*Journal of The Electrochemical Society*, **151** (1) G1-G7 (2004)
0013-4651/2003/151(1)/G1/7/$7.00 © The Electrochemical Society, Inc.



# A Simplified Picture for Transient Enhanced Diffusion of Boron in Silicon

**M. Y. L. Jung, R. Gunawan, R. D. Braatz, and E. G. Seebauer**[z]

*Department of Chemical Engineering, University of Illinois, Urbana, Illinois 61801, USA*

In recent years, transistor junction formation in complementary metal oxide semiconductor devices by ion implantation has encountered serious limitations due to transient enhanced diffusion (TED) during the annealing step. Current models of TED rely heavily on detailed simulations of the complex diffusion-reaction network that governs TED, and often rely on fitted parameters whose values are uncertain. The present work uses a more rigorous set of rate parameters obtained from a maximum likelihood estimation to develop a relatively simple analytical treatment of boron TED that is capable of estimating the degree of profile spreading and the temperature at which TED should begin to occur significantly. The treatment suggests that reduction of TED should focus on implantation schemes and heating programs designed to decrease the number of clusters slightly smaller than the very largest.
 [DOI: 10.1149/1.1628238] 

Manuscript submitted February 18, 2003; revised manuscript received July 1, 2003. Available electronically November 18, 2003.

Forming extremely shallow pn junctions in Si-based microelectronic logic devices is becoming increasingly critical as device dimensions continue to diminish. For example, advanced complementary metal oxide semiconductor devices will require junction depths $X_j$ between 13 and 22 nm in the source and drain extension regions by 2005.[1] Current technology for junction formation relies almost exclusively on ion implantation to introduce dopants into the substrate. Although junction depths can be made shallower by reducing the implant energy, the effectiveness of this approach has been limited by the need to anneal the resulting structure to activate the dopant electrically and to eliminate implant-induced defects in the crystal structure. As long as they exist, these defects mediate exceptionally fast transient enhanced diffusion (TED) of the implanted dopants, often leading to significant spreading of the original dopant profile.

Because experiments to measure TED kinetics are expensive and sometimes difficult to interpret unambiguously, many researchers have resorted to detailed modeling to aid process development. There have been several attempts in the literature to develop a comprehensive physical picture for TED,[2-5] and such attempts have been incorporated into various widely used profile simulators.[6] However, the existing models suffer important deficiencies.

One problem is that the predictive capability of most TED models outside their tested range is subject to serious doubt. Many elementary kinetic steps contribute to the experimental observable: typically a dopant depth profile obtained by secondary ion mass spectroscopy (SIMS). Hence, all models include numerous rate parameters, many of whose values are developed primarily according to their ability to fit experimental SIMS profiles. The large number of parameters provides many degrees of freedom for fitting, impeding the ability to develop a unique set.

A second problem is that profile simulators are not very conducive to developing simple, intuitive explanations for key features of TED that retain quantitative utility. Many kinetic processes take place in ways that vary strongly both in space and time, and the dominant elementary steps in the overall diffusion-reaction network vary accordingly. Thus, clear generalizations are often difficult to make. These difficulties are particularly pronounced when there is uncertainty about key rate parameters.

This work attempts to alleviate these problems by developing a relatively simple analytical treatment for TED that incorporates no fitted rate parameters. The parameters have been calculated in our laboratory[7,8] using maximum likelihood (ML) estimation[9] which surpasses simple fitting of experimental profiles in terms of mathematical rigor and justifiability of the derived parameter set. The resulting model has yielded analytical expressions for estimating the temperature at which TED should begin to occur significantly and the degree of profile spreading. An important conclusion of the work is that reducing TED should involve adjusting both the initial conditions for the clusters (during implant) and the subsequent heating procedure to maximize the number of clusters having dissociation energies just under the maximum.

[z] E-mail: eseebaue@uiuc.edu

## Basic Model

*Kinetic parameters and treatment of clusters.*—Table I and II reproduce the kinetics we have developed previously[7,8] regarding interstitial diffusion and interaction with clusters. It is well known that clusters of atoms form during postimplant annealing. These clusters release free interstitial atoms that drive TED. There is considerable evidence that clusters can comprise pure B, pure Si, or mixed B-Si. Due to the number of cluster species that can be tracked by the process simulator we used, Tables I and II show rate parameters limited to a maximum of five atoms for pure Si and mixed B-Si clusters. We doubt that this restriction imposes serious limitations on simulation accuracy because experimental observations from spike rapid thermal annealing (RTA) of sub-keV implanted show no evidence for large clusters.[10-12] Regarding boron, there is no evidence for pure B clusters larger than two atoms that involve interstitial atoms. Instead, B interstitials seem to be more effectively captured by mixed B-Si.[13,14] The model includes only point defects and excludes extended defects such as dislocation loops and {311} defects, which may also interact with interstitials in some cases.[11]

*Weaknesses of published kinetic parameters.*—It seems prudent to highlight a weakness in using published rate parameters that to our knowledge has not been discussed in the TED literature. The very short estimated time constants for $(B_s - Si_i)$ complex formation, of order tens to hundreds of picoseconds, call into question the kinetic validity of the rate expressions and parameters being used. The expressions in Tables I and II tacitly assume a rate law based on a transition state formulation,[15] in which an atom attempts to cross a well-defined potential barrier at a frequency comparable to a lattice vibrational frequency. Such formulations are usually applied under conditions at modest temperatures where $kT$ is much less than the barrier height. At 1050°C, however, $kT = 0.11$ eV, which approaches the same magnitude as several barriers in Tables I and II. Under such conditions, it becomes questionable whether the potential barrier remains well defined.

