

Stress-induced dislocations in silicon integrated circuits

by P. M. Fahey
S. R. Mader
S. R. Stiffler
R. L. Mohler
J. D. Mis
J. A. Slinkman

Many of the processes used in the fabrication of silicon integrated circuits lead to the development of stress in the silicon substrate. Given enough stress, the substrate will yield by generating dislocations. We examine the formation of stress-induced dislocations in integrated circuit structures. Examples are presented from bipolar and MOS-based integrated circuit structures that were created during developmental studies. The underlying causes of oxidation-induced stress and the effect on such stress of varying oxidation conditions are discussed. The knowledge thus gained is used to explain dislocation generation during the formation of a shallow-trench isolation structure. The importance of ion-implantation processes in nucleating dislocations is illustrated using structures formed by a deep-trench isolation process and a process used to form a trench capacitor in a DRAM cell. The effect of device layout geometry on dislocation generation is also examined. We show how TEM observations can be used to provide more information than

solely identifying those process conditions under which dislocations are generated. By combining TEM observations with stress analysis, we show how the sources of stress responsible for dislocation movement can be identified.

1. Introduction

Given enough stress, a silicon substrate will yield by generating dislocations. Although it would certainly be preferable for this not to occur during the fabrication of silicon integrated circuits (ICs), its occurrence appears to be an unavoidable by-product of the processes used to fabricate such circuits; dislocations often appear at some point during the development of a fabrication sequence. Once present in the silicon substrate, dislocations can lead to charge leakage and electrical shorting between elements—effects that can seriously degrade or prevent device operation [1, 2].

There are many sources of stress that arise during IC fabrication processes. Some important examples are the imbedding of materials with thermal expansion coefficients different from that of silicon, deposition of films with intrinsic stress, and oxidation of nonplanar surfaces. A

©Copyright 1992 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.

recent paper by Hu [3] gives an excellent overview of these and other stress-related problems. In the present paper, we present examples of stress-induced dislocations that occur during the fabrication of advanced bipolar and MOS-based integrated circuits. This is a timely subject because of the growing importance of stress-induced defects in evolving IC processes.

For each succeeding generation of integrated circuits, two major trends are evident: The packing density of circuits on a chip increases, and device dimensions are reduced. Increasing packing density permits a greater level of integration per chip; reducing critical device dimensions leads to improved device performance for both MOS and bipolar structures. Unfortunately, the two trends lead to processes in which stress levels increase.

Higher packing densities are achieved by developing devices that occupy smaller areas of the silicon substrate and by packing the devices closer together. However, such scaling is subject to the constraint that devices must be electrically isolated from one another by isolation regions. In general, the reduction of the area required for isolation has been found to result in larger substrate stresses. In addition, as the area occupied by an active device continues to shrink, more of the device in each isolated cell is in proximity to the edges of its isolation regions, and some of the largest stresses develop at these locations. Scaling of device dimensions may lead indirectly to increased substrate stress. To maintain critical device dimensions (e.g., the channel length of a MOSFET or the base width of a bipolar transistor), control of dopant diffusion is essential. A large degree of such control is accomplished by minimizing the temperatures of thermal processing. But for oxidation processes, lowering the temperature of oxidation leads to higher stress levels in the substrate [4–6]. The fundamental reason for this is that an oxide growing on a nonplanar silicon surface must constantly deform, and oxides become more resistant to strain as temperatures are lowered. As an oxide thus becomes more rigid, a greater amount of the stress that develops during oxidation is accommodated by the development of strain in the substrate. Given the pervasive use of oxidation steps in IC fabrication, oxidation-induced stress is likely to become an increasingly important concern in the continuing drive toward lower-temperature processing.

Since stress is responsible for the unwanted appearance of dislocations, a question that naturally arises is how much stress a silicon wafer can tolerate before dislocations are generated. There is no simple answer to this question. A variety of studies indicate that stress levels of the order of 10^7 dynes/cm² or higher should be considered significant. In practice, however, the strengths of silicon wafers vary depending on a number of factors—in particular, the oxygen content and thermal history of the

