Distributed Algorithms for Three-dimensional Semiconductor Device Simulations

MEI-KEI IEONG* and TING-WEI TANG†

Department of Electrical and Computer Engineering, University of Massachusetts, Amherst, MA 01003

We describe a newly developed parallel three-dimensional semiconductor device simulator. The program utilizes the message passing architecture and has been developed on a network of workstations. Two iterative algorithms to solve the nonlinear semiconductor device equations are presented. The algorithms are tested by simulating a three-dimensional MOSFET structure. A near-ideal speedup for the distributed algorithms has been obtained for Poisson's equation.

Keywords: Three-dimensional Device Simulation, Domain Decomposition, Parallel Algorithms, Message Passing, Iterative Methods

INTRODUCTION

In device simulation, certain physical phenomena cannot be modeled accurately in a two-dimensional domain. However, a three-dimensional device simulation generally requires significant computation time and memory storage. Parallel processing makes it possible to tackle such large-scale problems. In [1] and [2], stationary iterative method such as Gauss-Seidel and Successive Over Relaxation (SOR) were used to solve the linear equations. These methods are easy to understand and parallelize. However, they are computationally less effective in solving nonsymmetric linear systems when compared to nonstationary iterative methods such as Conjugate Gradient Squared (CGS) and Generalized Minimum Residual. In [3] and [4], various massively parallel algorithms

using direct and nonstationary iterative methods have been developed. On the other hand, the coarse grain parallelism using a network of high performance workstations has attracted more and more attention recently because of its higher performance-to-cost ratio. Coughran et al. [5] have reported some preliminary results using the domain decomposition method on a network of workstations connected via a high-speed network interface. In this work, we study the performance of a similar *coarse grained parallelism* for three-dimensional device simulations but using an ordinary Ethernet cluster of DEC-stations. Because of differences in network latency, new parallelization strategy must be devised for the overall optimum performance. Two iterative algorithms based on domain decomposition have been studied and their performances are reported here.

* E-mail: ieong@khushi.ecs.umass.edu

† E-mail: ttang@ecs.umass.edu

GOVERNING EQUATIONS AND DISCRETIZATION

A commonly adopted set of scaled semiconductor equations, with the drift-diffusion approximation and in the steady state, takes the following form [6] :

$$\lambda^2 \nabla^2 \psi = n - p - C. \quad (1)$$

$$\nabla \cdot J_n = G, \quad (2)$$

$$J_n = -n\mu_n \nabla \psi + D_n \nabla n, \quad (3)$$

where, ψ , n , p , C , J_n , G , μ_n and D_n are the *scaled* electrostatic potential, electron concentration, hole concentration, total doping density, electron current density, total generation and recombination rate, electron mobility and electron diffusivity, respectively. The continuity equation is discretized on a three-dimensional nonuniform rectangular (tensor product) grid using the control-volume and Scharfetter-Gummel [7] methods.

SOLUTION METHODS

Due to the memory constraint, the Gummel decoupled method [8] in which the Poisson and continuity equations are solved iteratively, is more appropriate for three-dimensional problems. The parallelization is based on the Domain Decomposition approach which partitions the domain of physical problem into a number of smaller subdomains. Since the subdomains can be handled independently, such methods are very attractive for coarse-grain parallel computers. Domain decomposition methods can be classified as either overlapping (also called Schwarz) or nonoverlapping methods [9]. The overlapping methods may be viewed as generalizations of block Jacobi (also called additive Schwarz) and block Gauss-Seidel (also called multiplicative Schwarz) methods. The block Jacobi method provides a higher parallelism and has been implemented in our program. In each subdomain, we need to solve the Poisson and continuity equations. The homogeneous Dirichlet boundary conditions are imposed on the artificial boundaries and original boundary conditions are imposed on the

true boundary. In general, these equations are nonlinear and a number of iterative techniques can be used to solve these nonlinear equations. The commonly used algorithm is the Newton-Raphson method [6] which linearizes the system of nonlinear equations to obtain a Jacobian matrix. The associated Jacobian matrix is then solved by a linear solver to give a correction vector for the variables. This algorithm is quite effective but the main drawback is the extra memory requirement for the large Jacobian matrix. An alternative way is to solve each of the nonlinear equations explicitly using a fixed point method [10], thereby avoiding the generation of a large Jacobian matrix. This algorithm also has an advantage of being easily parallelized.