Van Vechten has described this problem at length with respect to vacancies in Si,[16] but the basic picture remains valid for interstitial motion as well as reactions with clusters. Well above the Debye temperature (670 K in Si), the neighbors of the atom of interest fluctuate so strongly in position that they occasionally conspire to

**Table I. Activation energies for interstitial diffusion and cluster association.**

| Reaction | Symbol[a] | Activation energy (eV) | Reference | Method[b] |
|---|---|---|---|---|
| $B_i$ diffusion | $E_{\text{diff},B_i}$ | 0.37 ± 0.04[c] | 7 | ML |
| $Si_i$ diffusion | $E_{\text{diff},Si_i}$ | 0.72 ± 0.03[c] | 7 | ML |
| $B_i + Si_s \rightarrow (B_s - Si_i)$ | $E_{\text{ki}}$ | 0.50 ± 0.1 | 7 | ML |
| $B_m - Si_n + B_i \rightarrow B_{m+1} - Si_n$, $n,m \geqslant 0$[d] | $E_{\text{assoc,B}}$ | 0.37 ± 0.04[e] | 7 | Assumed = $E_{\text{diff}\,B_i}$ |
| $B_m - Si_n + Si_i \rightarrow B_m - Si_{n+1}$, $n,m \geqslant 0$[d] | $E_{\text{assoc}}$ | 0.72 ± 0.03[e] | 7 | Assumed = $E_{\text{diff},Si_i}$ |

[a] For clarity in focus on cluster effects, the present work uses slightly different notation than related publications from this laboratory.[7] Where two symbols are given here, the latter symbol appears in Ref. 7.
[b] ML = Maximum likelihood estimation.
[c] The pre-exponential factor for this diffusion has been assumed to be $1 \times 10^{-3}$ $cm^2/s$.[7]
[d] For $m \geqslant 1$, one of the boron atoms is presumed to be substitutional. Also, $m$ and $n$ must obey $m + n \geqslant 2$ and $m + n \leqslant 4$. Finally, if $n = 0$ then $m$ is assumed to obey $m \leqslant 2$. (No pure boron clusters larger than dimers form.)
[e] The pre-exponential factor for this association reaction has been assumed to be $3 \times 10^{-10}$ $cm^2/s$,[7] regardless of cluster size. Also, if two free Si interstitials recombine, the stochiometric factor of 2 has been neglected.

create an open ''tube'' for motion with essentially no barrier. If the atom in question possesses sufficient kinetic energy (of order 1 eV), that atom simply zips through the tube before it closes. The resulting rate parameters depend upon the energy and entropy associated with rapid translational motion, and exhibit activation energies much larger than barriers computed by quantum methods (at 0 K) and pre-exponential factors much larger than vibrational frequencies. This theory has to our knowledge, not been refined for calculations of the sort required here, and so our treatment will stick with the classical low temperature picture. Nevertheless, the results of this paper and others employing low temperature physics should be treated with due circumspection.

**Table II. Activation energies for cluster dissociation.[a]**

| Composition | Cluster size | Species liberated | Symbol | Activation energy (eV) | Reference | Method[b] |
|---|---|---|---|---|---|---|
| Pure B | 2 | B | $E_{2,B}$ | 1.70 ± 0.07 | 8 | ML |
| Pure Si | 2 | Si | $E_2$ | 1.41 ± 0.03 | 8 | ML |
| | 3 | Si | $E_3$ | 2.2 | 8 | Linear interpolation |
| | 4 | Si | $E_4$ | 3.0 | 8 | Linear interpolation |
| | 5 | Si | $E_{\text{large}}$ | 3.7 ± 0.1 | 8 | ML |
| Mixed B-Si | 2[c] | B | $E_{2,\text{mix}\rightarrow B} = E_{\text{ko}}$ | 0.50 | 8 | From dopant activation[d] |
| | 2[e] | Si | $E_{2,\text{mix}\rightarrow Si} = E_{\text{dis}}$ | 0.59 ± 0.06 | 7 | ML |
| | 3 | B | $E_{3,\text{mix}}$ | 2.2 | 8 | Assumed = $E_3$ |
| | 3 | Si | $E_{3,\text{mix}}$ | 2.2 | 8 | Assumed = $E_3$ |
| | 4 | B | $E_{4,\text{mix}}$ | 3.0 | 8 | Assumed = $E_4$ |
| | 4 | Si | $E_{4,\text{mix}}$ | 3.0 | 8 | Assumed = $E_4$ |
| | 5 | B | $E_{\text{large,mix}}$ | 3.5 | 35 | Density functional theory |
| | 5 | Si | $E_{\text{large,mix}}$ | 3.5 | 35 | Density functional theory |

[a] All pre-exponential factors are assumed equal to $6 \times 10^{12}$ $s^{-1}$.[7]
[b] ML = Maximum likelihood estimation.
[c] This represents the kick-out reaction $(B_s - Si_i) \rightarrow B_i + Si_s$.
[d] ML method in Ref. 7 yielded 1.05 eV. The value in the table is calculated from published data for dopant activation (equivalent to solid solubility) as discussed in text.
[e] This represents the dissociation reaction $(B_s - Si_i) \rightarrow B_s + Si_i$.

*Simulation method*.—Calculations were performed using the profile simulator FLOOPS 2000.[a,17] That simulator solves the coupled mass balance equations for interstitials, vacancies, and clusters. Those equations have the general form for species *j*

$$\frac{\partial N_j}{\partial t} = -\frac{\partial J_{N_j}}{\partial x} + G_j \qquad [1]$$

where $N_j$ denotes concentration and $G$ a net generation rate. The flux $J$ comprises terms due to both diffusion and drift in response to electric fields. The electric fields are obtained by solution of Poisson's equation. FLOOPS was implemented with the rate expressions and parameters shown in Tables I and II together with no-flux surface boundary conditions for all species[b,18] and no surface band bending. Concentrations of charged interstitial B and Si species were computed according to Fermi statistics as described elsewhere.[19]

Due to the number of cluster species that can be tracked by FLOOPS 2000, cluster sizes were limited to four atoms for pure B and Si clusters and five for mixed B-Si clusters. The entire distribution of cluster dissociation energies has been captured in the present model by equating the size-5 dissociation energy to that of the large interstitial clusters (~3.5-3.7 eV). This model telescopes the entire cluster dissociation cascade into a computationally manageable set of events. According to a detailed parameter sensitivity analysis that appears elsewhere,[8] the junction depth and degree of boron activation are not sensitive to the dissociation energy of size 5 clusters, but rather to the dissociation of intermediate-sized clusters. This finding, together with experimental observations from spike RTA of sub-keV implanted wafers indicating that large clusters do not form,[20] suggests that the limitation on cluster size does not impose serious restrictions on physical interpretations drawn from the simulation.