wafer [7–9]. Also, the generation of stress-induced dislocations is a two-step process. Dislocations must first be nucleated and must then grow or move into regions where they affect devices. The stress necessary to nucleate dislocations depends greatly on the particular process of nucleation. Nuclei may be found in the as-grown wafer [7–9] or may be introduced into the wafer during processing. Two common process steps that assist the nucleation of dislocations are oxidation and ion implantation. Oxidation generates silicon self-interstitials, which can coalesce preferentially in strained silicon [1]. Vanhellemont et al. [10, 11] have discussed in detail the homogeneous nucleation of dislocations by condensation of self-interstitials. Ion implantation disrupts the crystalline structure of the silicon lattice and creates an excess of point defects. Although the details of the nucleation process following ion implantation are not thoroughly understood, it is well established that implantation damage can lead to the nucleation of dislocations. The effect of implantation on dislocation generation in stressed material can be quite dramatic; we demonstrate this with examples from bipolar and DRAM technologies (Sections 4 and 5). Once nucleated, dislocations can move great distances under an applied shear stress, by the process of glide. (The level of stress necessary to move dislocations depends, as in the case of nucleation, on oxygen content and thermal history [12].) Thus, dislocations initially created in a locally stressed area can propagate to other parts of a device cell (Sections 3 through 5).

The examples presented in this paper are taken from structures used in developing IC processes. During the developmental stage, process conditions are often varied to extreme cases, yielding important information regarding relevant process windows and information regarding process extendability: for example, extendability by placing devices closer together (vs. changing the layout design or isolation scheme) or by lowering processing temperatures and thereby minimizing dopant diffusion. We have chosen examples gathered from process experiments performed as part of a few different chip development efforts at several IBM laboratories. The examples illustrate the underlying causes of stress, how they cause dislocation generation, and how solutions to defect generation can be found.

In Section 2 we examine the underlying causes for the development of stress in oxide layers on silicon substrates and relate its development to oxidation parameters (temperature, pressure, ambience). In Section 3 we present an example illustrating the increased susceptibility of a process to defect generation as oxidation temperatures are lowered. The example explores dislocation generation resulting from a shallow-trench isolation process compatible with the requirements of a 16Mb DRAM technology. In addition to the temperature of oxidation, the importance of cell layout geometry is shown to be a

key factor in determining the onset of defect generation. The very important process of defect nucleation by ion implantation is discussed in Section 4, where we examine the effects of implantation on a deep-trench isolation process used in a bipolar technology. Damage created by ion implantation can develop into dislocations that glide large distances in stressed material. Substrates that otherwise show no signs of defects can become heavily dislocated after ion implantation. The geometry of cell layout again plays a key role in affecting this behavior. Finally, in Section 5 we present a detailed stress analysis for dislocation generation and propagation in a 4Mb DRAM process.

2. Oxidation-induced stress

Oxidation can introduce stress into silicon substrates in two different ways. The first is associated with the volume expansion of SiO_2 from silicon. A given volume of silicon will produce about 2.2 times as large a volume of SiO_2 . If the oxide is not free to expand, it pushes on whatever material is constraining its growth. This occurs frequently during IC fabrication, for example, during oxidation along the sidewalls of polysilicon-filled trench structures or at the edges of regions masked with Si_3N_4 . A second cause of stress is the strain that oxide layers experience when oxidation is performed on nonplanar structures. As a nonplanar silicon surface is oxidized, the oxide layer is constantly forced to stretch or contract as it grows out of the silicon surface.

How much stress is translated from an oxide layer into the substrate is determined by how easily the oxide deforms as it grows. In the following discussion, we consider the deformation of an oxide layer growing on a nonplanar substrate. The parameters that determine the ability of an oxide layer to deform in response to stress are examined, with particular attention focused on the effect of oxidation conditions on oxide stress.

• Development of stress in SiO_2

Upon oxidation, the growth of oxide at the Si/ SiO_2 interface pushes the oxide above it away from the Si surface. On a convex-shaped surface, this causes the oxide to experience a tangential tensile stress; on a concave-shaped surface, compressive stress results. We are primarily interested in how stress develops in the silicon substrate during an oxidation step. This requires foremost an understanding of how an SiO_2 film deforms in response to the stresses that arise during its growth. Our knowledge of such stress-induced deformation processes is not yet satisfactory, and the subject remains an active area of research. However, in the past few years there has been a growing consensus that the best starting point for understanding nonplanar oxide growth is to view SiO_2 as a

