

FEDSS—A 2D semiconductor fabrication process simulator

by L. Borucki
H. H. Hansen
K. Varahramyan

The main features of the finite element semiconductor process simulator FEDSS are described, with emphasis on a recently added capability for generalized 2D oxidation with impurity redistribution in oxide and silicon. Examples are given that demonstrate the ability of the program to oxidize various structures using a model based on steady-state oxidant diffusion and incompressible viscous oxide flow. Impurity profiles and contours are also shown in both neutral and oxidizing ambients, along with several comparisons with data or with the program SUPREM II.

Introduction

Field effect transistors (FETs) and bipolar transistors are commonly employed in making high-density semiconductor memories and high-speed logic circuits. The fabrication processes for these circuits have become increasingly complex as the level of integration has increased. Additionally, the life of any technology has shortened considerably, thus reducing the time available for the development of new and denser memories and logic circuits. As device dimensions have decreased, the interaction

©Copyright 1985 by International Business Machines Corporation. Copying in printed form for private use is permitted without payment of royalty provided that (1) each reproduction is done without alteration and (2) the *Journal* reference and IBM copyright notice are included on the first page. The title and abstract, but no other portions, of this paper may be copied or distributed royalty free without further permission by computer-based and other information-service systems. Permission to *republish* any other portion of this paper must be obtained from the Editor.

between process variables and device behavior has also increased—the most commonly cited examples being the short- and narrow-channel effects in field effect transistors. Previously, one-dimensional modeling of processes and devices was adequate. However, previously ignored 2D phenomena have now become vitally important. The net result of all of these changes and effects is that 2D process and device modeling is a necessary and accepted part of the design process.

The semiconductor process simulator FEDSS (Finite Element Diffusion Simulation System [1, 2]) was initially created several years ago from the software base underlying the device simulator FIELDAY (Finite Element Device Analysis program [3]). FEDSS was needed to create realistic 2D doping distributions for FIELDAY and also to do stand-alone process simulation. The environment in which FEDSS was to be used required that it be able to model a wide variety of bipolar and FET device fabrication steps, and so be as technology independent as possible. Like FIELDAY, FEDSS was based on the finite element method, which is used to solve the diffusion equation and other required equations. Presently, FEDSS is part of a simulation package based on the finite element method (**Figure 1**). Pre- and post-processors, data bases, and device programs are all portions of this package. To use the FEDSS portion, the user begins with a process menu and a sketch of the problem geometry. A mesh generation program called TRIM (TRIangular Mesh generator) is run to create the finite element mesh. FEDSS input is coded using the process description, then the simulation is run—essentially step by step as is done in the actual processing, except faster. After

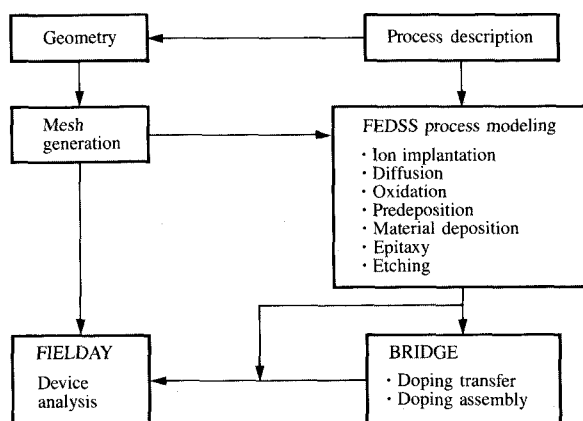


Figure 1

The finite element process and device analysis system, with particular emphasis on FEDSS.

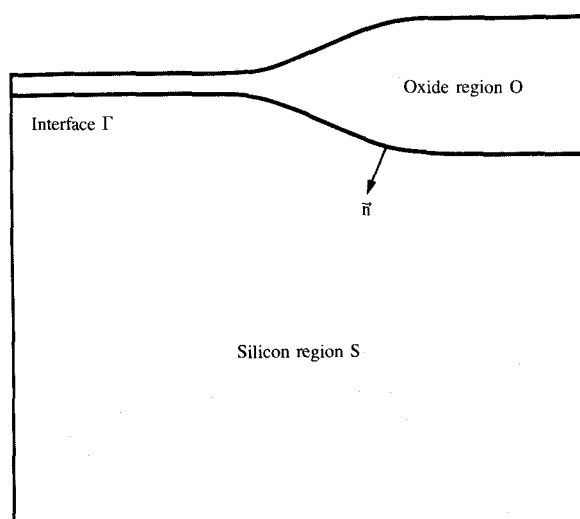


Figure 2

An idealized simulation region R , showing oxide and silicon sub-regions O and S , an oxide-silicon interface Γ , and the convention used for the interface normal $\vec{n}$.

completion of the run, post-processors enable the user to graphically view the results—cross sections, contours, and projections are all possible. Often it is necessary to make device calculations from these profiles. Usually process meshes are not suitable for device calculations, so a program called BRIDGE can be used to map FEDSS concentrations onto the FIELDAY mesh. Thus a suitable mesh can be used

for process calculations and a different mesh can be used for the device calculations. In cases where a device can be dissected into isolated components, BRIDGE can also be used to assemble high-resolution FEDSS simulations of the individual device pieces on the final FIELDAY mesh.

What is reported here is the latest version of FEDSS, which now offers significant enhancements over the original. The major accomplishment is the capability to model generalized 2D oxidation while simulating impurity redistribution in both oxide and silicon in the presence of a moving oxide-silicon interface. Even though there are a few other 2D process simulation programs that provide some oxidation capability, it is believed that the present oxidation-redistribution package in FEDSS is the most complete and generalized system of its kind. Specifically, the current program can model

- Ion implantation of any element by using 1D or 2D analytic models. Angled implants and irregular or planar surface geometries are allowed.
- 1D ion implantation of any element by using a Monte Carlo method.
- Diffusion of any element, with particular emphasis on arsenic, boron, and phosphorus. Segregation is modeled at interfaces between silicon dioxide and silicon or polysilicon. The diffusion of arsenic can be modeled either in equilibrium or kinetically, and the interaction of arsenic and boron is properly accounted for in the redistribution models.
- Oxidation and the diffusion phenomena related to it.
- Predeposition of impurities through an exposed wafer surface and evaporation through the same surface.
- Deposition of silicon dioxide, nitride, silicon, polysilicon, and up to three different user-defined materials.
- Etching (material deletion).
- Silicon epitaxy.

Numerical approach

The diffusion equation is the principal equation solved by FEDSS in any simulation of impurity migration during thermal processing, although chemical reaction equations and fluid flow equations are sometimes also solved. Before describing any of the more complex physical models used in FEDSS, the numerical technique adopted to solve model equations is illustrated in the case of a single impurity diffusing in oxide and silicon with segregation at a static interface. (Figure 2 shows an oxide region O , silicon region S , and interface Γ for a typical "bird's-beak" problem.) Static boundary segregation is a significant numerical problem by itself and is one of the important components in the FEDSS simulation of impurity redistribution during oxidation. The approach used in this problem will also help to clarify the numerical methods used in some of the more complicated models.

• Example segregation problem

The appropriate model partial differential equations for the static boundary segregation problem are

$$\partial C_o / \partial t = \vec{\nabla} \cdot (D_o \vec{\nabla} C_o) \text{ in the region O,} \quad (1a)$$

$$\partial C_s / \partial t = \vec{\nabla} \cdot (D_s \vec{\nabla} C_s) \text{ in the region S,} \quad (1b)$$

where C_o is the concentration of the impurity in oxide and C_s is the concentration in silicon. D_o and D_s are the effective diffusivities in the two materials. In general, D_s depends on C_s , so (1b) is nonlinear.

At the oxide-silicon interface Γ , segregation is modeled using a standard first-order kinetic model [4],

$$D_o \partial C_o / \partial n = -h(C_o - 1/m C_s), \quad (2a)$$

$$D_s \partial C_s / \partial n = D_o \partial C_o / \partial n. \quad (2b)$$

Here, $\vec{n}$ is a unit normal vector that points from the oxide to the silicon, h is the boundary transport rate, and m is the equilibrium segregation ratio. It is important to note that the impurity concentration distribution is discontinuous at the interface. Thus, in Eqs. (2), C_o is the concentration on the oxide side of the interface, while C_s is the value on the silicon side.