Initial conditions on the profiles for $Si_i$ were set to track the local concentration of total implanted boron. For total boron we employed experimental as-implanted profiles as initial conditions, with a fixed fraction of one-fifth of the total boron in substitutional sites, in accord with the suggestion of Caturla *et al.*[21] and Kobayashi *et al.*[22]

Figure 1 shows an example of heating program employed, drawn from a larger experimental data set to which the simulations were compared. The experimental wafers, obtained from International Sematech, were implanted with $2 \times 10^{15}$ ions/cm$^2$ of B at 0.60 keV with 0° tilt. Results for a comparable set of experiments also have been reported by Downey *et al.*[23] Simulations were performed using the temperature program shown in Fig. 1 at a ramp rate of 150°C/s and maximum temperature of 1050°C. Ramp rate effects are considered in another publication.[24]

Figure 2 shows a typical example comparing simulated total boron profiles and experiment. The agreement is excellent. The kinetic parameters were taken from published experimental and computational estimates using the ml method,[7,8] which gives the most likely values of the parameters as weighted average of the published estimates. Note that there has been no free parameter fitting in developing the ML parameter set. For descriptive purposes, the profile is divided into α, β, and γ regions that encompass, respectively, the dopant peak near the surface, the relatively flat region where most TED takes place, and the deep bulk beyond the "shoulder" in the diffused profiles. The junction at a $[B_s] = 1 \times 10^{18}$ cm$^{-3}$ falls within the γ region.

[a] FLOOPS 2000, by Mark E. Law of the University of Florida and Al Tasch of the University of Texas/Austin.

[b] No-flux boundary conditions represent an approximation that suffer deficiencies as discussed in Ref. 18. However, the simulations fit experimental data well and, in the regime of the simulations, the conclusions of the paper are insensitive to the details of these conditions.

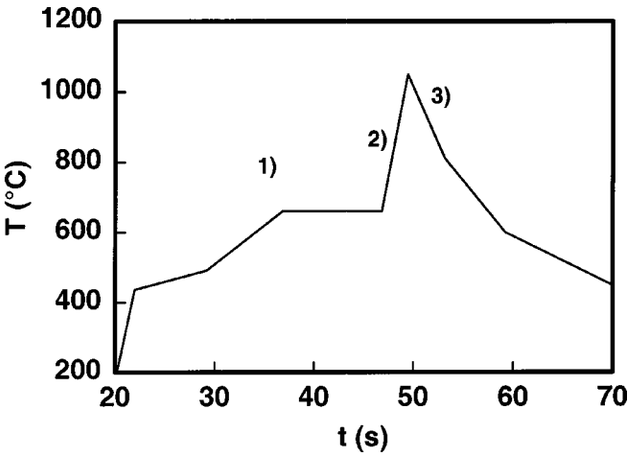


**Figure 1.** Temperature program for simulated spike anneals. (1) temperature stabilization between about 400 and 660°C, (2) main spike with β = 150°C/s, (3) maximum temperature $T_M$ = 1050°C, (4) radiative rampdown, initial rate of 64°C/s.

### Qualitative Model Predictions for Profile Spreading and Activation

It is well known[2-5] that TED and dopant activation behavior is determined primarily by the interplay between interstitials (both Si and B) and the reservoirs that render interstitials immobile, *i.e.*, lattice sites and various clusters between interstitials and substitutional boron [such as $(B_s - Si_i)$ and $(B_s - B_i)$] and pure Si interstitial clusters. As-implanted profiles would broaden very rapidly even at low temperature if there were no reactions that immobilize the interstitials in these reservoirs.

Based on our improved set of parameters, we can summarize the main features of the diffusion/activation process during implantation and annealing as follows.

*Implantation and temperature stabilization*.—During implantation at room temperature, numerous interstitials of B and Si are created. In agreement with previous discussions in the literature,[3] the rate parameters in Table I indicate that the interstitials diffuse as soon as they appear. Indeed, there exists both computational[25] and

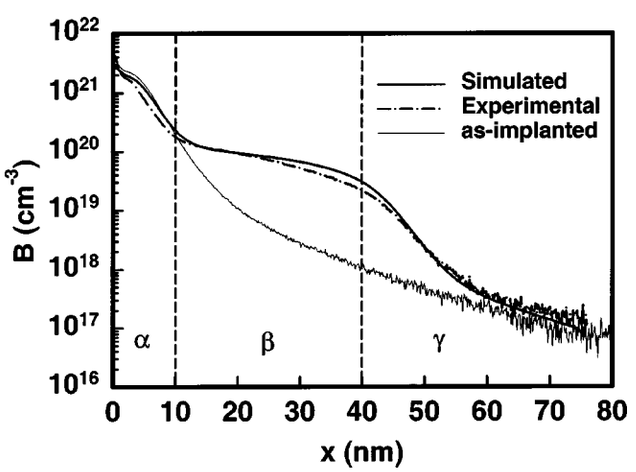


**Figure 2.** Simulated and experimental boron profiles for the temperature program shown in Fig. 1. For descriptive purposes, the profile is divided into α, β, and γ regions.

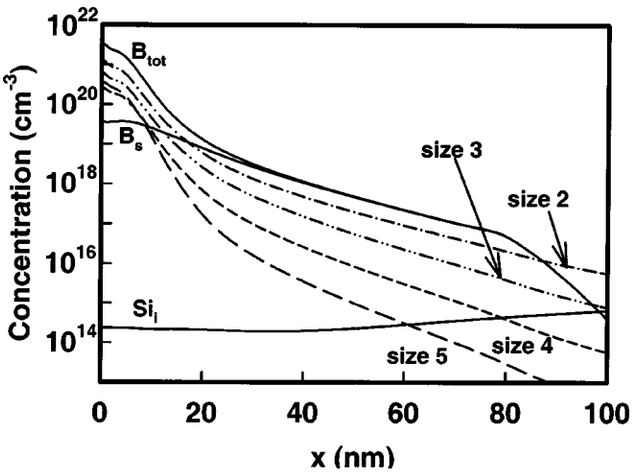


**Figure 3.** Concentrations of total boron, $B_s$, $Si_i$, and clusters ranging from size two to five as a function of position at the end of the 20 s resting period at 170°C just after implantation in Fig. 1. The cluster concentrations in this and the following two figures represent the sum of pure B (which includes size 2 only), pure Si, and mixed B-Si clusters. Note that all clusters are present in large numbers. In all regions in this and the following figures, $[B_i]$ roughly tracks $[Si_i]$, but remains two to five orders of magnitude lower.