nonlinear viscoelastic material [13, 14]. In recent papers, Rafferty, Borucki, and Dutton [13] and Hu [15] have provided an excellent illustration of the viscoelastic properties of SiO_2 . They considered the simple case of oxidizing a cylinder of silicon. The geometry of this problem lends itself to relatively simple analyses of stress induced in a growing oxide, yet the oxidizing cylinder example brings out most of the essential features of oxidation-induced stress on nonplanar surfaces. Hu has modeled oxide stress on a cylinder under a variety of process conditions. Rafferty et al. have modeled oxide stress under a more limited set of conditions than Hu, but with a more realistic viscoelastic model for SiO_2 . In the following treatment we use the oxide model of Rafferty et al. and reexamine the model predictions of Hu for a variety of process conditions. We also use the oxidizing cylinder as a vehicle to demonstrate how different model assumptions used in process modeling programs manifest themselves in predicting oxidation-induced stress.

Consider a cross section of a silicon cylinder having a thin surface oxide layer, the outer surface of the oxide located at a radial distance r_o . When the cylinder is oxidized, the newly formed oxide at the silicon surface will push out the original oxide layer. After a given oxidation time, the outermost layer of oxide moves from r_o to r , stretching out in the process. The stretching process requires a tangential strain of $\epsilon = (r - r_o)/r_o$. If the oxide grows at a rate \dot{X}_{ox} , it is a simple matter to show that $\dot{\epsilon} = (1 - \beta)\dot{X}_{\text{ox}}/r_o$, where $\beta = 0.44$ is the amount of silicon consumed for a given thickness of oxide grown. Therefore, the strain rate of the outermost oxide layer on the cylinder is given by

$$\dot{\epsilon} = \frac{(1 - \beta)}{r_o} \dot{X}_{\text{ox}}. \quad (1)$$

The question is how this strain is accomplished. If the oxide is highly resistant to deformation, large stresses develop in the oxide during the oxidation and, consequently, also in the underlying silicon. On the other hand, if the oxide can be stretched easily, little stress is developed during the oxidation. In general, stress is accommodated by a combination of elastic deformation and viscous flow. Following the treatment by Hu [15] and Rafferty et al. [13], we illustrate the viscoelastic deformation process by analogy to a spring and dashpot (i.e., damped piston) in series—a combination known as a Maxwell element (Figure 1).

Initially, the system is assumed to be in a state of zero stress. Stress is developed by applying a force that moves the end of the spring. The total strain in the system is $\epsilon = (A - A_o)/A_o$, where A_o is the original position of A . If the end of the spring moves at a speed v , the strain rate of the system is $\dot{\epsilon} = v/A_o$. The total strain is the sum of elastic

and viscous deformation processes, i.e., $\epsilon = \epsilon_{\text{elastic}} + \epsilon_{\text{viscous}}$. The elastic strain is given by

$$\epsilon_{\text{elastic}} = \frac{(A - B)}{A_0} = \frac{\sigma}{G}, \quad (2)$$

where G is the spring constant, which we identify with the elastic modulus of rigidity of the oxide. The stress-strain relation for viscous flow is expressed as

$$\dot{\epsilon}_{\text{viscous}} = \frac{\dot{B}}{A_0} = \frac{\sigma}{\eta}, \quad (3)$$

where η is the viscosity of the dashpot (i.e., the viscosity of the oxide). Equations (2) and (3) lead to the differential equation

$$\frac{\dot{\sigma}}{G} + \frac{\sigma}{\eta} = \dot{\epsilon}. \quad (4)$$

Equation (4) is a simple but useful expression for understanding many of the important factors that determine evolution of stress in oxides during growth. By using values for G and η corresponding to SiO_2 and using the expression for $\dot{\epsilon}$ in Equation (1), relative levels of stress generated with different oxidation conditions can be investigated for the oxidizing cylinder example [13, 15]. For example, one can calculate the relative stress levels in different thicknesses of an oxide film grown at a given temperature, or the relative stress levels in an oxide film of a given thickness grown at different temperatures.