Finally, for illustrative purposes, it is assumed that there is no outward normal flux of impurities through the exterior boundaries of the simulation region; i.e., that

$$\partial C_o / \partial n = \partial C_s / \partial n = 0 \text{ on } \partial R. \quad (3)$$

In a more realistic problem, (3) is applied at all points on the boundary except where evaporation, predeposition, or other transport processes are occurring. It is also assumed that some initial concentrations C_o^0 , C_s^0 are known in the oxide and silicon,

$$C_o = C_o^0, C_s = C_s^0 \text{ at time } t = 0. \quad (4)$$

• Discretization of the simulation region

Equations (1-4) are solved by a finite element-finite difference method that uses first-order triangular elements in space. The initial triangulation of the simulation region R is created by the user with the mesh generator TRIM, which is based on conformal mapping of simple "ideal" triangulations onto subregions of R . During generation of the initial mesh, the user may assign material properties to the elements that will distinguish the oxide region O of R from the silicon region S .

In the example problem, the interface between the oxide and silicon regions is found automatically by FEDSS, and a domain cut, or crack, is introduced at the interface if the segregation ratio m is to differ from one at some point during the simulation. The domain cut is created by simply assigning two node numbers to each point on the oxide-silicon interface, corresponding to the "top" and "bottom" of the cut. The existence of two node numbers then makes it

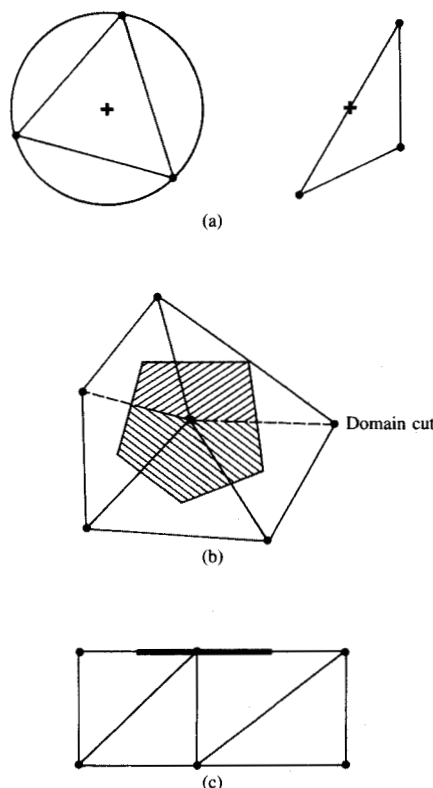


Figure 3

Numerical simplifications: (a) The diffusivity of an element is evaluated at the center of the circumcircle or at the center of the hypotenuse. (b) The weight of a node is the area of a polygon defined by the diffusivity evaluation points. At a domain cut, the weight is split between the two sides. (c) Boundary term weights at boundary nodes are taken to be half the sum of the lengths of the attached boundary edges.

possible to represent the concentration discontinuity implied by Eqs. (2).

• Discretization and solution of the equations

When Galerkin's method is applied in the standard way to Eqs. (1-3), the system of ordinary differential equations.

$$A_s \bar{c}_s + M_s d\bar{c}_s / dt + h/m K_s \bar{c}_s - h K_{so} \bar{c}_o = 0, \quad (5a)$$

$$A_o \bar{c}_o + M_o d\bar{c}_o / dt + h K_o \bar{c}_o - h/m K_{os} \bar{c}_s = 0 \quad (5b)$$

is obtained, where the vectors $\bar{c}_s$ and $\bar{c}_o$ are the concentrations at time t at each node in silicon and oxide, A_s and A_o are global stiffness matrices, M_s and M_o are mass matrices, and the K 's are matrices that contain geometrical information about the oxide-silicon interface. If the silicon region has s nodes and the oxide region has o nodes (counting interface nodes on the top and bottom of the cut),

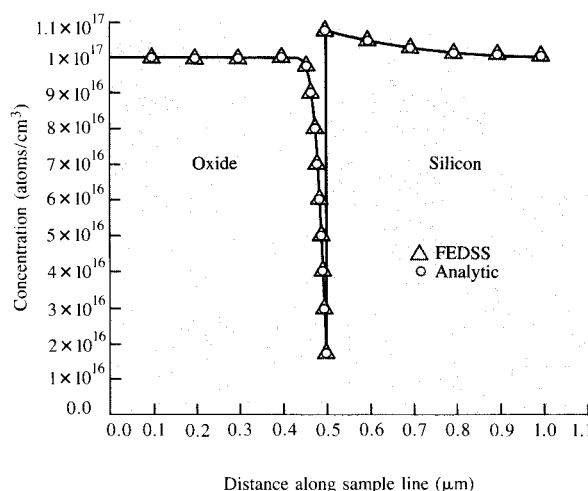


Figure 4
A high-resolution FEDSS static boundary phosphorus segregation profile, superimposed over an analytic solution.

then K_{so} and K_{os} are $s \times o$ and $o \times s$ matrices that can be interpreted as connecting the concentration distributions $\bar{c}_o$ and $\bar{c}_s$ across the interface.

Several simplifications (Figure 3) are introduced at this point that eliminate solution oscillations and a few other problems that sometimes occur when using standard "consistent" mass and boundary matrices, i.e.:

1. When the entries of the local stiffness matrix for a nonobtuse element are calculated, the diffusivity D_s is approximated on the element by its value at the center of the element circumcircle. For obtuse elements, the evaluation is done at the center of the hypotenuse [5]. This approximation helps to produce laterally uniform solutions in cases where the concentration varies only with depth and the nodes of the mesh form a rectangular lattice.
2. The mass matrices M_o and M_s are replaced by lumped diagonal matrices in which the weight attached to node i is the area of a polygon formed by joining the points where the diffusivities are evaluated [5]. This simplification eliminates oscillations that sometimes occur using the standard "consistent" mass matrices; it also contributes to lateral uniformity.
3. Matrices that are associated with a boundary are lumped by summing the entries in each row. Boundary nodes or node connections are thereby assigned a weight equal to half the sum of the lengths of the adjoining boundary edges. For example, if i_s and i_o are nodes on the top and bottom of the oxide-silicon interface cut, then the lumped K_s contains a node weight in the i_s diagonal position and the lumped K_{so} has an equal node connection weight in

the (i_s, i_o) position. As noted in 2), this simplification also eliminates oscillations and other undesirable effects at the boundary.

After these initial approximations are applied, Eqs. (5) are solved subject to initial conditions on $\bar{c}_o$ and $\bar{c}_s$ using a semi-implicit Euler method. In particular, considering Eq. (5a), the derivative is replaced by a difference quotient,

$$d\bar{c}_s/dt \approx (\bar{c}_s(k+1) - \bar{c}_s(k))/\Delta t,$$

where $\bar{c}_s(k)$ is the concentration at time step k . All other terms are evaluated at time step $k+1$, except for the stiffness matrix A_s , which is lagged to time step k in order to linearize the problem. The resulting equations are symmetrically structured but not symmetrically filled in, and are solved using one of the general sparse matrix techniques from SPARSPAK (Waterloo Sparse Linear Equations Package [6]). The overall numerical approach is quite fast and stable and is sufficiently accurate to yield an acceptable picture of dopant redistribution.

Figure 4 compares a high-resolution FEDSS simulation of phosphorus segregation with an analytic solution obtained for the special case in which $x < 0.5 \mu\text{m}$ is the oxide region, $x > 0.5 \mu\text{m}$ is the silicon region, and the phosphorus concentration is initially uniform at $1 \times 10^{17} \text{ cm}^{-3}$. The profiles correspond to 30 minutes of annealing in a neutral ambient at 1000°C . The agreement between the two curves is excellent.

Physical models

A general-purpose process simulator must be able to model a variety of process steps. FEDSS has evolved to the point where the key process steps can now be simulated with reasonable accuracy, although some restrictions still remain and verification continues. The essential features of each of these steps are now discussed.

• Ion implantation

Ion implantation is one of the most frequently used methods for introducing impurities into semiconductor materials. Commonly, dopants are implanted into silicon and silicon dioxide. However, silicon nitride, photoresist, aluminum, polysilicon, or various silicides are also used as target materials. FEDSS contains both analytic and Monte Carlo ion implantation models so that a wider range of the required physical situations can be realistically simulated.