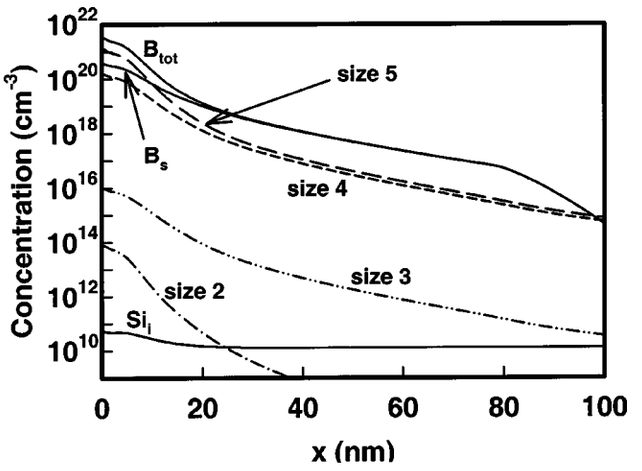


**Figure 4.** Concentrations of total boron, $B_s$, $Si_i$, and clusters ranging from size two to five as a function of position at the end of the temperature stabilization step at 660°C. Note that clusters of sizes two and three have largely disappeared, and that $[Si_i]$ is low and rather flat throughout the profile.

experimental[26-28] evidence suggesting that B and Si diffuse significantly even at temperatures much lower than room temperature. Random-walk motion can be described by

$$x^2 = 6Dt \qquad [2]$$

where $x$ denotes diffusion length and $t$ diffusion time. Simple estimates of interstitial diffusivities based on this equation and the data of Table I show that almost all B and Si interstitials have accreted into clusters (or entered lattice sites) at the end of implantation and before any heating.

Figure 3 shows cluster concentrations at various sizes as a function of depth. For simplicity, the concentrations represent the sums of pure B, pure Si and mixed B-Si clusters at the nominal sizes 2 through 5. Note that in this scheme the pure Si clusters of a given nominal size actually contained one more interstitial defect than the other cluster types (which constellate around a substitutional B atom). However, Fig. 3 seeks primarily to show concentrations of clusters having similar dissociation energy, for which nominal cluster size was a good proxy in our model.

Very early in the heating cycle on the way to the stabilization temperature, large numbers of clusters exist at all sizes. $B_i$ remains immobilized until dissociation of clusters other than ($B_s$ − $Si_i$) began to release $B_i$. All dimers and trimers have dissociation energies low enough for this process to proceed at or below typical stabilization temperatures on the order of 660°C. The dissociation energy for trimers probably varies somewhat with their composition and whether they release B or Si. Given the data of Table II, however, the time constant for release decreases to 1 s near 280°C for pure Si dimers, near 400°C for pure B dimers, and near 590°C for trimers. By 660°C, the concentrations of these species decrease by roughly five orders of magnitude in our simulations, as shown in Fig. 4. The interstitials released by the disintegration of dimers and trimers diffuse until recapture by larger clusters having larger dissociation energies. The mean spacing between larger clusters remains fairly small, however, on the order of 1, 5, and 27 nm in the α, β, and γ regions, respectively. Thus, in α and β regions where most boron resides, very little boron motion is possible, in agreement with much literature.

One might expect somewhat more boron motion in the γ region where the cluster spacing is significant. However, the kick-in reaction of $B_i$ via ($B_s$-$Si_i$) inhibits this motion as follows. The ($B_s$-$Si_i$) complex forms with the rate constant $k_{ki}$ and the time constant $1/k_{ki}$. Insertion of this time constant into Eq. 2 together with the diffusivity for interstitial boron yields

$$x^2 = \frac{6D_{\mathrm{diff,B_i}}}{k_{\mathrm{ki}}} \qquad [3]$$

If the vibrational frequencies characteristic of hopping and kick-in are assumed to be identical (as is often done in treatments of this type[29]), then Eq. 3 becomes

$$x^2 = 6\lambda^2 \exp[(E_{\mathrm{ki}} - E_{\mathrm{diff,B_i}})/kT] \qquad [4]$$

where $k$ denotes Boltzmann's constant and λ denotes the hop length of an interstitial. In the present case, λ equals roughly 0.3 nm.[7] At 660°C, Eq. 4 indicates that $B_i$ can diffuse only 1.5 nm before this reaction takes place. Once the complex forms, it quickly redissociates by liberating either $B_i$ or $Si_i$. Liberation of $B_i$ frees the boron atom again.

Liberation of $Si_i$ immobilizes the boron as $B_s$, however, until another $Si_i$ comes along to reform the ($B_s$-$Si_i$) complex and provide a mechanism for releasing $B_i$. The branching ratio $b$ between the dissociation reaction forming $B_i$ and $Si_i$ is

$$b = \frac{r_{2,\mathrm{mix}\rightarrow\mathrm{Si}}}{r_{2,\mathrm{mix}\rightarrow\mathrm{Si}} + r_{2,\mathrm{mix}\rightarrow\mathrm{B}}} \qquad [5]$$

At 660°C, this branching ratio is about 0.25, meaning that, on average, an immobilizing dissociation event that releases $Si_i$ occurs after only three dissociation events releasing $B_i$. Put another way, boron is free to move only for a time equal to $3/k_{ki}$, meaning that immobilization occurs after a diffusion distance given by

$$x^2 = 18\lambda^2 \exp[(E_{\mathrm{ki}} - E_{\mathrm{diff,B_i}})/kT] \qquad [6]$$

This equation yields $x = 2.6$ nm at 660°C.