In the simplest case, both the growth rate and viscosity are assumed constant. In this case, Equation (4) has the solution

$$\sigma = \eta \dot{\epsilon} (1 - e^{-Gt/\eta}). \quad (5)$$

This equation states that the oxide starts at zero stress and asymptotically approaches a value of $\eta \dot{\epsilon}$. Physically, this means that the oxide deforms elastically in the initial stages of growth; then, as the oxide continues to grow, its stress is relieved by viscous flow. In steady state, the stress in the oxide is directly proportional to the oxidation rate times the viscosity. Obviously, the less viscous the oxide, the smaller the saturation value of its stress. Oxide viscosity decreases with higher temperatures and increasing hydroxyl content (see the papers by Hu [15] and Stiffler [14] for a summary of data). Thus, performing oxidations at the highest acceptable temperature and in a wet rather than a dry oxygen ambience would seem to be preferable. However, oxidation rates increase with increasing temperature and are higher in a wet oxygen ambience than in a dry oxygen ambience. Faster oxide growth means that the strain rate is higher and viscous flow must increase to avoid stress buildup. Predictions regarding such trade-offs in processing conditions can be obtained by solving Equation (4). In that regard, Equation

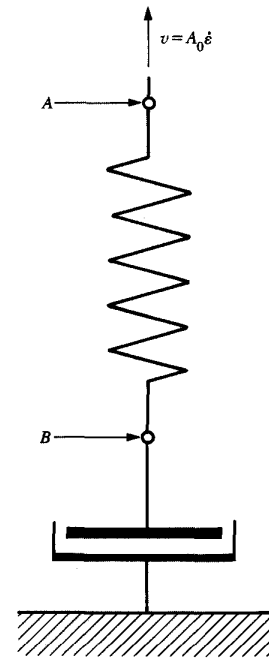
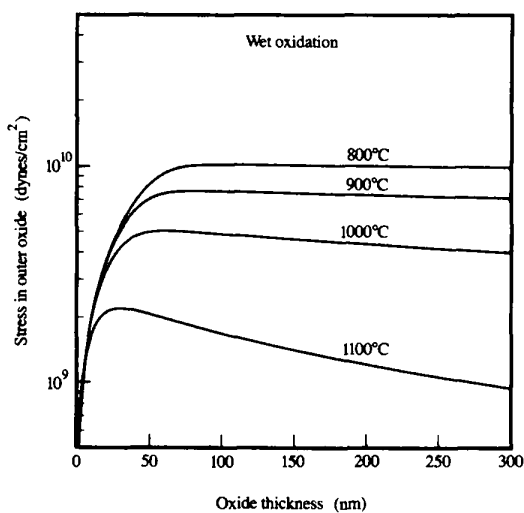


Figure 1

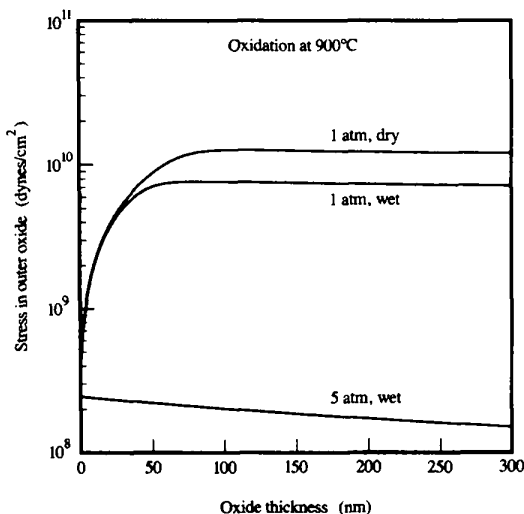
Viscoelastic model of oxide strain. Analogy is made to a spring and dashpot (i.e., damped piston) system.

(5) predicts that growing oxides at high temperatures in a wet oxygen ambience does indeed minimize stress levels, despite growth at relatively high rates. Hu [15] has shown that this same qualitative behavior occurs for the more general situation of linear-parabolic oxide growth. In addition, the analysis of Hu predicts that high-pressure steam oxidation should lead to further reductions in oxide stress from the levels of atmospheric oxidation—again, despite the increased oxidation rates at high pressure.

The above analysis applies under the assumption that oxide viscosity remains constant during growth. The viscosity of SiO_2 was considered a function of temperature and hydroxyl content only. However, it has been shown experimentally that the viscosities of glasses decrease under high stress [16–18]. Recent treatments [13, 14, 19, 20] of oxidation-induced stress have attempted to take account of this fact. Sutarjda and Oldham [19, 20], Rafferty et al. [13], and Stiffler [14] have favored the use of Eyring's model [21] to describe viscosity. In Eyring's model, the viscosity of the oxide decreases dramatically at



(a)



(b)

Figure 2

Calculated oxide stress as a function of oxide thickness at different conditions, for growth on a silicon cylinder, initially $1 \mu\text{m}$ in radius: (a) Stress in the outermost oxide layer for different oxidation temperatures. (b) Stress developed at 900°C for different ambient conditions.