The analytic models in FEDSS are based on tables relating the energy of an implant to various statistical fitting parameters, such as the range and vertical and lateral standard deviations. For a given impurity, energy, and dose, a fit is made to the vertical implant profile using a Gaussian, joined half-Gaussian [7], or Pearson IV [8] distribution. The default profiles are the joined half-Gaussian for arsenic and phosphorus [7, 9] and the Pearson IV for boron. For arsenic,

it is also possible to use a Pearson IV distribution for which limited experimental data have been published [10]. Exponential tails are added to the implanted distributions if required. All of the distributions are modified in 2D simulations to include Gaussian lateral scattering. In all cases, the user may override the implant tables by giving the implant parameters directly rather than specifying the energy.

An example of an analytic implant into an arbitrary surface is shown in **Figure 5**. In general, a mesh like the one in the figure may be the initial mesh or the end result of some sequence of deposition, etch, oxidation, or epitaxy steps. Whatever the case, the simulation region is automatically analyzed to locate the wafer surface. The implant is then done either over the entire top surface or between specified left and right infinite mask edges. Variations in surface topography and the effects of mask edges are handled analytically using the method of Runge [11]. Fictitious extensions of the wafer surface to infinity are added to more correctly model the concentrations at the sides of the mesh. Note from the example that undercuts can be accommodated. Angled implants are also possible.

In addition to the analytic implant models, a Monte Carlo ion implantation program [12, 13] is available to use when the effects of multiple layers must be modeled accurately, the implant falls out of the range of the analytic tables, or the tables are suspected to be inaccurate. In the Monte Carlo method, the individual trajectories of a number of ions (the default is 1000) are followed in the target layers. Ions change direction due to binary nuclear collisions and move in straight-line, free-flight paths between collisions. Nuclear and electronic interactions result in energy loss that continues until the ions stop. The particular program used in FEDSS considers the target to be amorphous, so that directional properties of the crystal lattice are ignored (further details can be found in the references). The results of the Monte Carlo implant are smoothed using local least squares with third-degree polynomials relative to five points. Repeated use of this procedure approaches, in the limit, a least-squares polynomial fit of the third degree which uses the total set of given points. This smoothing requires only a small amount of CPU time.

At present, the Monte Carlo calculation allows for 1D implants through a target of up to three horizontal layers, each of uniform thickness and uniform material properties. These layers may be composed of at most seven different atoms. No lateral scattering or effects due to mask edges are included as yet. Future work on the Monte Carlo program will generalize it to the same level as the analytic implant techniques.

A third approach to ion implantation for 1D implants and some 2D implants is to run the Monte Carlo program as a stand-alone simulator and use the four resulting statistical moments in the 2D FEDSS Pearson IV analytic model.

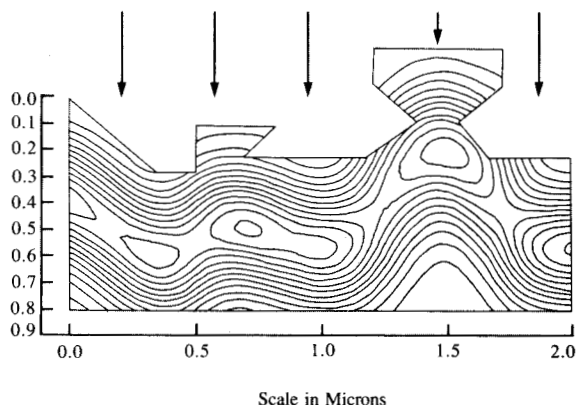


Figure 5

An analytic implant through a complicated surface. Scale is in μm .

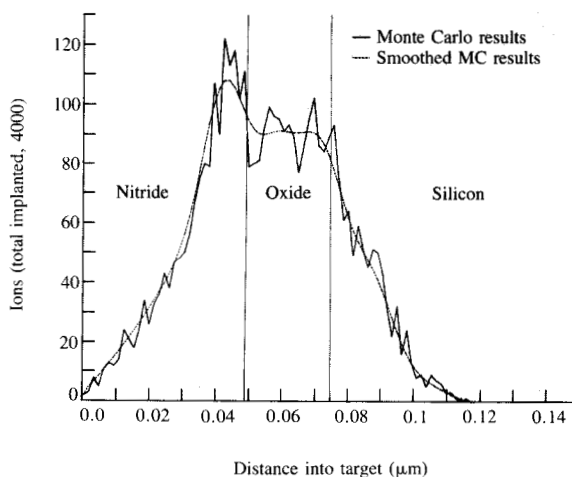


Figure 6

A profile from a Monte Carlo ion implantation simulation.

Figure 6 shows a result from the Monte Carlo ion implantation program for a boron implant at 20 keV into silicon through 0.0495 μm of nitride and 0.0255 μm of oxide. The resulting smoothed profile can be used directly in FEDSS.

• Predeposition and evaporation

Predeposition refers to the introduction of an impurity from a constant source at the wafer surface, while evaporation refers to the loss of impurity from the wafer surface to the surrounding gas.

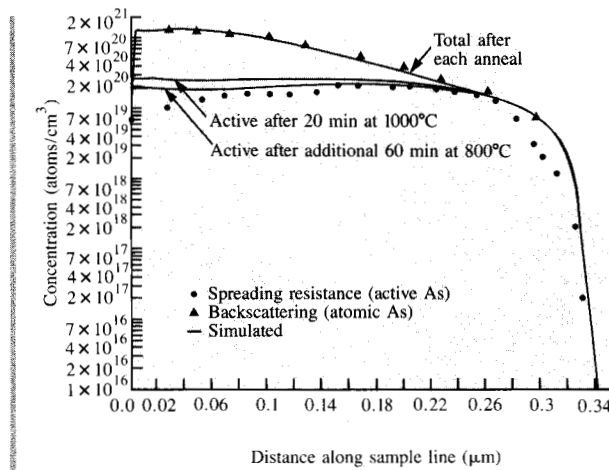


Figure 7

A comparison between the arsenic-charged three-cluster kinetic model and experimental profiles [14] showing the decrease in the active fraction during an 800°C annealing step.

Predeposition and evaporation are both modeled using the boundary condition [9]

$$D_{\text{eff}} \partial C / \partial n + hC = hC_{\text{ss}},$$

where C_{ss} is set equal to the solid solubility limit in the case of predeposition and to zero for evaporation, and D_{eff} is an effective diffusivity that depends on the physical model being used. The mesh boundary at which each transport process is applied is automatically taken to be the wafer surface. If required, however, the program allows the user to restrict transport to between one and five subsections of the surface.

• Recently added diffusion models

FEDSS contains a large menu of diffusion models: two arsenic equilibrium models, two arsenic kinetic models, one model each for boron and phosphorus, two arsenic-boron interaction models, and a p- or n-type "generic" diffusion model with either an Arrhenius or vacancy-assisted diffusivity. The numerical implementations of two of the more difficult models that were recently added are reported here.

Arsenic kinetic model

One of the arsenic kinetic models in FEDSS assumes that in silicon, at high arsenic concentrations, clusters form that consist of three arsenic ions, an electron, and a vacancy, thereby reducing the final electrical activity of the doped region. The equations involved in the model [14] can be written

$$\partial C_T / \partial t = \vec{\nabla} \cdot (D_{\text{iso}} (\vec{\nabla} a + a/n \vec{\nabla} n)), \quad (6a)$$

$$\partial C_{\text{clus}} / \partial t = K_C a^3 n - K_D C_{\text{clus}}, \quad (6b)$$

$$C_T = a + 3C_{\text{clus}}, \quad (6c)$$

where

C_T = total arsenic concentration in silicon,

a = active concentration in silicon,

C_{clus} = concentration of clusters in silicon,

n = electron concentration, given by

$$\begin{cases} np = n_i^2, \\ n - p = a + 2C_{\text{clus}}, \end{cases}$$

where p is the hole concentration and n_i is the intrinsic electron concentration,

$D_{\text{iso}} = (1 + \beta n/n_i + \gamma(n/n_i)^2)/(1 + \beta + \gamma)$, where β and γ are constants,

K_C, K_D = clustering and declustering reaction rates.

Equations (6) must be solved subject to boundary and initial conditions. The most general boundary conditions used in silicon, which cover evaporation, predeposition, epitaxy, and reflecting boundaries, are

$$D_{\text{eff}} \partial C_T / \partial n + k_1 C_T + k_2 \partial C_T / \partial t = k_3, \quad (7a)$$

$$\partial C_{\text{clus}} / \partial n = 0, \quad (7b)$$

where $D_{\text{eff}} = D_{\text{iso}}(1 + a/(n + p))$ and k_1, k_2, k_3 are piecewise constant on the boundary. Note that Eqs. (6) involve two independent variables (C_T and C_{clus} , for example), so boundary conditions must be given in silicon for both of them. In materials other than silicon or polysilicon, only the standard diffusion equation (1a) for the total arsenic concentration is solved, subject to appropriate boundary conditions. If an oxide-silicon interface is present, segregation boundary conditions [Eqs. (2)] are imposed on C_T at the interface in place of (7a).

