Simulation Of Point Defect Clustering In Cz-Silicon Wafers On The Cray T3E Scalable Parallel Computer: Application To Oxygen Precipitation

Faouzia Sahtout Karoui*, Abdennaceur Karoui, and George A. Rozgonyi
*: karoui@eos.ncsu.edu

Department of Materials Science and Engineering,
North Carolina State University, Raleigh, North Carolina 27695-7916

ABSTRACT

The oxygen precipitation in high purity CZ-silicon for ULSI is investigated with regard to the LO-HI and HI-LO-HI annealing processes used for denuded zone formation. The precipitation can be treated as a stochastic phenomenon and described using the chemical Rate Equations (RE) for small precipitates and Fokker-Planck Equation (FPE) for larger size domain. These key equations are connected to the point defect continuity equations (CE). The latter describe the depth and time dependency of the point defect concentrations inside the wafer. This paper presents a robust, stable and accurate numerical simulation of oxygen precipitation and annihilation in silicon. The main parameter, precipitate size distribution, is calculated as a function of depth and time. A C^{++} parallel program was developed and implemented on the Cray T3E Scalable Parallel Computer. MPI message passing interface was used for the inter-processor communication. The simulation results are compared to the experimental data obtained by FTIR[1] and OPP[2] measurement.

Keywords: Silicon, Precipitation, Fokker-Planck, Rate-Equation, Parallel Program

1 PHYSICAL MODEL

The kinetic of oxygen precipitation in CZ-Si is described by the rate equations (RE) [1] and uses the homogenous nucleation principle. These provide the evolution of the size distribution of oxygen precipitates, SiO_2 clusters containing n oxygen atoms. The precipitates grow at a rate $g(n,x,t)$ and concurrently dissolve at a rate $d(n,x,t)$. The size distribution function is related to the oxygen atom fluxes as follows

$$\frac{\partial f(n,x,t)}{\partial t} = J(n,x,t) - J(n+1,x,t) \quad (\mathrm{n} = 2,3,...) \tag{1}$$

$$J(n,x,t) = g(n-1,x,t).f(n-1,x,t) - d(n,x,t).f(n,x,t) \tag{2}$$

$g(n,x,t)$ and $d(n,x,t)$ are expressed using Gibbs free energy of the system [12]. Note that $\mathrm{f}(1,\mathrm{x},\mathrm{t}) = \mathrm{C_O}(\mathrm{x},\mathrm{t})$ where $C_O(x,t)$ is the oxygen interstitial concentration.

Precipitates of single SiO_2 molecule up to $n_{\max} = 10^{11}$ oxygen atoms are considered. The number of RE is equal to the maximum precipitate size and thus must be limited, practically to $n_0 = 20$. Above this number the discrete representation RE is extended by the Fokker-Planck equation (FPE), which is a representation of the system in a continuum size domain [1]

$$\frac{\partial f(n,x,t)}{\partial t} = -\frac{\partial I(n,x,t)}{\partial n} \tag{3}$$

$$I(n,x,t) = -B(n,x,t).\frac{\partial f(n,x,t)}{\partial n} + A(n,x,t).f(n,x,t) \tag{3a}$$

$$B(n,x,t) = g(n,x,t) + d(n,x,t)/2 \tag{3b}$$

$$A(n,x,t) = g(n,x,t) - d(n,x,t) - \frac{\partial B(n,x,t)}{\partial n} \tag{3c}$$

where A, B are respectively the advective and diffusion terms and are function of the growth and dissolution rates. The diffusion term describes a random change in precipitate size, which results in a flux from highly populated to less populated sizes. The advective term describes the growth of the precipitates for lowering the system total free energy.

