstations so they can be done in parallel. The control variables of the optimization problem are mapped by the framework to specified parameters in the simulation flow. The framework also collects the extracted responses and evaluates the target function.

Figure 6 shows the minimization of the target function versus the number of main steps, consisting of several single steps and one gradient calculation. This doping profile optimization with these 62 parameters required 3249 simulation steps. The knee between the 10th and 45th main step is by the active constraint. During these steps the optimizer moved along the boundary of the feasible region where the leakage-current is below a given value.

Figure 7 shows the resulting acceptor doping profile of the optimized transistor. With this highly asymmetric acceptor doping profile the drive-current was increased by 50% compared to the initial device with the same off current.

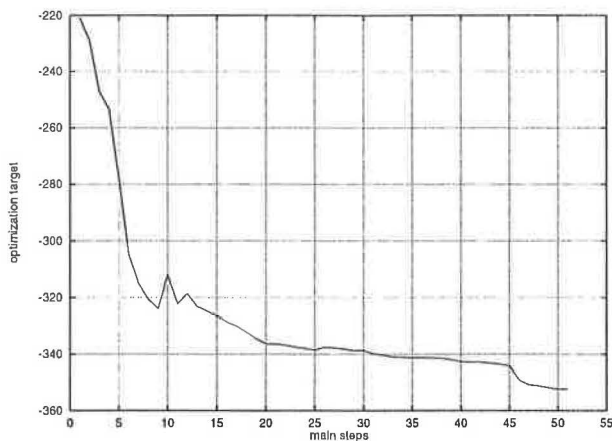


FIGURE 6: MINIMIZATION OF THE TARGET FUNCTION VERSUS THE NUMBER OF MAIN STEPS.

CONCLUSION

We presented an universal environment for automated optimization. In this environment parallel execution of tasks is performed to reduce computation time which is very important for the gradient evaluations. Sophisticated management of initialization data increases the simulator's convergence and speeds up the simulations. In an example of a microelectronics application the efficiency and robustness of this framework was shown. Based on this environment with the flexible framework features, tasks like reverse engineering of doping profiles can be realized.

ACKNOWLEDGMENT

This work was significantly supported by the "Christian Doppler Forschungsgesellschaft", Vienna, Austria and Austria Mikro Systeme, Unterpremstätten, Austria.

REFERENCES

- [1] C. Pichler, R. Plasun, R. Strasser, and S. Selberherr, "High-Level TCAD Task Representation and Automation," *IEEE Trans.Semiconductor Technology Modeling and Simulation*, May 1997. <http://www.ieee.org/journal/tcad/accepted/pichler-may97/>.
- [2] P. Spellucci, "Solving General Convex QP Problems via an Exact Quadratic Augmented Lagrangian with Bound Constraints." <http://www.mathematik.th-darmstadt.de/ags/ag8/spellucci>, 6 1996.
- [3] S. Halama, F. Fasching, C. Fischer, H. Kosina, E. Leitner, C. Pichler, H. Pimingstorfer, H. Puchner, G. Rieger, G. Schrom, T. Simlinger, M. Stiftinger, H. Stippel, E. Strasser, W. Tuppa, K. Wimmer, and

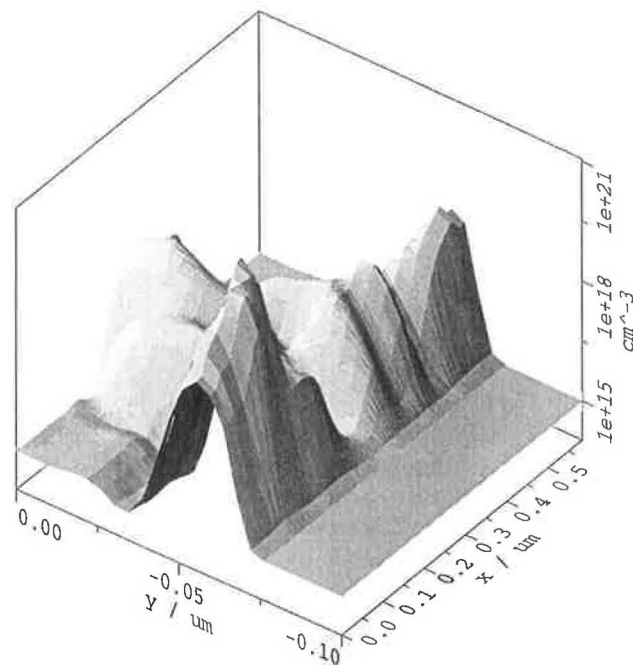


FIGURE 7: ANALYTICAL ACCEPTOR DOPING PROFILE OF THE OPTIMIZED N-MOS TRANSISTOR.

S. Selberherr, "The Viennese Integrated System for Technology CAD Applications," in *Technology CAD Systems* (F. Fasching, S. Halama, and S. Selberherr, eds.), pp. 197–236, Springer, 1993.

- [4] Institut für Mikroelektronik, Technische Universität Wien, Austria, *VISTA Documentation 1.3-1, VLISP Manual*, Jan. 1996.
- [5] T. Simlinger, H. Kosina, M. Rottinger, and S. Selberherr, "MINIMOS-NT: A Generic Simulator for Complex Semiconductor Devices," in *25th European Solid State Device Research Conference* (H. de Graaff and H. van Kranenburg, eds.), pp. 83–86, Editions Frontières, 1995.

BIOGRAPHY

Richard Plasun was born in Vienna, Austria, in 1969. He studied electrical engineering at the Technical University of Vienna, where he received the degree of 'Diplomingenieur' in 1994. He joined the 'Institut für Mikroelektronik' in March 1995. He is currently working towards his doctoral degree. From October to December 1996, he was with the National Semiconductor Cooperation in Santa Clara, where he concentrated on tool binding and process optimization aspects. His scientific interests include linear and nonlinear optimization, technology CAD framework aspects and software development.