The finite element solution of (6) subject to (7) is not straightforward because the chemical reaction equation (6b) does not contain a Laplacian. The approach adopted involves first expressing $\vec{\nabla} a$ and $\vec{\nabla} n$ in (6a) in terms of $\vec{\nabla} C_T$ and $\vec{\nabla} C_{\text{clus}}$,

$$\partial C_T / \partial t = \vec{\nabla} \cdot (D_{\text{eff}} \vec{\nabla} C_T) - \vec{\nabla} \cdot ((D_{\text{eff}} + 2D_{\text{iso}}) \vec{\nabla} C_{\text{clus}}).$$

The spatial discretization procedure and semi-implicit Euler method described earlier are next applied to the above equation and boundary conditions, yielding a matrix equation of the general form

$$A_1(k) \bar{c}_T(k+1) = \bar{r} + A_2 \bar{c}_T(k) + A_3(k) \bar{c}_{\text{clus}}(k+1),$$

where k is the time step index. The term $\bar{c}_{\text{clus}}(k+1)$ is now evaluated in terms of known quantities by simply using a semi-implicit approximation to (6b) at each node i ,

$$(c_{\text{clus}}^i(k+1) - c_{\text{clus}}^i(k))/\Delta t = K_C a^i(k)^3 n^i(k) - K_D c_{\text{clus}}^i(k+1).$$

All concentrations in the above equation are normalized by

dividing by n_i to prevent overflows. Finally, an automatic time selection scheme is used to reduce Δt during periods of rapid clustering or declustering in order to improve tracking of the active and clustered fractions.

Figure 7 compares the above model and experimental profiles [14] measured after implanting arsenic at 140 keV through 0.025 μm of SiO_2 at 2×10^{16} ions/ cm^2 and then annealing at 1000°C for 20 min and at 800°C for 60 min. Note that the total profile is essentially unchanged during the 800°C step but that the active fraction decreases due to clustering.

Arsenic-boron interaction models

The two arsenic-boron interaction models in FEDSS predict the effects of arsenic on the diffusion of boron, including the retardation of boron diffusion under a highly doped arsenic emitter and the depletion of boron at an n-p junction due to a steep arsenic profile. Both of the interaction models have the general form [15]

$$\partial A / \partial t = \vec{\nabla} \cdot (D_{11} \vec{\nabla} A + D_{12} \vec{\nabla} B), \quad (8a)$$

$$\partial B / \partial t = \vec{\nabla} \cdot (D_{21} \vec{\nabla} A + D_{22} \vec{\nabla} B), \quad (8b)$$

where A is the total arsenic concentration, B is the boron concentration, and D_{11} , D_{12} , D_{21} , D_{22} depend on A , B , and the physics assumed. The chief mathematical difficulty in Eqs. (8) is the coupling between the two impurities. If the numerical approach described above is applied to (8), then a system of linear equations of the form

$$C_{11}(k)\bar{a}(k+1) + C_{12}(k)\bar{b}(k+1) = \bar{r}_1(k), \quad (9a)$$

$$C_{21}(k)\bar{a}(k+1) + C_{22}(k)\bar{b}(k+1) = \bar{r}_2(k) \quad (9b)$$

is obtained that must be solved at each time step k for $\bar{a}$ and $\bar{b}$. Although the individual coefficient matrices in (9) are symmetric, with C_{11} and C_{22} usually being positive definite, the global coefficient matrix does not have such nice characteristics. Therefore, rather than trying to solve the global system directly, a block Gauss-Seidel iterative method was adopted that uses C_{11} , $\dots$, C_{22} as the blocks and SPARSPAK as a direct solver of the blocked equations. This method has been observed to converge sufficiently within 2 to 4 iterations (with a gain of 2 to 3 digits per iteration) and has never diverged in any of the problems that have been run.

Figure 8 compares an arsenic-boron interaction model that uses the equilibrium version of (6) with SIMS data. Boron was first implanted at 50 keV through 0.005 μm of oxide and annealed at 800°C for 20 min. Arsenic was then implanted at 60 keV and the sample was further annealed for 35 min at 1000°C in N_2 . FEDSS implant parameters and tails were adjusted to match SIMS profiles measured prior to the 1000°C step.

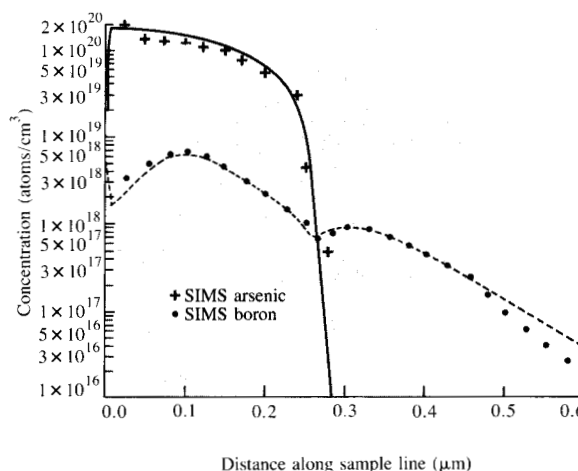


Figure 8

A comparison between the FEDSS arsenic-boron interaction model (with equilibrium arsenic 3-clustering) and SIMS data. Boron depletion at the junction is correctly predicted.

• Oxidation

Redistribution in an oxidizing ambient is unquestionably the most complicated and resource-intensive step simulated by FEDSS. Although FEDSS is capable of oxidizing rather general structures, as the examples below show, work is in progress to extend the range of the algorithm, improve its reliability, and reduce resource requirements. The description given below refers to the program in its current state.

Oxidation model and algorithm

An earlier version of FEDSS modeled oxidation by solving the diffusion equation for the transport of oxygen through oxide and silicon. Silicon elements at the interface were changed to oxide when they had accumulated a sufficient concentration of oxidant. At the same time, matching "air" elements were turned into oxide to simulate volume expansion. A deposited silicon nitride layer acted as a diffusion barrier in regions where oxidation was not desired. This algorithm, although theoretically fast, was not physically realistic and applied to only a small variety of problems.

Because of the general-purpose requirements of FEDSS, it was apparent that another model was needed. The physical model now used is due to Chin and co-workers [16], who developed a method for doing 2D oxidation and implemented it in a stand-alone program called SOAP (Stanford Oxidation Analysis Program [17]). This model of 2D oxide growth is based on steady-state oxidant diffusion and slow incompressible viscous oxide flow. The numerical method in SOAP uses a velocity/pressure iteration algorithm

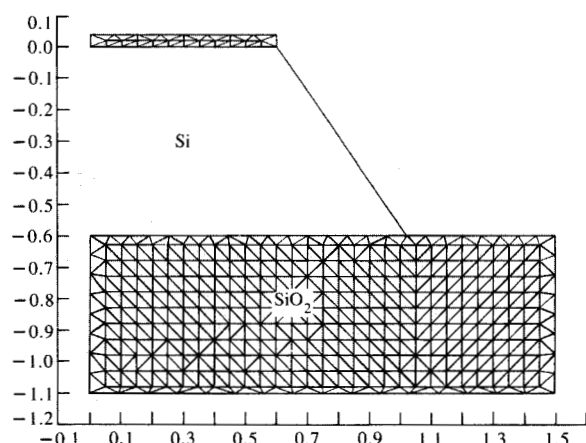


Figure 9

The initial structure used in a stand-alone OX2D oxidation simulation. Scale is in μm .

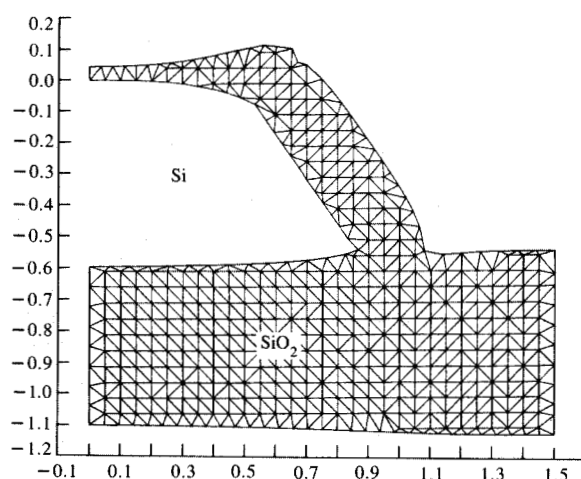


Figure 10

Simulated oxide shape after 5-37-5-minute dry-wet-dry oxidation at 950°C. Scale is in μm .

due to Chorin [18] and a boundary value method for solving Poisson's equation. SOAP does not account for impurity redistribution.