Oxygen, self-interstitials and vacancies (O, I, V) interact at the matrix-precipitate interface according to the reaction

$$(1+2a)Si + 2O + 2b\,V \leftrightarrow SiO_2 + 2a\,I + stresss \tag{4}$$

Where a and b are the fraction of I and V emitted or absorbed per precipitated oxygen atom. The system is then diffusion-reaction limited. The point defects concentration in the matrix is given by the following continuity equations [1]

$$\frac{\partial C_X}{\partial t} = D_X \frac{\partial^2}{\partial x^2} C_X + p\frac{\partial}{\partial t} C_{O,OP} + qC_X + C \quad \text{with } X = O,I,V \tag{5}$$

Where $p = 1$; $q = 0$; $C = 0$ if $X = O$

$p = a$; $q = -k_{IV} C_V$; $C = k_{IV} C_I^{eq} C_V^{eq}$ if $X = I$

$p = -b$; $q = -k_{IV} C_I$; $C = k_{IV} C_I^{eq} C_V^{eq}$ if $X = V$

C_O^{eq}, C_I^{eq}, C_V^{eq} point defect equilibrium values, and

k_{IV} is the rate for the generation of Frenkel pairs.

The precipitated oxygen concentration $C_{O,OP}$ is given by

$$C_{O,OP} = \sum_{2}^{n_0} n f(n,x,t) + \int_{n_0+1}^{n_{\max}} n f(n,x,t)\,dn \tag{6}$$

[1] FTIR: Fourier Transform Infra Red spectroscopy
[2] OPP Oxygen Precipitate Profiler

2 INITIAL AND BOUNDARY CONDITIONS

CE: The initial oxygen concentration $C_O(t=0)$, is a known process parameter, measured by FTIR. We have assumed $C_I(t=0)=C_I^{eq}$ and $C_V(t=0)=1E15\,cm^{-3}$ corresponding to the vacancy rich wafers used [10, 13]. At the surface of the wafer, point defect concentrations were assumed at their equilibrium value

$$C_X(x=0,t)=C_X^{eq} \qquad \text{for} \qquad \text{X} = (\text{O}, \text{I}, \text{V}) \tag{7}$$

whereas at the half of the wafer thickness (200-300 μm) the Neumann condition ($\partial C_X/\partial x = 0$) is applied.

FPE: Available experimental tools cannot provide the size distribution in the as-grown wafer, due to the small precipitate size, two orders below their detection limit. Therefore a power law variation of the initial size distribution is used [8].

The FPE requires two boundary conditions in n-space. The fluxes calculated with RE and FPE must match at n_0

$$J(n_0,x,t)=I(n_0,x,t) \tag{8}$$

The no-flux condition at n_{max} expresses the mass conservation. It provides a mean for limiting the precipitate size. Also it encompasses the stability feature of big precipitates. The number, n_{max} value previously mentioned, is estimated from TEM[3] measurements [9].

$$I(n_{\max},x,t)=0 \tag{9}$$

3 NUMERICAL METHOD

The oxygen precipitation is a time dependent problem, in two dimension (n, x)-plane. The FPE and CE are parabolic in regard to the space coordinate and precipitate size. The unknowns in this problem are the size distribution f(n,x,t) for $n \geq 2$ and the point defect concentrations $C_O(x,t), C_I(x,t), C_V(x,t)$. To ensure the convergence for very long annealing times, implicit finite difference methods were used for the two sets of equations. These are known for their convergence and unconditional stability [4,5]. Nevertheless, this stability is at the price of greater computational complexity since we have to solve for the coupled PDEs, a large set of simultaneous linear systems.