Two iterative algorithms for the solution of nonlinear equations in each subdomain have been implemented. The first is the Newton-Raphson method and the linear system is solved by the Bi-CGSTAB [11] algorithm with ILU preconditioner. The global problem is solved by the block Jacobi method. We refer this algorithm as the Jacobi Bi-CGSTAB (JBi-CGSTAB) method. The second one solves the local problem by a nonlinear Gauss-Seidel with red-black ordering while the global problem is also solved by the block Jacobi algorithm. This scheme is referred to as the Jacobi nonlinear Gauss-Seidel (JNG-S) method.

These two algorithms have been implemented on a network of workstations connected via Ethernet. The program partitions the physical domain according to their coordinates along the axis that has the most number of grid points. The LPARX [12] C++ class library has been adapted to simplify the data distribution. The inter-process communication is accomplished by a popular public domain message passing library, Parallel Virtual Machine (PVM) [13]. The programming language used is mainly C++ with some routines written in other programming languages such as C and Fortran. This particular approach exploits the respective strength of various programming languages, and at the same time, any optimized numerical library can be easily incorporated into the code [12].

FIGURE 1 Distribution of Electron Current Density Vectors

SIMULATION RESULTS

We have simulated a three-dimensional submicron N-MOSFET, with a dimension of $1\text{ }\mu\text{m} \times 1\text{ }\mu\text{m} \times 1\text{ }\mu\text{m}$, an oxide thickness of 25nm and a channel length of $0.4\text{ }\mu\text{m}$. The channel doping concentration is $1.7 \times 10^{17}\text{ cm}^{-3}$. To show the three-dimensional electron flow, we deliberately positioned the source and drain contacts as shown in Fig.1. The current density vector clearly indicates the three-dimensional feature of the result. With this simulation tool, the dependence of the parasitic resistance on the size and location of the source and drain contacts can be readily extracted. To show the performance of the distributed algorithms, we compare the total elapsed time versus number of processors in solving Poisson's equation. The bias condition is $V_g = 2.0$ volts and all the other terminals are grounded. Table 1 and Table 2, respectively, show the performance of the JBi-CGSTAB and JNG-S algorithms. As mentioned earlier, the JBi-CGSTAB algorithm requires more memory to store the Jacobian matrix. Also, the size of problem that can be solved by a single workstation without extensive memory swapping is limited by the size of the main memory of the local machine. These are indicated by a "*" in Table 1. The cluster is simultaneously shared by other users and some fluctuation in

the performance has been observed. Yet, we have obtained a near ideal speedup when the network is only lightly loaded.

TABLE I Total Elapsed Time in Seconds and speedup for a MOSFET simulation using JBi-CGSTAB algorithm on a cluster of DECstation5000. Here P is the number of processors

Size $NX \times NY \times NZ$	$P = 1$ Time/ Speedup	$P = 2$ Time/ Speedup	$P = 4$ Time/ Speedup	$P = 8$ Time/ Speedup
$22 \times 22 \times 30$	749/1.00	394/1.90	191/3.91	145.7/5.14
$22 \times 22 \times 50$	*	653/2.0	373/3.5	203/6.44
$22 \times 22 \times 100$	*	*	982/4.0	556/7.08

TABLE II Total Elapsed Time in Seconds and speedup for a MOSFET simulation using JNG-S algorithm on a cluster of DECstation5000. Here P is the number of processors.

Size $NX \times NY \times NZ$	$P = 1$ Time/ Speedup	$P = 2$ Time/ Speedup	$P = 4$ Time/ Speedup	$P = 8$ Time/ Speedup
$22 \times 22 \times 30$	307/1.00	156/1.97	80.2/3.83	69/4.45
$22 \times 22 \times 50$	663/1.00	340/1.95	187/3.55	114/5.82
$22 \times 22 \times 100$	3410/1.00	1740/1.96	917/3.72	576/6.02

DISCUSSION AND CONCLUSIONS

We have presented our newly developed distributed three-dimensional semiconductor device simulation program. The algorithm is based on the domain decomposition method. Two iterative algorithms in solving the nonlinear semiconductor device equations are tested by simulating a three-dimensional MOSFET structure. A very good speedup for both algorithms has been obtained in low bias conditions. However, as the drain voltage is increased, the number of iterations required to solve the problem also increases. The deterioration of the convergence rate is due to the fact that the drift-term in the current equation becomes more and more significant at high biases. It has been reported [14] that the block Jacobi method when applied to the convection-diffusion problem exhibits poor convergence properties for

convection-dominated flow. A domain decomposition preconditioning together with a conjugate-gradient type method may be more appropriate in this situation [5], [14]. This approach is currently being investigated. Another advantage of the networked computing is that very large problems that cannot be handled by a single workstation previously can now be solved in a relatively short time. Although the project is still in its preliminary stage, we see the distributed computing by workstation clusters, even with relatively slow network interface, as an attractive alternative for cost-effective multi-dimensional semiconductor device simulations.