Because of the complexity of the oxidation-redistribution problem, a two-stage approach was used. First a stand-alone

oxidation program called OX2D was written, then OX2D was interfaced with FEDSS, where a prototype redistribution algorithm had been developed that worked well with the 1D Deal-Grove oxidation model [19]. The finite element implementation of the new 2D oxidation model was formidable because of mesh generation problems and high numerical accuracy requirements. The major steps in the OX2D algorithm are as follows:

1. To solve the mesh generation problem, an automatic grid point triangulation program is used at each time step to create a mesh with the required boundaries and some specified grid spacing.

2. The steady-state oxygen diffusion problem

$$\nabla^2 C = 0,$$

$$-D\partial C/\partial n = kC \text{ at the SiO}_2\text{-Si interface,}$$

$$-D\partial C/\partial n = h(C - C_o) \text{ at the SiO}_2\text{-gas interface}$$

is solved to get the concentration of oxidant at the oxide-silicon interface. Here, C is the concentration of O_2 in oxide, D is the diffusivity of oxygen in oxide, k is the SiO_2 -Si interface reaction rate, and C_o is the concentration of O_2 in the surrounding gas. Quadratic elements are used here and in steps 4b and 4d in order to provide increased numerical accuracy.

3. The oxidant concentrations from step 2 are converted to boundary velocities using a standard first-order chemical reaction model [19],

$$VN = kC,$$

where V is the speed of the interface as measured from the oxide and N is the number of O_2 molecules in a unit volume of SiO_2 . The velocity at each interface point is assumed to be normal to the interface.

4. Consistent velocity and pressure fields are found using a version of Chorin's algorithm:

- a. A guess is made at the pressure field p .

- b. The velocity field $\vec{v}$ is calculated by solving

$$v\nabla^2 \vec{v} = \vec{\nabla} p,$$

$$-v\partial \vec{v}/\partial n = h(\vec{v} - (1 - \alpha)V\vec{n}) \text{ at the SiO}_2\text{-Si interface,}$$

where v is the viscosity of the oxide, α is the ratio of a volume of silicon to the resulting volume of SiO_2 , and V is the boundary velocity distribution from step 3.

- c. The pressure field is corrected with

$$p(i, k + 1) - p(i, k) = -2v\vec{\nabla} \cdot \vec{v}(i, k)/M(i),$$

where $p(i, k)$ is the pressure at node i at iteration k , $\vec{v}(i, k)$ is the corresponding velocity, and $M(i)$ is the mass at node i . Linear elements and a lumped mass matrix are used in this case.

- d. The pressure field is recalculated by solving $\nabla^2 p = 0$ subject to corrected boundary pressures found in step c on all silicon boundaries. Pressure effects due to nitride bending are currently being incorporated.

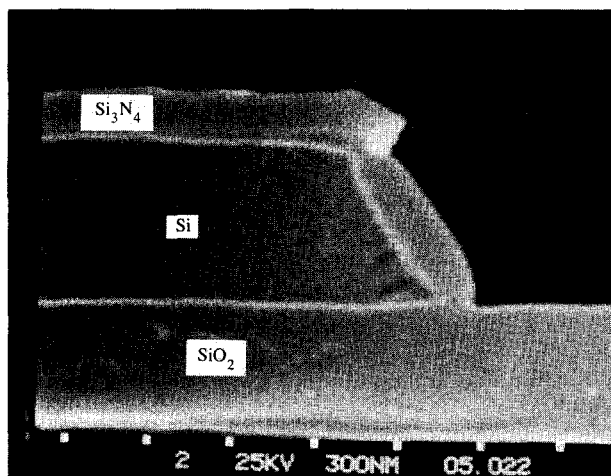


Figure 11

SEM picture of actual oxide shape after 5-37-5-minute dry-wet-dry oxidation at 950°C.

- e. Steps b to d are now repeated until the velocities converge. A fully converged solution satisfies $\vec{\nabla} \cdot \vec{v} = 0$.

The stand-alone OX2D program was used to simulate the oxidation of the structure shown in the cross section view of **Figure 9**. This figure shows the initial mesh generated by the program from the user's description of the boundary. In this case, a 0.6- μm -thick silicon layer on a silicon dioxide substrate has a sidewall slope of approximately 60 degrees. The top horizontal surface of the silicon is covered with 0.04 μm of silicon dioxide under a silicon nitride oxidation barrier film. The Si_3N_4 layer does not appear in any of the mesh plots since nitride deformation is not currently being modeled. All oxidation cycles were run at 950°C. The linear and parabolic growth rates are the same as those used in FEDSS and SUPREM II for (111) silicon.

Figure 10 shows the simulated oxide profile after an oxidation cycle of 5 minutes dry plus 37 minutes wet plus 5 minutes dry at 950°C. With this cycle, the top corner of the silicon is rounded due to a bird's-beak phenomenon and a slight point is evident. The oxide grown at the bottom corner is thinner than on the sidewall and is beginning to form a cusp. This result is attributed to the viscous flow of the oxide. Finally, the silicon at the bottom corner comes to a sharp point due to the oxidation of the underside of the silicon. **Figure 11** is an SEM cross section of a sample which was oxidized with the above cycle. In this view the silicon nitride is present. Note that the model accurately predicts the observed features, except that the top and bottom of the silicon are over-oxidized due to the uniform application of (111) rate constants on all oxidizing surfaces.

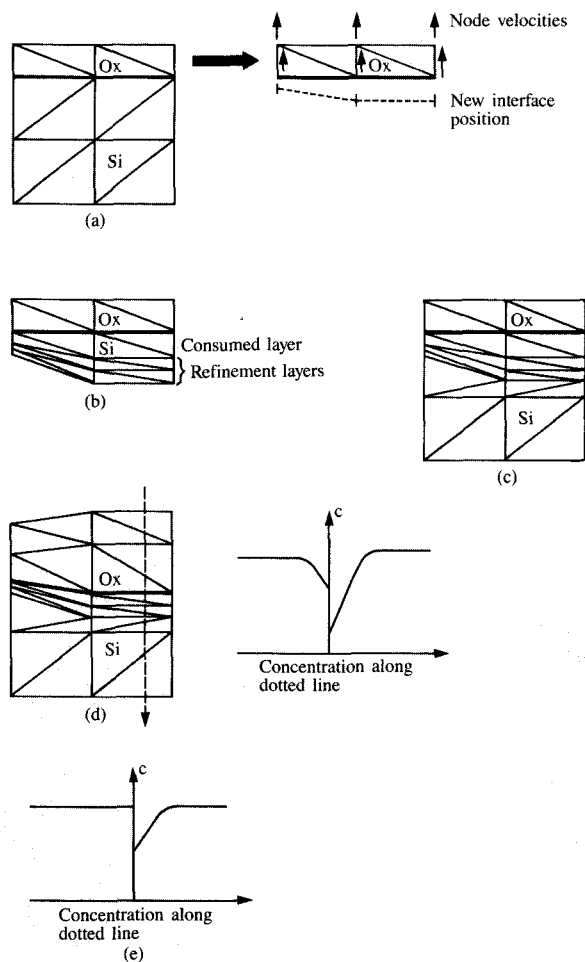


Figure 12

The boundary movement/impurity redistribution algorithm in FEDSS: (a) Current oxide region is removed from mesh at start of time step; oxide node velocities and new oxide-silicon interface position are calculated by OX2D. Thickened line in figure is a thin row of elements at interface. (b) Initial and refinement layers of silicon elements are added forward of oxide-silicon interface. (c) Remaining silicon region is reattached to mesh. (d) Silicon is consumed and oxide nodes are allowed to flow. Impurities are redistributed at interface. (e) Impurities are globally redistributed with segregation at interface.