Tests realized with Crank-Nicholson Method (CNM) showed that finite oscillations close to the boundary values, often occur when the time step increases. CNM is not dissipative [5] which prevents short wavelength noise from decaying away. These oscillations were found to be due to a large differences of several orders in the point defect diffusivity ($D_I >> D_V >> D_O$). The use of Douglas implicit approximation [4] for the CE, has removed these oscillations. Douglas implicit method is similar to CNM except that additional terms utilized in the central-difference series make it more accurate in space. The Douglas method approximates the CE at the point $\{i\Delta x, (j+1/2)\Delta t\}$ by

[3] TEM: Transmission Electron Microscopy

$$(1-6r)u_{i-1}^{j+1}+(10+12r-\frac{1}{2}q)u_i^{j+1}+(1-6r)u_{i+1}^{j+1}=(1+6r)u_{i-1}^{j}$$
$$+\left(10-12r+\frac{1}{2}q\right)u_i^{j}+(1+6r)u_{i+1}^{j}+p(v_i^{j+1}-v_i^{j})+Ck\Delta t \tag{10}$$

with $r=D_X\dfrac{\Delta t}{\Delta x^2}$ and $0 \leq i \leq 130$

Douglas scheme is unconditionally stable and has a local truncation error of $O(\Delta t^2)+O(\Delta x^4)$.

For FPE the n-space operators is approximated using the fully implicit Chang-Copper (CC) method [2,3], which was proved to be a robust finite difference scheme for this kind of equation [3]. The CC method proposes the following differentiation scheme for Eq.3 and uses a centered difference on the diffusion term and a weighted difference on the advective term

$$\frac{f_i^{j+1}-f_i^{j}}{\Delta t}=-\frac{I_{i+1/2}^{j+1}-I_{i-1/2}^{j+1}}{\Delta n_i} \qquad \text{for} \quad n_0 \leq n_i \leq n_{\max} \tag{11}$$

The flux at the point $n_{i+1/2}$ is sampled as following

$$I_{i+1/2}^{j+1}=A_{i+1/2}f_{i+1/2}^{j+1}-B_{i+1/2}\frac{f_{i+1}^{j+1}-f_i^{j+1}}{\Delta n_{i+1/2}} \tag{12}$$

where $f_{i+1/2}^{j+1}=(1-d_{i+1/2})f_{i+1}^{j+1}+d_{i+1/2}f_i^{j+1}$

$$j_{i+1/2}=\frac{j_i+j_{i+1}}{2},\ j=A,B,n,d \quad \textit{and} \quad \Delta n_{i+1/2}=n_{i+1}-n_i$$

This scheme adjusts the weight $d_{i+1/2}$ so that the differenciation of the advective term is always "upwind". This property makes the CC method first-order accurate in both space and time. d_m monotonically decreases from 1/2 to 0. The value of d_m given below, ensures the mass conservation property, as well as the positiveness of the solution, for all values of Δt and Δn

$$d_m=\frac{1}{w_m}-\frac{1}{\exp(w_m)-1} \tag{13}$$

where $w_m=\dfrac{A_m}{B_m}\Delta n_m \qquad n_0 \leq n_m \leq n_{\max}$

Subsequent choice of the size-step becomes only dictated by the accuracy. In order to build up the solution we need to scan small sizes as well as reach very large ones; this is made possible by using a logarithmic size mesh which resulted in a significant decrease in the actual computation time. Note that the linear systems obtained from CE and RE/FPE are tridiagonal and are solved by the Gauss elimination algorithm.

4 AUTOMATIC TIME-STEP CONTROL

A critical issue arises from long time annealing [9] and the need of a common time stepping procedure for CE and FPE sets of equations, which must preserve the accuracy.

In most physical problems formulated by parabolic PDE, the solution tends to a steady state, which is understood as a gradual decay of the exponential terms [4,5]. Therefore, as the annealing time progresses, larger steps must be taken. The program must adjust the time step in order to maintain a specific accuracy. The method of step doubling and local extrapolation was applied. We select the step size in order to keep the "per step local error" below a predetermined value.
First, the time integration goes over a compound of two half steps $\Delta t/2$ to predict the value u_i at point x_i; Then, the solution $\tilde{u}_i$ at x_i is estimated using a time step equal to Δt. If a differentiation scheme is of order p, its local truncation error is $T = c\Delta t^p$. Richardson extrapolation [4] has been used for the estimation of T at x_i, which was found to be

$$T = \frac{2^{p-1}(\tilde{u}_i - u_i)}{\left(1 - 2^{p-1}\right)} \tag{14}$$

The estimated error is compared to a tolerance e. If T is much greater than e, the compound step is rejected and a smaller time step is taken. A random number weights the current time step in order to eventually break the stationary states of the algorithm.

