Reliability limits for the gate insulator in CMOS technology

by J. H. Stathis

Aggressive scaling of the thickness of the gate insulator in CMOS transistors has caused the quality and reliability of ultrathin dielectrics to assume greater importance. This paper reviews the physics and statistics of dielectric wearout and breakdown in ultrathin SiO_2-based gate dielectrics. Estimating reliability requires an extrapolation from the measeurment conditions (e.g., higher voltage) to normal operation conditions. To reduce the error in this extrapolation, long-term (>1 year) stress experiments have been used to measure the wearout and breakdown of ultrathin (<2 nm) dielectric films as close as possible to operating conditions. Measured over a sufficiently wide range of stress conditions, the time to breakdown (T_{BD}) does not obey any simple "law" such as exponential dependence on electric field or voltage, as has been commonly assumed in reliability extrapolations. Thus, the interpretation of T_{BD} data remains somewhat controversial. Present research is aimed at better understanding the nature of the electrical conduction through a breakdown spot, and the effect of the oxide breakdown on device and circuit performance. In some cases an oxide breakdown may not lead to immediate circuit failure, so more research is needed in order to develop a quantitative methodology for predicting the reliability of circuits.

Introduction

The microelectronics industry owes a great deal of its success to the existence of the thermal oxide of silicon, i.e., silicon dioxide (SiO_2). A thin layer of SiO_2 forms the insulating layer between the control gate and the conducting channel of the transistors used in most modern integrated circuits. As circuits are made more dense, all of the dimensions of the transistors are reduced ("scaled") correspondingly [1], so that nowadays the SiO_2 layer thickness (t_{ox}) is 2 nm or less. **Figure 1** shows the historical trend in oxide thickness for high-performance logic applications over the past decade.[1] The oxide thickness has been decreasing quasi-exponentially, but clearly this trend must saturate at some point, since there are physical and practical limits on how thin an oxide film can be made. Where will this point be? What is the ultimate oxide thickness limit, and when will it be reached? This paper reviews the limitations to SiO_2 film thickness, with particular emphasis on reliability. We describe the physics and statistics of oxide breakdown, illustrated by recent long-term stress data. Finally, we explore the implications of oxide breakdown for the

[1] The sources for these data are the technical digests of the International Electron Devices Meeting and the Symposium on VLSI Technology. Data prior to 1995 compiled with the help of Ping Yang, Texas Instruments Corporation.

0018-8646/02/$5.00 © 2002 IBM

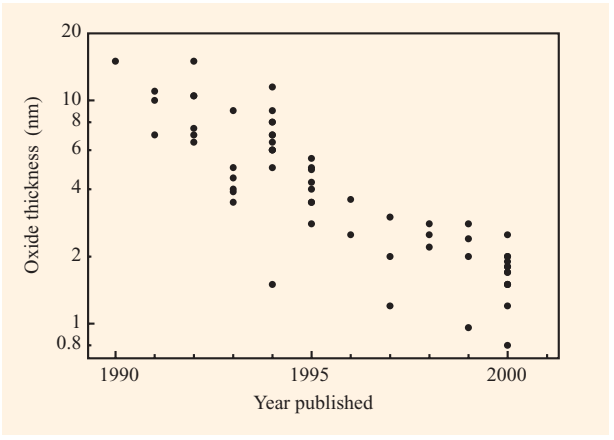


Figure 1

Historical trend in the thickness of SiO_2 used as the gate insulator in CMOS logic vs. year of publication. These data are from research reports, not necessarily manufacturable technologies.

reliable functioning of circuits, and ways in which this might be improved.

Limitations to oxide thickness reduction

The essential physical limitations on gate insulator thickness, ignoring "extrinsic" concerns related to manufacturability and yield (e.g., pinhole formation and film uniformity) are related to the exponentially increasing gate current as t_{ox} is reduced, and the effect of this current on both the functionality and reliability of devices and circuits. Although no insulator is ideal, amorphous SiO_2, with a bulk resistivity of $\sim 10^{15}$ Ω-cm and a dielectric breakdown strength of $\sim 10^7$ V/cm, is one of the best nature has provided. Its usefulness in silicon technology is due to the fact that films of this material can be grown by thermal oxidation of clean silicon surfaces, and that the resulting films can be made with sufficiently low densities of charge traps. In addition, the energy barrier (conduction-band discontinuity) between Si and SiO_2 is 3.1 eV, only about 1 eV lower than the Si/vacuum barrier. The leakage current is controlled by quantum-mechanical tunneling, either Fowler–Nordheim (FN) tunneling [2] at high fields, or, for applied bias less than about 3 V, by direct tunneling. At a typical operating field of 5 MV/cm, the current density for FN tunneling is $<10^{-9}$ A/cm^2 [3]. However, for $t_{ox} \lesssim 4$ nm, the direct tunneling current increases exponentially by about one order of magnitude for every 0.2–0.3-nm reduction in oxide thickness [4].