Moving boundary/impurity redistribution algorithm

A realistic simulation of impurity redistribution during oxidation must take into account the physical motion and changes occurring at the oxide-silicon interface. These include the advancement of the interface into silicon as the oxidation proceeds, volume expansion of the newly oxidized silicon, and transport and segregation of impurities across the interface. Oxidation-initiated diffusion phenomena, such as oxidation-enhanced or retarded diffusion, also need to be considered.

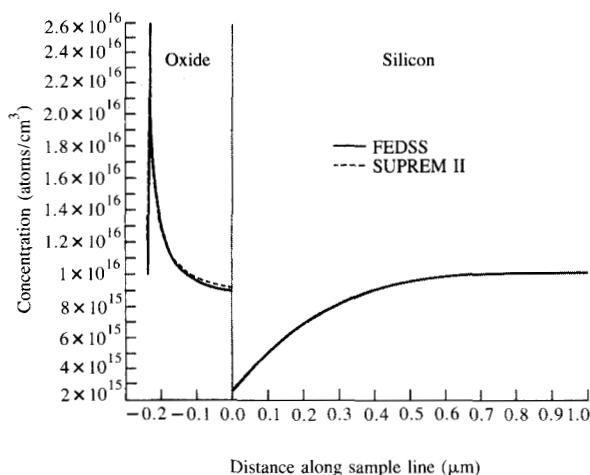


Figure 13

Comparison of numerical methods for impurity redistribution in FEDSS and SUPREM II. Thirty minutes of wet oxidation at 1000°C was simulated using identical data beginning with an initially flat boron background and 0.002 μm of surface oxide.

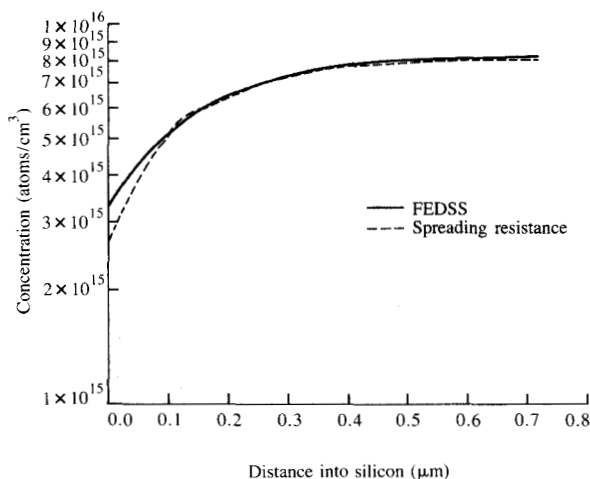


Figure 14

Comparison between a FEDSS-generated profile and spreading resistance data showing the redistribution of boron in silicon after 150 minutes of dry oxidation at 950°C.

A heuristic algorithm has been developed for the simulation of impurity redistribution during thermal oxidation. The simulation results agree well with those from SUPREM II and analytic models, thus verifying the correctness of the numerical method. Details of the algorithm during a single time step are as follows (Figure 12):

1. At the beginning of the time step, the oxide region present is removed from the mesh and passed to OX2D, which finds the velocity at each oxide node and the new position of the oxide-silicon interface.
2. The mesh is next regenerated around the oxide region in a way that is determined by the new interface location.
 - a. An initial band of elements, corresponding to the silicon region to be oxidized, is attached to the oxide region ahead of the interface.
 - b. If impurities are present, several more silicon bands are added ahead of the initial band in order to adequately resolve the impurity profiles.
 - c. All remaining silicon nodes in the mesh are attached to the last refinement band by using an automatic grid point triangulation program.
3. All oxide nodes that are not on the oxide-silicon interface are allowed to flow to their new positions.
4. The elements in the initial band are changed to oxide and the interface is moved almost all the way to the new position. If impurities are present, their concentrations at the new interface are adjusted using a finite difference version of the mass conservation boundary condition [4]

$$D_o \vec{\nabla} C_o \cdot \vec{n} + V(C_o - \alpha C_s) = D_s \vec{\nabla} C_s \cdot \vec{n}, \quad (10)$$

where V is defined above. The effect of oxide volume expansion on impurity concentrations is accounted for at this point. Since (10) does not completely determine the interface concentrations at the new position, it is also required that the impurity ratio C_s/C_o at the interface should be maintained.

5. Elements are removed from oxide near the interface in a way that ensures that the oxide region will have a given maximum grid spacing in the direction normal to the interface. This slows the rate of growth of the number of nodes and elements in the mesh. The final oxide grid is also nearly independent of the time step.
6. Finally, impurities are redistributed in the oxide and silicon and at the interface between them using the domain cut algorithm explained earlier. At present, the default values used for the segregation ratio m are from Refs. [20, 21]. Oxidation-enhanced or -retarded diffusion is also accounted for by using a model based on the findings of several investigators [22–24].

Apart from rounding errors, only steps 2a, 2b, 3, and 5 may introduce numerically induced changes in impurity doses. In numerous simulations, the total mass of any impurity has been observed to change only slightly (generally much less than one percent) due to interpolation, errors in oxide flow, or oxide element deletion.

Figure 13 shows FEDSS and SUPREM II boron profiles after 30 min of wet O_2 oxidation at 1000°C. For both simulations, the initial grid has 0.002 μm of surface oxide, and both oxide and silicon are initially doped with boron at

a uniform concentration of $1 \times 10^{16} \text{ cm}^{-3}$. The two programs were run using the same oxidation, segregation, and diffusion parameter data in order to verify the numerical method in FEDSS. Note that the match is excellent, particularly in silicon. Results of similar quality were obtained for arsenic, boron, and phosphorus between 800°C and 1100°C. **Figure 14** compares a FEDSS-generated profile with spreading resistance data in silicon that was uniformly doped with boron at $8.1 \times 10^{15} \text{ cm}^{-3}$ and subsequently subjected for 150 min to dry O_2 oxidation at 950°C. The FEDSS profile was obtained by using the default values of various oxidation-redistribution parameters in the program. There is reasonably good agreement between the data and the simulation.

• Epitaxy

The silicon epitaxy model in FEDSS is preliminary, pending further work on a generalized 2D model. Dopant incorporation during epitaxy is currently modeled via the boundary condition [25]

$$D_{\text{eff}} \partial C / \partial n + (K_{\text{mf}} / K_p + g) C + K_A \partial C / \partial t = K_{\text{mr}} P_D^0,$$

where g is the growth rate of the epitaxial layer, P_D^0 is the partial pressure of the dopant in the gas phase, and K_{mf} , K_p , and K_A are parameters that depend on the dopant, the reactor geometry, and other physical parameters. Uniformly thick silicon layers can currently be grown. When the new layer requires additional storage, a restart feature in FEDSS is invoked. This is done to increase the dynamic storage limit if the space required during automatic mesh generation exceeds the amount initially allocated.

• Material deposition and etching

Another processing step that has been incorporated into FEDSS is low-temperature deposition. FEDSS allows the layering of polysilicon, Si_3N_4 , SiO_2 , silicon, and up to three user-defined "generic" materials over an existing region. This is accomplished by specifying the region of interest, where all of the mesh elements are changed to the requested material type. The deposited layer may be doped as desired. Diffusion models are available for all impurities in all materials. However, parameter data for the models must be supplied by the user in "exotic" cases.

FEDSS also contains an etch step that works like a deposition, except that elements in an existing region are either changed to "air" or removed from the mesh.

Applications

• CMOS N-well simulation

A technologically important process is the creation of the N-well in bulk CMOS processing along with the adjacent field isolation region (see the circled area in **Figure 15**). Simulation of a CMOS process in 2D was reported earlier

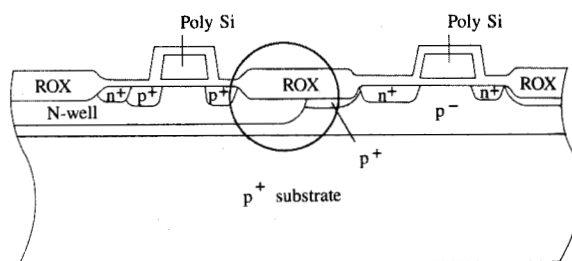


Figure 15

Cross section of a CMOS device. The simulated region is circled.