5 PARALLEL IMPLEMENTATION

At each time step we solved 130 linear systems of size 150x150 for RE and FPE and three linear systems for the CE. A Matlab version developed for Sun Sparc workstation indicated that few days will be necessary to entirely solve the time and depth dependent problem of the first step annealing (8 hours) in the LO-HI process. This serious problem was overcome by developing a parallel version implemented on the Cray T3E supercomputer. The Cray T3E is a parallel scalable MIMD [6] system including 64 processors totaling one GFLOPS computation power. The entire program was rewritten in C^{++} language using the Linear Algebra Package LAPACK. MPI message passing interface was used for the inter-processor communication.

The three distinct stages required for the design of the parallel process are: task partitioning, data dependence and communication analysis, and mapping [6,7]. Since the physical problem is computationally extensive and possesses several coupled systems, we have adopted a coarse grain functional partitioning. We focused on task decomposition, which results in a specific data stream for each processor such that the overall communications are minimized. A parallel task T_k $(1 \le k \le 130)$ compute $g(n,x,t)$, $d(n,x,t)$ and one linear system associated to FPE for a given depth x. A master-slaves mapping approach was used. Each task T_k is assigned to a slave processor in a "round robin" way. The master processor solves the CE, estimates the local truncation error and the new time step. The tasks T_k are independent [7]. The communication occurs only between the master and each slave. The program is parameterized such that it can run on any number of processors p. As p increases, the CPU time decreases toward an asymptotic value, Fig. 1, which is reflected on the speedup and the efficiency, Fig. 2. The chosen mapping led to a linear speed-up up to 16 processors, with a maximum efficiency [6] equal to 0.9. The speed-up and efficiency presented in Fig. 2 are independent of the thermal process since a static scheduling [7] was adopted.

6 EXPERIMENTAL RESULTS

Simulation of the HI-LO-HI process and the effect of the nucleation step on subsequent annealing steps was carried out and compared to the experimental data by T. Sasaki [10]. The annealing cycles considered are:

1. LO-HI anneals: a nucleation step at 750°C for different times t_{LO}= 8, 16, 36, and 64 hr followed by a growth anneal at 1050°C for 16 Hrs in nitrogen.
2. HI-LO-HI anneals: the same as above but prior to the nucleation step an out-diffusion pre-anneal was made at high temperature, 1250°C for 1 hour in Ar.

Note that the critical radius of a precipitate increases with the heat treatment temperature. Precipitates larger than the critical size grow while small ones shrink. The simulated size distribution behavior appears to be in good agreement with the experimental results, compare the shape of the curves in Fig. 3 (a) and (b). The simulated distributions for the LO-HI annealing in Fig. 3 (a) show that the precipitate density increases with the nucleation time (duration of the LO step). The precipitate size is affected by the LO annealing times while the HI time is maintained identical for all LO-HI tests. As the LO time increases, the size distribution after the HI step tends to reach a maximum value, compare the 36 hr and 64 hr distribution in one side and the 16 hr in the other side, see Fig. 3 (a). The saturation indicates that the precipitates tend to reach their equilibrium size, this situation occurs in a similar way for all precipitate sizes.

Figure 4 (a) reports the simulation results of the three-step annealing treatments, HI-LO-HI thermal process. Like the published experimental data, only the half of the size distribution curves are presented. Note the 5 nm shift in the integrated density. This is attributed to the OPP new technique, which has required a deconvolution of the signal in order to separate the true size distribution from the ghost image [11]. Both the physics and the deconvolution algorithm are not yet established, and might need improvements.