$B_s$ is freed again only after reassociating with a $Si_i$ atom that happens by, which occurs with a time constant equal to $(k_{\mathrm{assoc}}[Si_i])^{-1}$. Simulations show that at 600°C, $[Si_i]$ in the γ region is about $10^{10}$ cm$^{-3}$, leading to a time constant of about 80 min. This time is much longer than the total spike cycle, meaning that once $B_i$

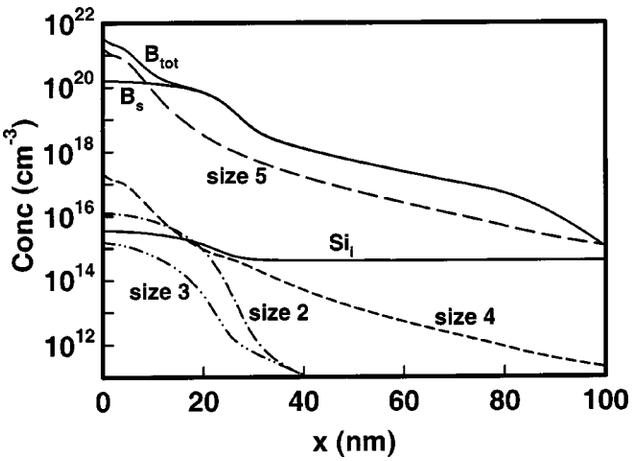


**Figure 5.** Concentrations of total boron, $B_s$, $Si_i$, and clusters ranging from size two to five as a function of position at the top of the spike at 1050°C. Note that TED is taking place, and that $[Si_i]$ has risen dramatically.

is immobilized in the γ region (after $B_i$ diffusing only about 2-3 nm), it remains immobile during the remainder of temperature stabilization.

*Main spike*.—During the main spike, the rising temperature dissociates clusters of progressively larger sizes. The profile begins to spread, as shown in Fig. 5, and the concentration of Si interstitials rose dramatically. Note that clusters of size four had mostly dissociated, leaving only the largest clusters intact.

Clusters [excluding ($B_s$-$Si_i$)] dissociate with energies ranging from a low of 1.4 eV for dimers up to 3.7 eV for large clusters. The number of dissociation pathways between those extremes is actually quite large if the structural isomers of the B-containing clusters are taken into account. This large number of pathways is plausible to describe according to a nearly continuous distribution of dissociation energies. At a given temperature during the spike, there exist three classes of clusters. The first, with low dissociation energy, have large dissociation rate constants but exist in negligibly small concentrations because most of them have already dissociated. The contribution to $[B_i]$ and $[Si_i]$ is also negligible. Clusters in the second class, with high dissociation energy, exist in large concentrations but have negligibly small rate constants for dissociation and therefore also do not contribute appreciably to $[B_i]$ and $[Si_i]$. The third class has both intermediate rate constants and intermediate concentrations and does contribute. Thus, during the main ramp, clusters of progressively larger sizes and therefore dissociation energies disappear. Note that once a cluster begins to dissociate, complete dissolution follows quickly because subsequent dissociation events have lower activation energies.

The freed interstitials may exchange with ($B_s$-$Si_i$) or briefly coalesce into new clusters, but then quickly accrete onto larger clusters having large dissociation energies. This process may be considered as a form of Ostwald ripening. Such ripening and the quasi-equilibrium it implies has been discussed at length in conjunction with constant-temperature experiments,[2,3,30,31] but to our knowledge has not been rigorously justified for the conditions characteristic of fast ramps. The formation of ($B_s$-$Si_i$) and addition of interstitials to clusters is extremely rapid during the main spike. At a typical maximum spike temperature of 1050°C, for example (with $[B_s] \sim 10^{20}$), the time constants for creating ($B_s$-$Si_i$) are 13 ps and 18 ns for $B_i$ and $Si_i$, respectively. Dissociation proceeds at comparable speeds. At 1050°C, time constants for interstitial release from ($B_s$-$Si_i$) are 13 and 30 ps for B and Si, respectively. Thus, concentrations of free interstitials may be considered to be in quasi-equilibrium with those of complexes and small clusters, even under

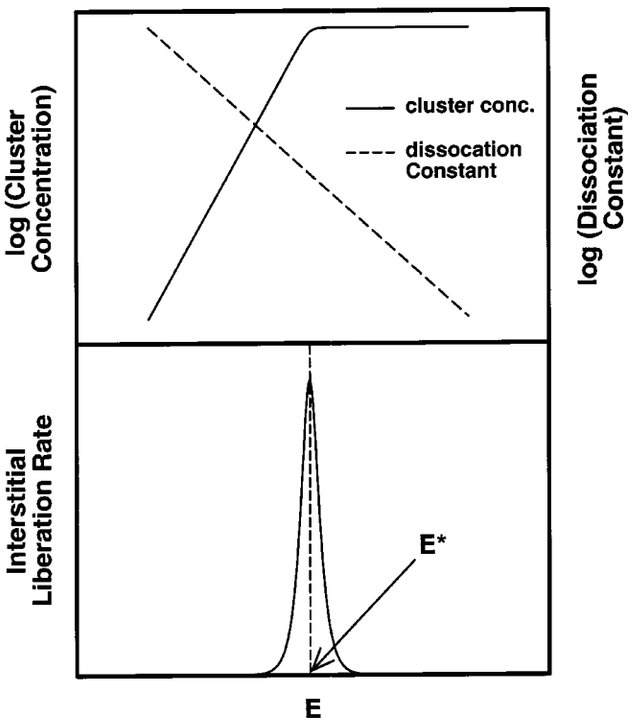


**Figure 6.** Schematic diagram showing three classes of clusters as a function of dissociation energy at a fixed temperature during the ramp. Clusters from only a narrow range of energies near $E^*$ contribute significantly to the instantaneous liberation rate of interstitials because the concentration of clusters with $E < E^*$ is too low, and the dissociation rate of clusters with $E > E^*$ is too low.

the extremely fast ramping conditions of "flash" rapid thermal processing, where heating rates approach $10^6$ C/s.[32]

There are many cluster dissociation pathways due to the varying stochiometries of larger clusters. It is therefore plausible to mathematically approximate the activation energies of those pathways using a nearly continuous distribution of dissociation energies. This description permits the use of simplifying mathematical expressions developed for related kinetic systems in which dissociation is described by a distribution of activation energies.