CMOS N-WELL AND FIELD OXIDATION

```

-----
ALLOCATE STORAGE, SCALE THE MESH, AND DEFINE THE SUBSTRATE.
&ALLEY MELEMS=8000, MNODES=4200, MAXS=230000 &END
&DIM XL=8, YL=5 &END
&SUBS CON=1.0E15, ORNT=100, ELEM='+' &END

-----
DEPOSIT .04 MICRONS OF OXIDE.
&DEPO MAT='OXIDE', XY=0,4.96, 8,4.96, 8,5, 0,5, STOP=T &END

-----
IMPLANT PHOSPHORUS. SPECIFY THE DISTRIBUTION PARAMETERS AND USE
A GAUSSIAN MODEL.
&START RECORD=3 &END
&IMPL ELEM='P', DOSE=5.0E12, RPY=.4330, SY=.1340, SX=.1701,
MODL='LATG', XWMIN=0.0, XWMAX=4.401, SAVTOT=1 &END

-----
DRIVE FOR 360 MIN AT 1160 C.
&DRV TIME=360, DELTIM=1, TEMP=1160, AMBIENT='NIT', SAVTOT=1 &END

-----
IMPLANT BORON. USE RPY AND SY FROM SMITH, SX FROM GIBBONS. POSITION
THE WINDOW TO SIMULATE A MASK MISALIGNMENT.
&IMPL ELEM='B', DOSE=1.0E13, RPY=.2990, SY=.0760, SX=.1203,
XWMIN=4.849, XWMAX=7.99999, MODL='LATG', SAVTOT=1 &END

-----
COVER THE N-WELL AREA WITH .02 MICRONS OF NITRIDE.
&DEPO MAT='NIT', XY=0,4.98, 4,4.98, 4,5, 0,5, SAVTOT=1 &END

-----
ETCH AWAY .02 MICRONS OF NITRIDE TO THE RIGHT OF THE N-WELL.
&ETCH XY=4,4.98, 8,4.98, 8,5, 4,5, SAVTOT=1 &END

-----
GROW THE FIELD OXIDE--WET PORTION ONLY. OXGRID IS THE GRID
SPACING IN OXIDE; DYNIT IS THE NITRIDE THICKNESS.
&DRV TIME=292, TEMP=950, AMBIENT='WET',
OXGRID=.05, DYNIT=.02, SAVTOT=1 &END

-----
SUM OF REMAINDER OF 1000C TIME-TEMP CYCLES.
&DRV TIME=115, TEMP=1000, DELTIM=1, AMBIENT='NIT', SAVALL=1 &END

```

Figure 16

FEDSS input for the CMOS example. All process step cards apart from the 292-minute wet oxidation step are exactly as run, except that controls used to stop and restart the program have been omitted. (A stop after the first DEPO step and a restart at the next IMPL step are shown, however.) The 292-minute oxidation step was broken into two smaller steps in order to vary the time step. Model cards used during the simulation reside in a different data set and are not shown here. Input cards use the FORTRAN namelist input facility. All text that does not appear between & <NAME> and &END (where <NAME> is a namelist name like DEPO) is regarded as a comment.

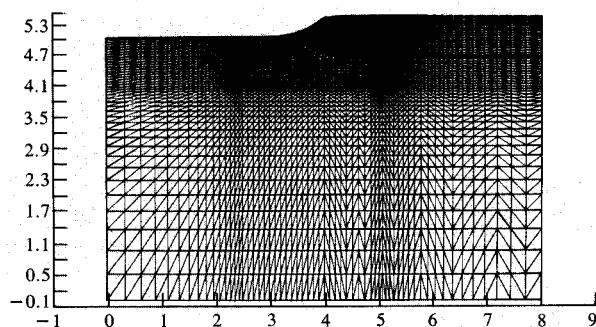


Figure 17

The mesh at the end of the CMOS process. Scale is in μm .

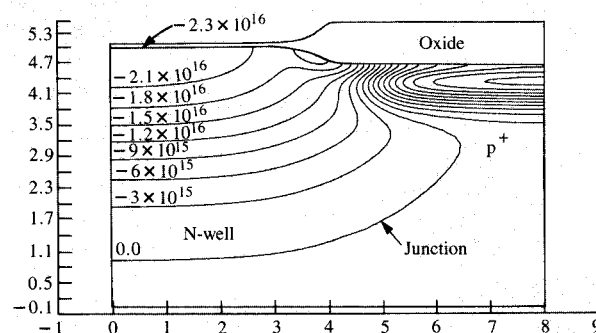


Figure 18

The net active concentration at the end of the CMOS process, showing the phosphorus N-well and the lateral penetration of boron into the N-well. Scale is in μm .

[2], before oxidation was available in FEDSS. The steps involved in the process are shown in **Figure 16**, which also illustrates the input language currently used by the program. Note that a phosphorus implant is used to form the N-well region and a misaligned boron implant is used to form the field region. Of interest is not only the junction depth of the N-well region, but also the extent of the lateral penetration of the boron into the N-well due to the mask misalignment. These and similar results may be used in the development of ground rules for designers.

As mentioned above, the earlier work did not include the oxidation step of growing a semi-recessed oxide isolation (SROX) region. This oxidation step generates a "bird's-beak"

which alters the profile of the boron. With the addition of the oxidation feature in FEDSS, the SROX growth is now possible. **Figure 17** illustrates the shape of the oxide region at the end of the process. The final mesh has approximately 4200 nodes and 8000 elements. **Figure 18** is a corresponding contour plot of the final net active concentration in silicon. The concentration distribution in oxide is modeled but is not shown here. This figure clearly shows the extent of vertical phosphorus diffusion and lateral boron penetration. For this simulation, most of the computation time and storage allocation were used during the 292-minute oxidation step, which required 131 minutes and 9.5 megabytes on an IBM 3084 computer.

• Trench sidewall oxidation

The generality of the oxidation-redistribution package developed in FEDSS allows a wide variety of advanced structures to be investigated. The example in **Figures 19** and **20** demonstrates the application of FEDSS to the simulation of boron redistribution after trench sidewall oxidation. The structure in Fig. 19 shows the redistributed boron equilibrium concentration contours in the oxide and silicon prior to the trench oxidation step. The edge of a window where boron was initially ion-implanted is indicated by an arrow. The junction with the n-type silicon substrate is at the $2.2 \times 10^{16} \text{ cm}^{-3}$ contour line. The boron contours as they appear after 0.05 μm of trench sidewall oxidation are shown in Fig. 20. When these are compared with the contours in Fig. 19, the depletion of boron due to segregation in the vicinity of the region being oxidized becomes evident. The final redistribution over the entire silicon region has also been affected by oxidation-enhanced diffusion.

Summary and future work

In this paper, the most advanced models and up-to-date features of the semiconductor process simulation program FEDSS have been presented. Specific application examples have also been provided to better illustrate some of the capabilities of the program. At present, the main feature of FEDSS is its ability to model generalized 2D oxidation while simulating impurity redistribution in both oxide and silicon. Even though there are a few other 2D process simulation programs that provide some oxidation capability, it is believed that the present oxidation-redistribution package in FEDSS is the most complete and generalized system of its kind. As part of future enhancements, work is under way to develop the next generation of process models. These models need to be based on improved physics and lend themselves to faster and more efficient numerical solutions.

Acknowledgments

Special thanks are due to Paul Murray for writing a crucial and difficult portion of the FEDSS program. We would also like to express our appreciation to Evan Lamp, Frank White,

Steve Laux, Jim Ziegler, Fred Morehead, Al Michel, and many other individuals for the time, advice, code, or data that they generously contributed to the project.