Since a large precipitate critical size is associated with the annealing temperature of 1250°C, during the HI step (pre-annealing treatment) a large part of the latent micro-defects shrinks and vanishes. Unlike the LO-HI, in the first step of the HI-LO-HI process, the grown-in defects dissolve, and during the subsequent step (LO) oxygen clusters homogeneously giving rise to new nuclei. These are expected to have lower density than the grown-in sites, which were used for a heterogeneous nucleation in the LO-HI cycle. The difference in nuclei site density resulted in a lower precipitate density calculated after the three step annealing. This is supported by the experimental data [10], where the variation in the size distributions is larger than in the case of LO-HI anneal (compare for instance the gaps between 36 hr and 64 hr curves for both cycles in Fig. 3 (b) and Fig. 4 (b)).

7 CONCLUSION

Parallel computing on the Cray T3E appeared valuable and efficient for simulating precisely the oxygen precipitation

in CZ-Silicon which is a time, depth and size dependent problem of and to handle long annealing times. Our simulation matches the main features of available experimental data. We found that the actual precipitation in the LO-HI process does not depart from a homogeneous nucleation. The simulation was instrumental in calibrating the newly developed OPP technique for micro- and nano-defects measurements in high purity silicon.

Acknowledgement: The authors are indebted to the North Carolina Supercomputer Center which has awarded the first author with 1200 CPU hours on the CrayT3E.

REFERENCES

[1] M. Schrems, in Semiconductors and Semimetals, ed. F. Shimura, **42**, 391, 1994.
[2] J. S. Chang, and G. Cooper, J. of Comp. Physics, **6**, 1, 1970.
[3] B. T. Park and V. Petrosian, AspJS, **103**, 255, 1996.
[4] G. D. Smith, Numerical solution of partial differential equations, finite difference methods, Oxford Univ. Press, 1985.
[5] John C. Strikwerda, Finite difference schemes and partial differential equations, Chapman & Hall, 1989
[6] Hesham El-Rewini, Ted Lewis, and Hesham Ali, Task Scheduling in Parallel and Distributed Systems,Prentice-Hall, 1994.
[7] Faouzia S. Karoui, Ph.D. Thesis, Univ. of Tunis, 1996.
[8] S.M. Hu, Defects in Semiconductors. MRS, 333, 1981.
[9] T. Ono, E. Asayama, H. Horie, M. Hourai, K. Sueoka, H. Tsuya, G. A. Rozgonyi, J. Electrochem. Soc., **146** (6), 2239, 1999.
[10] T. Sasaki, T. Ono, and G. A Rozgonyi, Electrochem. and Solid-State Letters, 2, 1999.
[11] K. Nakai, A. Matsumura, and H. Harada, Rev. Sci, Inst., **69**, 3283, 1998.
[12] W. A. Tiller et al, J. Appl. Phys., 64 (1), 375, 1988.
[13] S. Dannefaer and T. Bretagnon, J. Appl. Phys, 77(11), 5584,1995.

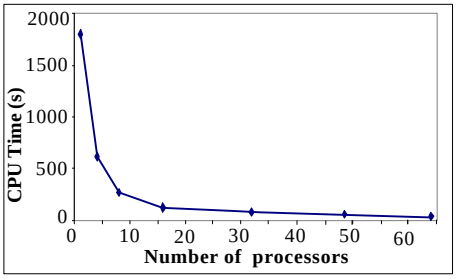


Fig.1: Performances of the parallel program for the simulation of oxygen precipitation in silicon: Execution time