In such systems, there exist three classes of dissociating species during a linear ramp: almost fully dissociated, presently dissociating, and not yet dissociated. At any given temperature, only the class that is presently dissociating contributes significantly to interstitial release. This point is illustrated schematically in Fig. 6, which shows that this class constitutes only a small fraction of the total distribution. This kinetic situation has been treated extensively in the literature in the context of gas desorption from surfaces. A clean, closed-form analytical expression can be obtained[33] connecting each temperature in a linear ramp with the dissociation energy $E^*$ of the most active dissociating species. The number of clusters at a given value of $E^*$ controls the rate of interstitial release at a given temperature.

The following conditions must be satisfied to obtain the analytical expression:

1. The rate of reassociation of interstitials with the most actively dissociating clusters can be neglected. Examination of the association rates in Table I together with cluster concentrations suggested by simulations indicates that this condition holds in the present case.

2. The distribution must have a sufficiently broad standard deviation σ, that obeys $\sigma/kT > 1.5$. For clusters, the distribution of dissociation energies has a standard deviation on the order of 1 eV,

while kT ~ 0.1 eV at 1300 K. Thus, the breadth condition is easily satisfied in the present case.

3. The dissociation entails a single-step reaction. Although cluster dissolution involves a sequence of desorption events, release of the first atom has by far the highest activation energy. Thus the subsequent dissolution cascade is very rapid (because dissociation energy decreases substantially with size) and liberates interstitials essentially instantaneously. Thus, the overall dissociation for practical purposes can be treated as a single step.

Under conditions 1-3, it can be shown that $E^*$ obeys the following transcendental equation[33]

$$(E^*/kT - 1/2)\exp(E^*/kT - 1/2) = \nu T/\beta \qquad [7]$$

For simplified calculation of $E^*$, this transcendental expression can be replaced[34] by the analytical expression

$$E^*/kT = 1/2 + Y - \ln Y + (\ln Y)/Y - (2 - \ln Y)(\ln Y)/(2Y^2) \qquad [8]$$

where $Y \equiv \ln(\nu T/\beta)$. The approximation has negligible error on the order of $[\ln Y/(Y)]^3$, which is $1.5 \times 10^{-3}$ for a typical value of $Y$ near 30.

At $\beta = 50$ K/s and a typical spike maximum of 1050°C = 1323 K, Eq. 7 or 8 yield $E^* = 3.4$ eV, which is just below the dissociation energy for the maximum-sized mixed Si-B cluster. Raising the temperature much higher than 1050°C dissociates these large clusters and creates a flood of interstitials that greatly accelerate TED, as is well known from experiments. Note in Fig. 5 that all clusters except for the largest had largely dissociated by the top of the spike.

### Estimating TED Distances

These considerations lead to a simple way to estimate likely TED distances from cluster concentrations and $[Si_i]$. In other publications,[7,8] we showed that diffusion of boron during TED is most likely controlled by free $B_i$ rather than $(B_s - Si_i)$ or $(B_s - B_i)$. As our prior discussion suggests, there are two primary kinds of reservoirs that can appreciably hinder motion of $B_i$: nondissociated clusters and the lattice (as $B_s$). The importance of the nondissociated clusters scales with their concentration, while that of the lattice scales as the concentration of interstitial silicon.

*Hindrance of $B_i$ motion by nondissociated clusters*.—At a given temperature, the mean distance over which $B_i$ can diffuse freely equals the mean spacing between nondissociating reservoirs. By about 1000°C, the Ostwald ripening process described earlier has eliminated small clusters in favor of those having dissociation energies near the maximum of 3.5 eV (for B-containing clusters) or 3.7 eV (for pure Si clusters). Near the top of the spike, the total concentration of clusters that can add interstitials decreases dramatically. Figure 7 shows this effect computationally. In the simulations, clusters sizes two, three, and four can accrete interstitials, but not size five, because the model constrains clusters to sizes less than five. In the vicinity of the top of the spike the concentrations of size two, three, and four clusters in the β region where most TED takes place decrease to a sum of about $4 \times 10^{15}$ cm$^{-3}$, meaning that the spacing rises to roughly $(4 \times 10^{15}\ \text{cm}^{-3})^{-1/3} = 63$ nm. In the γ region near the junction, the corresponding number is 520 nm. These numbers are too large to account for the modest change in junction depth (14 nm) observed in Fig. 2.

Note that the artificiality of the simulations due to cluster size limitation to five or less can be instructive. In contrast to real clusters that can both add and lose interstitials, the largest clusters in the simulations can dissociate but not accrete interstitials. The primary effect of the size five clusters is therefore to accurately reproduce the huge release of interstitials if the temperature gets too high. The largest clusters do not dissociate appreciably in the simulations, so their concentration remains high, near $6 \times 10^{20}$, $8 \times 10^{18}$, and $3 \times 10^{16}$ cm$^{-3}$ in the α, β, and γ regions, respectively, at 1050°C. If these clusters were active in accreting boron interstitials, the arguments given above would yield 1.2, 5, and 32 nm for the respective diffusion distances. Mobility in the α region is difficult to assess from the boron profiles, since most boron there is tied up in large clusters that never dissociate. However, the numbers for β and γ regions form a range that brackets the observed change in junction depth of 14 nm, and therefore might lend support to the cluster mechanism for inhibiting boron diffusion.