high stresses. The high-stress behavior is described by the relation

$$\eta = \eta_0 \frac{(\sigma V_m / 2kT)}{\sinh(\sigma V_m / 2kT)}, \quad (6)$$

where η_0 is the zero-stress viscosity and V_m is a characteristic volume associated with the space cut out by a diffusing molecule during flow. Eyring [21] has analyzed the dependence of viscosity on stress from the viewpoint of absolute rate theory. However, the quantitative predictions of this atomistic approach do not agree well with experiment [16]. To obtain agreement with experiment, V_m has been viewed as a fitting factor. Donnadieu et al. [18] have empirically determined a value for V_m of approximately 180 \AA^3 .

As an example of how oxide stress varies as a function of process variables, we show in Figure 2 calculations for stress in the outermost layer of oxide grown on a silicon cylinder $1 \mu\text{m}$ in radius. To produce these plots, we have solved Equation (4) using the expression for viscosity η given in Equation (6). We have used a value of $G = 3 \times 10^{11} \text{ dynes/cm}^2$ [22, 23] and viscosity values for η_0 from Stiffler [14], and have assumed oxidation rates for {100} surfaces using the Deal-Grove formulation [24]

$$\frac{X_{\text{ox}}^2}{k_p} + \frac{X_{\text{ox}}}{k_\ell} = t, \quad (7)$$

where k_p and k_ℓ are the parabolic and linear growth-rate constants. For wet oxidations, we have used values for k_p and k_ℓ from Razouk et al. [25]; for dry oxidations, we have used values from Hess and Deal [26, 27].

In Figure 2(a) we show calculated stress values for varying thicknesses of oxide grown at different temperatures. This plot indicates that stress should decrease as the oxidation temperature is increased. This is primarily due to the decrease of viscosity with increasing temperature. Similar plots by Hu [15], using a stress-independent viscosity model, show a stronger temperature dependence on stress than that shown in Figure 2(a). A secondary effect involves the quicker transition at higher temperatures from linear growth rate ($\dot{X}_{\text{ox}} \propto t$) to parabolic growth rate ($\dot{X}_{\text{ox}} \propto t^{-1/2}$). In the viscous flow regime, stress is expected to decrease as the strain rate decreases [Equation (3)]. Therefore, as the oxidation rate decreases during parabolic growth, stress should also decrease. The effect of parabolic oxide growth is most pronounced for the 1100°C case in Figure 2(a), where more of the oxide growth takes place in the parabolic regime compared to the lower temperatures of oxidation. Similar plots by Rafferty et al. [13], using the same type of viscoelastic oxide model, do not show this behavior. For simplicity, Rafferty et al. ignored the parabolic growth regime in their treatment and assumed oxidation rates to be constant with time.

Including full linear-parabolic growth-rate behavior for the oxide, Figure 2(a) shows that for a given temperature there should be a maximum value in the amount of stress in the oxide. Once the oxide growth has moved from the elastic to viscous regimes, no further increase in stress is

Explore Litigation Insights

Docket Alarm provides insights to develop a more informed litigation strategy and the peace of mind of knowing you're on top of things.

Real-Time Litigation Alerts



Keep your litigation team up-to-date with **real-time alerts** and advanced team management tools built for the enterprise, all while greatly reducing PACER spend.

Our comprehensive service means we can handle Federal, State, and Administrative courts across the country.

Advanced Docket Research



With over 230 million records, Docket Alarm's cloud-native docket research platform finds what other services can't. Coverage includes Federal, State, plus PTAB, TTAB, ITC and NLRB decisions, all in one place.

Identify arguments that have been successful in the past with full text, pinpoint searching. Link to case law cited within any court document via Fastcase.

Analytics At Your Fingertips



Learn what happened the last time a particular judge, opposing counsel or company faced cases similar to yours.

Advanced out-of-the-box PTAB and TTAB analytics are always at your fingertips.

API

Docket Alarm offers a powerful API (application programming interface) to developers that want to integrate case filings into their apps.

LAW FIRMS

Build custom dashboards for your attorneys and clients with live data direct from the court.

Automate many repetitive legal tasks like conflict checks, document management, and marketing.

FINANCIAL INSTITUTIONS

Litigation and bankruptcy checks for companies and debtors.

E-DISCOVERY AND LEGAL VENDORS

Sync your system to PACER to automate legal marketing.