Performance and power consumption

The gate leakage current causes increased power consumption and may affect device and circuit functionality. Thus, the leakage current imposes a practical limit on oxide thickness. It is instructive to look at the history of this concept. In 1972, some claimed [5] that the lower limit would be ≈5 nm. Since this was about 100 times thinner than the devices in use at that time, it must have seemed a rather futuristic limit; nonetheless, the researchers projected that this limit would be reached by the mid-1980s on the basis of the rate of technology improvement at that time. The data of Figure 1 suggest that the rate of oxide thickness reduction was somewhat lower than was projected, but certainly the improvement in MOS technology did not cease in 1980, nor has 5-nm SiO_2 turned out to be a fundamental limit. As early as 1988, researchers in Japan [6] fabricated MOSFETs with 2.3-nm gate oxide, noting that the device characteristics were not significantly perturbed as long as the gate leakage was significantly less than the channel current, which is achieved by decreasing the gate length. The same consideration was later shown to permit a 1.5-nm oxide [7]. The lesson from this is that it is very difficult to make accurate predictions about the limitations of technology, possibly because engineers and scientists regard such predictions as a challenge.

A thickness of ~1.5 nm, corresponding to a leakage current of ~10 A/cm^2, is now viewed as a sort of practical limit for SiO_2 [8], although some researchers have recently claimed that leakage current as high as 100 A/cm^2, corresponding to ~1.0–1.3 nm SiO_2, is feasible for both static and dynamic logic [9]. (Clearly, such high leakage currents are acceptable only for high-performance server systems, not for low-power portable electronics.) A recent analysis claims that 0.8-nm SiO_2 is a lower limit in MOSFET applications, based on consideration of threshold voltage fluctuations when the resistance of the oxide becomes comparable to that of the gate metal [10]. These limits are not fundamental, but depend on circuit and system design, as well as materials properties such as film uniformity and interface roughness. However, below $t_{ox} \approx 1$ nm, overlap of the electronic wave functions of the silicon conduction bands may become large enough to effectively reduce the tunneling barrier [11]. Other performance limitations, e.g., the decreasing n-FET drive current observed for oxide thickness below 1.3 nm, may also represent limitations of fundamental physics [12, 13].

Wearout and breakdown

The current flowing through an ultrathin gate oxide is not merely a nuisance parasitic, but also causes reliability problems by leading to long-term parameter shifts (wearout) and eventually to oxide breakdown. The stress-induced parameter shifts are gradual, and the degradation is predictable on the basis of experimental data and physical models. The impact on device design therefore involves an engineering tradeoff between short-term and

long-term performance. Oxide breakdown, on the other hand, is a sudden discontinuous increase in conductance, often accompanied by increased current noise. Breakdown is generally understood to be the result of a gradual and predictable buildup of defects such as electron traps in the oxide, but the precise point at which breakdown occurs is statistically distributed so that only statistical averages can be predicted.

The rate of defect generation in the oxide is proportional to the current density flowing through it, and therefore the reliability margin for gate-oxide breakdown has been drastically reduced as a consequence of device scaling. At present, predicted reliability "limits" for the gate-oxide thickness range from less than 1.5 nm [14] to 2.8 nm [15]. Even if the smallest estimates were correct, the latest international roadmap for the industry [16] anticipates that $t_{ox} \sim 1$ nm will be needed by 2005 for 60–70-nm gate lengths (90-nm lithography node) operating at ~1 V in order to meet the desired performance targets. One way this might be achieved is by replacing the SiO_2 with another dielectric with a higher dielectric constant, such as metal oxides, metal silicates, and epitaxial perovskites. This avenue is being aggressively pursued but faces large hurdles, since neither the materials science nor the electrical properties of these materials are understood nearly as well as for SiO_2 [17, 18]. A more conservative approach is to use oxynitrides, which have been reported to show somewhat improved reliability characteristics compared to pure SiO_2 [19–21].

Oxide lifetime and product lifetime

The reliability of SiO_2 in microelectronics, i.e., the ability of a thin film of this material to retain its insulating properties while subjected to high electric fields for many years, has always been a concern and has been the subject of numerous publications over the last 35–40 years, ever since the realization that SiO_2 could be used as an insulating and passivating layer in silicon-based transistors [22, 23]. (For a recent review, see [24].) In a recent survey of nine manufacturers asked to rate seventeen different reliability-related technology constraints in order of importance, gate-dielectric reliability was the top concern overall and was ranked among the top five concerns by every manufacturer surveyed [25].