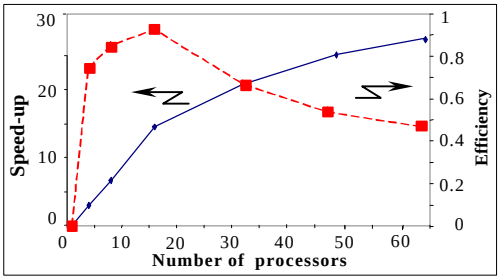
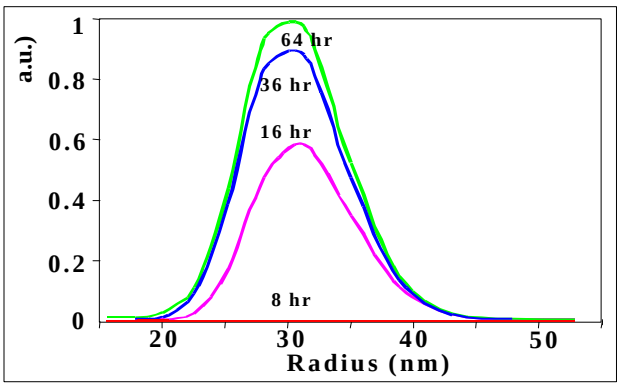
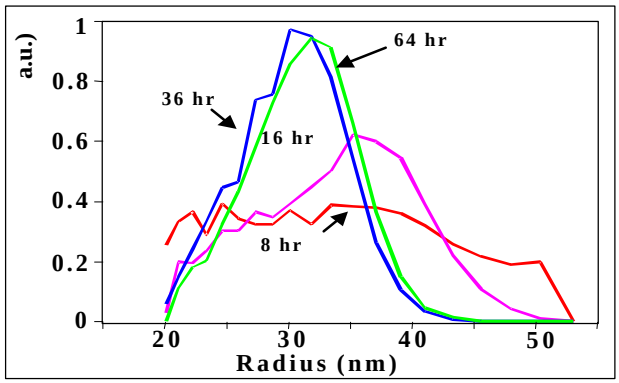


Fig.2: Performances of the parallel program for the simulation of oxygen precipitation in silicon: Speed-up and Efficiency.

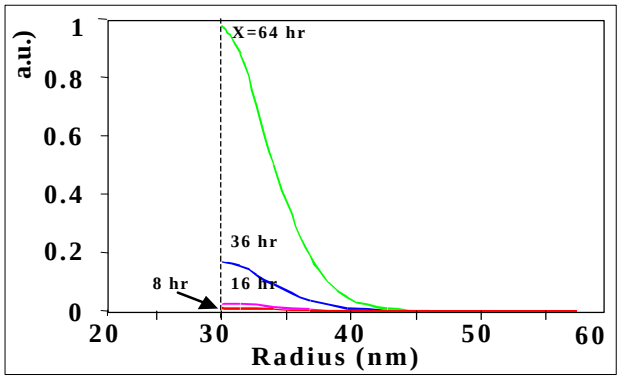


(a)

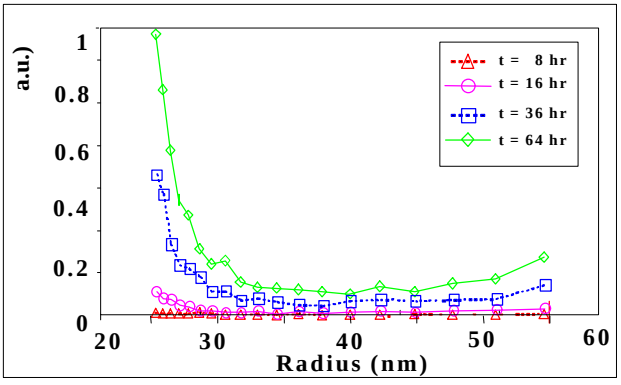


(b)

Fig.3: Oxygen precipitate profiles for the **LO-HI** annealing process. (a): simulated, (.b): measured with OPP [10].



(a)



(b)

Fig. 4: Oxygen precipitate profiles for the **HILOHI** annealing process. (a): simulated, (.b): measured with OPP [10].