References

1. H. H. Hansen and K. A. Salsburg, "Two-Dimensional Process Modeling Using the Finite Element Method," presented at the First International Symposium on Very Large Scale Integration Science and Technology, Detroit, MI, October 17-22, 1982.
2. K. A. Salsburg and H. H. Hansen, "FEDSS—Finite Element Diffusion Simulation System," *IEEE Trans. Electron Devices* **ED-30**, 1004 (1983).
3. E. M. Buturla, P. E. Cottrell, B. M. Grossman, and K. A. Salsburg, "Finite-Element Analysis of Semiconductor Devices: The FIELDAY Program," *IBM J. Res. Develop.* **25**, 218-231 (1981).
4. D. A. Antoniadis, M. Rodoni, and R. W. Dutton, "Impurity Redistribution in SiO₂-Si During Oxidation: A Numerical Solution Including Interfacial Fluxes," *J. Electrochem. Soc.* **126**, 1939-1945 (1979).
5. C. Price, "Device Analysis for GaAs Technology," *Integrated Circuits Laboratory Report: Computer-Aided Design of IC Fabrication Processes—Process and Device Simulation*, Stanford University, Stanford, CA, August 19, 1982.
6. A. George and J. W.-H. Liu, *Computer Solution of Large Sparse Positive Definite Systems*, Prentice-Hall, Inc., Englewood Cliffs, NJ, 1981.
7. K. Inoue, T. Hirao, Y. Yaegashi, and S. Takayanagi, "Asymmetrical Profiles of Ion Implanted Phosphorus in Silicon," *Jpn. J. Appl. Phys.* **18**, A367 (1979).
8. W. K. Hofker, D. P. Oosthoek, N. J. Koelman, and H. A. M. De Grefte, *Radiat. Effects* **24**, 223 (1975).
9. D. A. Antoniadis and R. W. Dutton, "Models for Computer Simulation of Complete IC Fabrication Processes," *IEEE Trans. Electron Devices* **ED-26**, 490 (1979).
10. F. Jahnel, H. Ryssel, G. Prinke, K. Hoffmann, K. Muller, J. Biersack, and R. Henkelmann, "Description of Arsenic and Boron Profiles Implanted in SiO₂, Si₃N₄ and Si Using Pearson Distributions With Four Moments," *Nucl. Inst. & Meth.* **182/183**, 223 (1981).
11. H. Runge, "Distribution of Implanted Ions Under Arbitrarily Shaped Mask Edges," *Phys. Status Solidi A* **39**, 595 (1977).
12. J. P. Biersack and L. G. Haggmark, *Nucl. Inst. & Meth.* **174**, 257 (1980).
13. J. F. Ziegler, J. P. Biersack, and U. Littmark, *The Stopping and Range of Ions in Solids*, Vol. 1, Pergamon Press, New York, 1984.
14. M. Y. Tsai, F. F. Morehead, J. E. E. Baglin, and A. E. Michel, "Shallow Junctions by High-Dose As Implants in Si: Experiments and Modeling," *J. Appl. Phys.* **51**, 3230-3235 (1980).
15. S. M. Hu and S. Schmidt, "Interactions in Sequential Diffusion Processes in Semiconductors," *J. Appl. Phys.* **39**, 4272 (1968).
16. D. Chin, S. Y. Oh, S. M. Hu, R. W. Dutton, and J. L. Moll, "Two-Dimensional Oxidation," *IEEE Trans. Electron Devices* **ED-30**, 744-749 (1983).
17. D. Chin, S. Y. Oh, and R. W. Dutton, "A General Solution Method for Two-Dimensional Nonplanar Oxidation," *IEEE Trans. Electron Devices* **ED-30**, 993 (1983).
18. A. J. Chorin, "A Numerical Method For Solving Incompressible Viscous Flow Problems," *J. Comput. Phys.* **2**, 12-26 (1967).
19. B. E. Deal and A. S. Grove, "General Relationship for the Thermal Oxidation of Silicon," *J. Appl. Phys.* **36**, 3770-3778 (1965).
20. R. B. Fair and J. C. C. Tsai, "Theory and Direct Measurement of Boron Segregation in SiO₂ During Dry, Near Dry, and Wet O₂ Oxidation," *J. Electrochem. Soc.* **125**, 2050-2058 (1978).
21. A. S. Grove, O. Leistiko, Jr., and C. T. Sah, "Redistribution of Acceptor and Donor Impurities During Thermal Oxidation of Silicon," *J. Appl. Phys.* **35**, 2695-2701 (1964).
22. T. Y. Tan, U. Gosele, and F. F. Morehead, "On the Nature of Point Defects and the Effect of Oxidation on Substitutional Dopant Diffusion in Silicon," *Appl. Phys. A* **31**, 97-108 (1983).
23. S. Matsumoto, Y. Ishikawa, and T. Niimi, "Oxidation Enhanced and Concentration Dependent Diffusions of Dopants in Silicon," *J. Appl. Phys.* **54**, 5049-5054 (1983).
24. D. A. Antoniadis and I. Moskowitz, "Diffusion of Substitutional Impurities in Silicon at Short Oxidation Times: An Insight into Point Defect Kinetics," *J. Appl. Phys.* **53**, 6788-6796 (1982).
25. R. Reif and R. W. Dutton, "Computer Simulation in Silicon Epitaxy," *J. Electrochem. Soc.: Solid State Sci. & Technol.* **128**, 909-918 (1981).

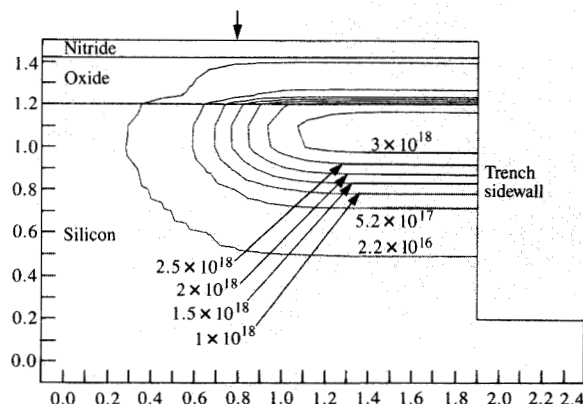


Figure 19

Boron contours before trench sidewall oxidation. Scale is in μm .

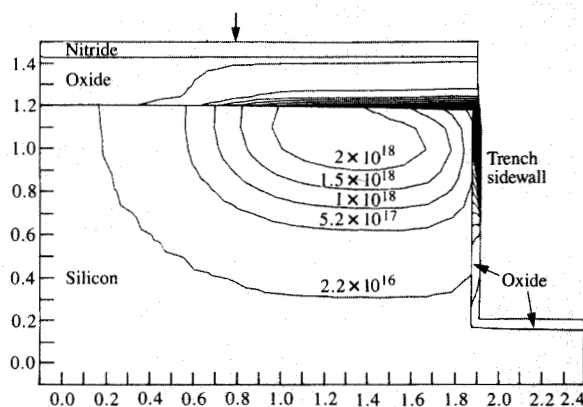


Figure 20

Boron contours after trench sidewall oxidation, showing increased depletion near the oxidized region and the effects of oxidation-enhanced diffusion. Scale is in μm .

Received October 2, 1984; revised December 6, 1984

Leonard Borucki *IBM General Technology Division, Burlington facility, P.O. Box A, Essex Junction, Vermont 05452.* Dr. Borucki received a Ph.D. in mathematics from Rensselaer Polytechnic Institute, Troy, New York, in 1978 after writing a thesis in number theory that was supported by an IBM research fellowship. He then taught courses in applied and pure mathematics at Lafayette College in Easton, Pennsylvania, until 1982, when he left to join the Advanced Mathematics and Engineering Analysis Department at the IBM Essex Junction laboratory. His current interests include semiconductor process modeling and the design and implementation of numerical algorithms. He received a GTD Achievement Award in 1983 for his work on FEDSS. Dr. Borucki is a member of the American Mathematical Society and the Mathematical Association of America.

Howard H. Hansen *IBM General Technology Division, Burlington facility, P.O. Box A, Essex Junction, Vermont 05452.* Dr. Hansen received a B.S. degree in mathematics from the University of Michigan in 1962 and a Ph.D. in physics from the State University of New York at Binghamton in 1973. He joined IBM in 1962, working first as a systems engineer and later as an applications programmer writing computer simulation programs for mechanical engineering design. After completing graduate work he joined the IBM Manassas laboratory and worked on epitaxial and ion implantation processes. Since 1977 he has been with the IBM GTD laboratory in Essex Junction, working on advanced VLSI processes with particular emphasis on 2D process modeling. From 1982 to 1984, Dr. Hansen managed a department responsible for bipolar technology modeling. He is currently a development engineer managing a department responsible for process and device modeling development in the base technology area of the Burlington laboratory. Dr. Hansen is a member of the American Physical Society and the Electron Devices Section of the Institute of Electrical and Electronics Engineers.

Khodadad Varahramyan *IBM General Technology Division, Burlington facility, P.O. Box A, Essex Junction, Vermont 05452.* Dr. Varahramyan joined IBM in 1982 at Burlington, where he has worked in the area of advanced modeling systems for semiconductor technology development. His present interests include development of advanced process models and application of computer tools in the design and fabrication of integrated circuits. He received his B.S. from the University of Illinois in 1977, and his M.S. and Ph.D. in electrical and systems engineering from Rensselaer Polytechnic Institute, Troy, New York, in 1979 and 1983. Dr. Varahramyan is a member of the Institute of Electrical and Electronics Engineers and Sigma Xi.