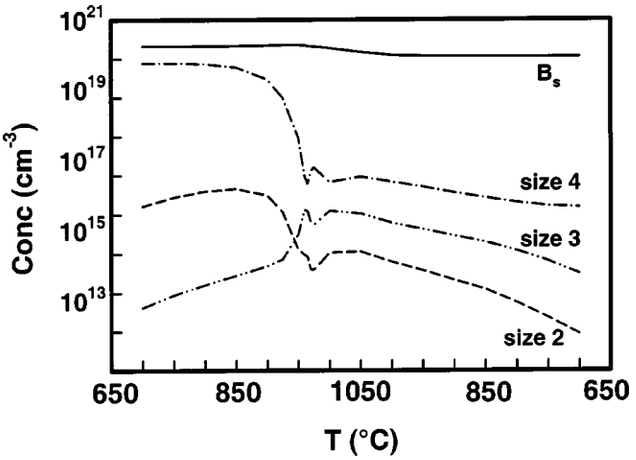


**Figure 7.** Concentrations of $B_s$, $Si_i$, and the sum of clusters that can add interstitials (two through four) as a function of $T$ for the temperature program shown in Fig. 1. Cluster concentrations are averaged over the α region and include the sum of the concentrations of pure B, pure Si, and mixed B-Si species. Note the large decrease in clusters above about 900°C. These qualitative behaviors also appear in β and γ regions, although all concentrations are lower and $[B_s]$ rises rather than falls.

It is not presently known whether large clusters (on their way to resembling {311} defects) can accrete boron, or what the dissociation energy for release might be. Hence, an appeal to the physics of cluster formation cannot resolve the issue. However, closer examination of the profiles in Fig. 2 provides a clue. Although the junction depth in the γ region changes by 14 nm, the degree of motion actually increases toward the β region. For example, profile spreading at the constant-concentration level of $2 \times 10^{19}$ cm$^{-3}$ level is close to 30 nm. At higher concentrations approaching those of the α region, spreading decreases again. However, the argument that nondissociated clusters constitute the primary brakes on boron diffusion would predict that spreading should increase uniformly with increasing depth, because cluster spacing increases this way. Thus, spreading should be smaller in the β region and larger in the γ region. This trend is observed neither in experiments (where large clusters in principle can accrete boron) nor in simulation (where they cannot), thereby arguing against the importance of clusters as primary inhibitors of boron diffusion.

*Hindrance of $B_i$ motion by the lattice reservior*.—When this mechanism dominates, boron moves while it is a free interstitial (slightly impeded by exchange with $(B_s - Si_i)$ via the pathway that releases $B_i$), but is rapidly immobilized when $(B_s - Si_i)$ releases $Si_i$ and can move again only after lengthy periods of waiting for association with another free Si interstitial. The distance over which boron can move before immobilization comes from the branching ratio $b$ defined in Eq. 5 and first used in Eq. 6. Equation 6 can be written in the following more general form

$$x^2 = 6\lambda^2 \exp[(E_{ki} - E_{diff,B_i})/kT](1 - b)/b \qquad [9]$$

During temperature stabilization, this ratio is about 0.25; it rises to 0.35 at 1050°C, so that $x$ decreases to 1.7 nm.

However, $[Si_i]$ increases to an average of about $5 \times 10^{14}$ and $2 \times 10^{14}$ $cm^{-3}$ in the β and γ regions in the vicinity of the spike maximum, so that the time constants for liberation decrease to $(k_{assoc}[Si_i])^{-1} = 0.0036$ and 0.009 s, respectively. Thus, during the approximately 1 s that $T$ remains within about 50°C of the top, liberation takes place roughly 280 and 110 times. Equation 9 can be further modified to take this liberation into account by defining a time, $t_{max}$, as a characteristic time over which the wafer remains near the peak temperature. Then

$$x^2 = 6\lambda^2 k_{assoc}[Si_i] t_{max} \exp[(E_{ki} - E_{diff,B_i})/kT](1 - b)/b \quad [10]$$

Equation 10 yields a total distance of motion equal to 27 and 17 nm in the β and γ regions near the top of the spike. The numerical estimates agree well (10-20%) with the degrees of profile spreading in these regions. Unlike the cluster braking mechanism, lattice braking also correctly predicts the decrease in profile spreading when moving deeper into the bulk from the β to the γ region.

Equation 10 helps to interpret why TED is observed only at the top of the spike. During heating, interstitial silicon is produced mainly by the dissociation of clusters. Since there is no lattice reservoir for $Si_i$ equivalent to that for $B_i$ (because the lattice holds primarily Si atoms), $Si_i$ diffuses rapidly over large distances in the β and γ regions where cluster concentrations are modest. Thus, as shown in Fig. 3-5, profiles for $[Si_i]$ are flat compared with most other species in the implanted system. During the main spike, $[Si_i]$ rose dramatically because of increasing release rate from clusters (due to the larger average number of Si atoms per cluster and the larger number of stoichiometric permutations of mixed clusters) and decreasing net capture rates due to the decreasing total number of clusters. This increase in $[Si_i]$ permitted $x$ to grow in Eq. 10.

## Implications for TED Reduction

This picture suggests that a key to reducing TED is to adjust both the initial conditions for the clusters (during implant) and the subsequent heating procedure to minimize the number of clusters having dissociation energies just below the maximum. Such minimization would decrease the release rate of silicon interstitials near the top of the spike where TED takes place. Such minimization can be accomplished by aiming to produce primarily (*i*) very small clusters, or (*ii*) large clusters. Producing very small clusters may be accomplished by very rapid implant at well below room temperature where interstitial motion is inhibited. A sudden, mild warm up would encourage formation of a large number of very small clusters. The problem, of course, is that conventional heating programs up to near 1050°C to improve dopant activation would dissociate the vast majority of these clusters (except a few that happen to grow by Ostwald-like ripening), producing a flood of $Si_i$ that would promote TED. Producing large clusters is best accomplished by a slow Ostwald-ripening anneal prior to the final spike. This procedure has the best chance of working if mixed B-Si clusters have slightly lower dissociation energy than pure Si clusters. Thus, appreciable amounts of boron could be freed while leaving most Si within clusters. We have assumed such a difference of activation energy in our model, although the details of the size and stoichiometry dependence of cluster dissociation are poorly known. A better understanding of this dependence would help greatly in designing suitable heating programs.

## Conclusion

Using a set of rate parameters based on firmer grounds than previously available, we developed a relatively simple analytical treatment of TED that is capable of (*i*) predicting the temperature at which TED should begin to occur, and (*ii*) estimating the TED distance in terms of the concentration of clusters slightly smaller than the largest. The picture suggests that reduction of TED should focus on decreasing the number of clusters in this size range in favor of larger clusters. A better understanding the size and stoichiometry dependence of cluster dissociation would help greatly in exploiting such effects.