References

- [1] C.S. Tsang-Ping, D.M. Barry and C.M. Snowden, "Parallel Implementation of A GaAs MESFET Electro-Thermal Simulation on a Transputer-based System", *The Proc. of the Third International Workshop on Computational Electronics*, pp. 266-269, 1994.
- [2] A. Asenov, A.R. Brown, S. Roy, and J.R. Barker, "Topologically Rectangular Grids in the Parallel Simulation of Semiconductor Devices", *This Issue of VLSI Design*, Fall 1996.
- [3] D. Webber, E. Tomacruz, R. Guerrier, T. Toyabe, and A. Sangiovanni-Vincentelli, "A massively Parallel Algorithm for Three Dimensional Device Simulation", *IEEE Trans. Computer-Aided Design of Integrated Circuits and Systems*, Vol.10, pp. 1201-1209, No.9, Sep 1991.
- [4] E. Tomacruz and A. Sangiovanni-Vincentelli, "A parallel Iterative Linear Solver for Solving Irregular Grid Semiconductor Device Matrices", *The Proc. of Supercomputing 94*, pp. 24-33, 1994.
- [5] W.M. Coughran, JR. and E. Grosse, "Domain Decomposition Applied to the Drift-Diffusion Equations", *The Proc. of the Third International Workshop on Computational Electronics*, pp. 247-247, 1994; P. E. Bjorstad, W.M. Coughran, JR. and E. Grosse, "Parallel Domain Decomposition Applied to Coupled Transport Equations", *The Proc. of Domain Decomposition Method 7*, 1993.
- [6] S. Selberherr, *Analysis and Simulation of Semiconductor Devices*, Springer-Verlag Wien New York, 1984.
- [7] D.L. Scharfetter and H.K. Gummel, "Large-signal analysis of a silicon Read diode oscillator", *IEEE Trans. on Electron Device*, Vol.ED-16, no.1, pp. 64-77, 1969.
- [8] H.K. Gummel, "A Self-Consistent Iterative Scheme for One-Dimensional Steady State Transistor Calculations", *IEEE Trans. on Electron Device*, Vol.ED-11, pp. 455-465, 1964.
- [9] B.F. Smith, "Domain Decomposition Methods for Partial Differential Equations", *the Proceedings of an ICASE/LaRC Workshop on Parallel Numerical Algorithms*, Kluwer Publishers, 1995.
- [10] C.E. Korman and I.D. Mayergoyz, "A Globally Convergent Algorithm for the Solution of the Steady-State Semiconductor Device Equations," *Journal of Applied Physics*, Vol.68, No.3 pp.1324-1334, 1990.
- [11] H. Van Der Vorst, "Bi-CGSTAB : A fast and smoothly converging variant of Bi-CG for the solution of nonsymmetric linear systems", *SIAM J. Sci. Statist. Comput.*, Vol.13, pp. 631-644, 1992.
- [12] S.B. Baden, S.R. Kohn, S.M. Figueira and S.J. Fink, *The LPARX User's Guide v2.0*, Department of Computer Science and Engineering, University of California, San Diego, 9 November, 1994.
- [13] A. Beguelin, J. Dongarra, A. Weist, R. Manchek, V. Sunderam, *User's Guide to PVM : Parallel Virtual Machine*, ORNL/TM-11826. Mathematical Science Section, Oak Ridge National Laboratory. Sep. 1991.
- [14] E. Brakkee and P. Wilders, "A Domain Decomposition method for the advection-diffusion equation", *Report 94-08*, Faculty of Technical Mathematics and Informatics, Delft University of Technology, Netherlands. 1994.

Biographies

Mei-Kei Ieong is with IBM SRDC, Hopewell Junction, N.Y. 12533, USA. He received Ph.D in Electrical and Computer Engineering from University of Massachusetts, Amherst. His research interests include simulation and modeling of semiconductor devices.

Ting-wei Tang is Professor of Electrical and Computer Engineering at University of Massachusetts, Amherst. His research interests include semiconductor device physics and numerical modeling.