## Acknowledgments

This work was supported by NSF (CTS 98-06329 and CTS 02-03237) and by International Sematech.

## References

1. See, *The International Technology Roadmap for Semiconductors*, http://public.itrs.net/Files/2001ITRS/Home.htm
2. L. Pelaz, M. Jaraiz, G. H. Gilmer, H.-J. Gossmann, C. S. Rafferty, D. J. Eaglesham, and J. M. Poate, *Appl. Phys. Lett.*, **70**, 2285 (1997).
3. L. Pelaz, G. H. Gilmer, H.-J. Gossmann, C. S. Rafferty, M. Jaraiz, and J. Barbolla, *Appl. Phys. Lett.*, **74**, 3657 (1999).
4. A. D. Lilak, M. E. Law, K. S. Jones, M. D. Giles, E. Andideh, M.-J. Caturla, T. Diaz de la Rubia, J. Zhu, and S. Theiss, *Tech. Dig. - Int. Electron Devices Meet., IEEE*, **1998**, 493.
5. A. D. Lilak, S. K. Earles, K. S. Jones, M. E. Law, and M. D. Giles, *Tech. Dig. - Int. Electron Devices Meet., IEEE*, **1997**, 493.
6. For example, see, M. E. Law and S. M. Cea, *Comput. Mater. Sci.*, **12**, 289 (1998).
7. M. Y. L. Jung, R. Gunawan, R. D. Braatz, and E. G. Seebauer, *J. Electrochem. Soc.*, **150**, G766 (2003).
8. R. Gunawan, M. Y. L. Jung, R. D. Braatz, and E. G. Seebauer, *J. Electrochem. Soc.*, **150**, G758 (2003).
9. J. V. Beck and K. J. Arnold, *Parameter Estimation in Engineering and Science*, Wiley & Sons, New York (1977).
10. L. H. Zhang, K. S. Jones, P. H. Chi, and D. S. Simons, *Appl. Phys. Lett.*, **67**, 2025 (1995).
11. G. Hobler and C. S. Rafferty, *Mater. Res. Soc. Symp. Proc.*, **568**, 123 (1999).
12. A. D. Lilak, S. K. Earles, M. E. Law, and K. S. Jones, *Appl. Phys. Lett.*, **74**, 2038 (1999).
13. M. J. Caturla, M. D. Johnson, and T. D. de la Rubia, *Appl. Phys. Lett.*, **72**, 2736 (1998).
14. S. Chakravarthi and S. T. Dunham, *J. Appl. Phys.*, **89**, 3650 (2001).
15. See, for example, K. J. Laidler, *Chemical Kinetics*, 3rd ed., Chap. 4, Harper & Row, New York (1987).
16. J. A. Van Vechten, *Phys. Rev. B*, **38**, 9913 (1988).
17. See, M. Law, http://www.swamp.tec.ufl.edu/
18. H.-H. Vuong, C. S. Rafferty, S. A. Eshraghi, J. Ning, J. R. McMacken, S. Chaudhry, J. McKinley, and F. A. Stevie, *J. Vac. Sci. Technol. B*, **18**, 428 (2000).
19. M. Y. L. Jung, R. Gunawan, R. D. Braatz, and E. G. Seebauer, *J. Appl. Phys.*, Submitted.
20. L. H. Zhang, K. S. Jones, P. H. Chi, and D. S. Simons, *Appl. Phys. Lett.*, **67**, 2025 (1995).
21. M. J. Caturla, M. Foad, and T. D. de la Rubia, in *Proceedings of the International, Conference on Ion Implantation Technology*, Vol. 2, IEEE, p. 1018 (1999).
22. H. Kobayashi, I. Nomachi, S. Kusanagi, and F. Nishiyama, *Mater. Res. Soc. Symp. Proc.*, **669**, J 5.3 (2001).
23. D. F. Downey, S. W. Falk, A. F. Bertuch, and S. D. Marcus, *J. Electron. Mater.*, **28**, 1340 (1999).
24. M. Y. L. Jung, R. Gunawan, R. D. Braatz, and E. G. Seebauer, *J. Electrochem. Soc.*, **150**, G838 (2003).
25. W.-C. Lee, S.-G. Lee, and K. J. Chang, *J. Phys.: Condens. Matter*, **10**, 995 (1998).
26. G. D. Watkins, J. W. Corbett, and R. M. Walker, *J. Appl. Phys.*, **30**, 1198 (1959).
27. A. Hallen, N. Keskitalo, L. Josyula, and B. G. Svensson, *J. Appl. Phys.*, **86**, 214 (1999).
28. V. Privitera, S. Coffa, F. Priolo, K. K. Larsen, and G. Mannino, *Appl. Phys. Lett.*, **68**, 3422 (1996).
29. See, for example, G. C. Bond, *Catalysis by Metals*, Academic Press, New York (1962).
30. P. A. Stolk, H.-J. Gossmann, D. J. Eaglesham, D. C. Jacobson, C. S. Rafferty, G. H. Gilmer, M. Jaraiz, J. M. Poate, H. S. Luftman, and T. E. Haynes, *J. Appl. Phys.*, **81**, 6031 (1997).
31. N. E. B. Cowern, G. Mannino, P. A. Stolk, F. Roozeboom, H. G. A. Huizing, J. G. M. Van Berkum, F. Cristiano, A. Claverie, and M. Jaraiz, *Mater. Sci. Semicond. Process.*, **2**, 369 (1999).
32. D. M. Camm, J. C. Gelpey, T. Thrum, G. C. Stuart, and J. K. Elliott, in *Proceedings of the 10th Conference on Advanced Thermal Processing of Semiconductors*, RTP 2002, p. 5 (2002).
33. E. G. Seebauer, *Surf. Sci.*, **316**, 391 (1994).
34. Z. Du, A. F. Sarofim, and J. L. Longwell, *Energy Fuels*, **4**, 296 (1990).
35. X.-Y. Liu, W. Windl, and M. P. Masquelier, *Appl. Phys. Lett.*, **77**, 2018 (2000).