In retrospect, it is not clear that this concern is justified. First, for the thicker oxides used in the past with $t_{ox} \gtrsim$ 3–4 nm, recently published models [4, 14, 15, 26] have indicated that intrinsic gate-oxide reliability has probably not been a real issue at the operating voltages employed. The difficulty has always been to extrapolate oxide lifetime data from accelerated stress conditions (high voltage), under which data can be collected within a reasonable time, to the lower operating voltage, at which the failure rate should be very low. The oxide wearout and breakdown mechanisms change as a function of voltage and oxide thickness [27], so that earlier extrapolations, before this was understood, were easily in error. A controversy still exists over the proper extrapolation from ~2 V to 1 V [14, 26, 28, 29], and this has major implications for reliability projection. This is the subject of a large part of the present paper. Second, for the ultrathin oxides at lower voltages now being employed, a potentially nondestructive "soft" breakdown mode is frequently observed [30]. This is discussed in the penultimate section of this paper. Thus, the true implications of oxide breakdown for the reliability of integrated circuits are still unclear. Indeed, microprocessors have already been successfully manufactured with sub-2-nm gate oxide [31, 32], with no immediate catastrophic result. The next section of this paper deals with the phenomenon of oxide breakdown from a fundamental physical viewpoint, and describes in detail the uncertainties surrounding oxide reliability projection. The possible implications for product reliability are revisited later.

The physics and statistics of oxide breakdown

Defect generation and breakdown at low voltage

The essential elements of our present understanding of oxide wearout and breakdown are illustrated in **Figure 2**. Electrons flowing across the oxide trigger several processes depending on their energy, which is determined by the gate voltage for thin oxides where electron transport is ballistic or quasi-ballistic [33, 34]. At least three defect-generation mechanisms have been identified: The first two, impact ionization and anode hole injection, occur at higher voltages and lead to hole trapping and hole-related defect generation [27, 35]. The lowest-energy process so far identified, which dominates at the voltages at which present MOSFETs operate, is the so-called "trap-creation" process attributed to hydrogen release from the anode [33], illustrated in **Figure 3**. This process has a threshold gate voltage of about 5 V, but continues in the subthreshold region even at operating voltages down to 1.2 V or lower [34, 36]. Eventually, damage builds up to the point at which the oxide breaks down destructively.

The evidence for hydrogen involvement in defect generation and breakdown is circumstantial but strong [24, 33, 37]—notably the observation of substrate dopant passivation [38] and hydrogen redistribution [39–41] during hot-electron stress, the enhanced degradation rate of hydrogen-soaked films [42], and experiments showing that exposure of bare SiO_2 films to atomic hydrogen radicals, in the absence of any electric field, produces electrically active defects essentially identical to those produced by electrical stress or radiation [37, 43–55]. The desorption rate of hydrogen from Si surfaces was

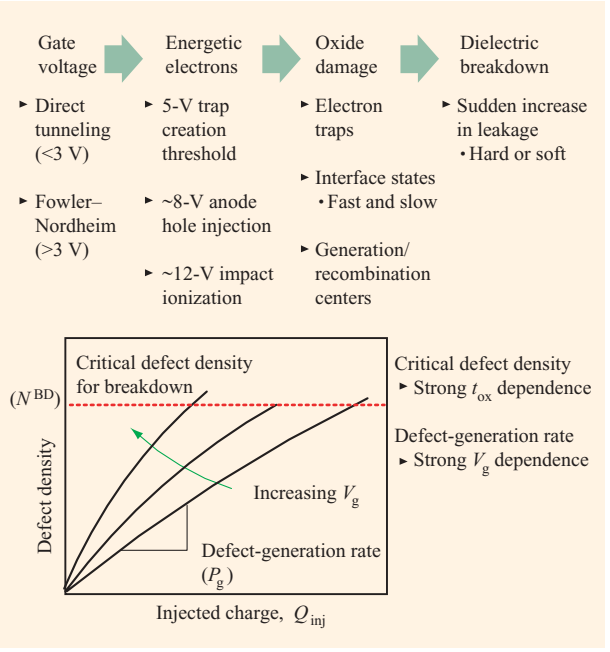


Figure 2

Outline of the mechanism of defect generation leading to breakdown in SiO_2. The gate voltage causes energetic electrons to flow across the oxide, leading to a buildup of damage in the form of microscopic defects which act as electron traps. When sufficient defect density is reached, the oxide no longer acts as a good insulator. Reprinted with permission from [24], ©2001 IEEE, and [114].

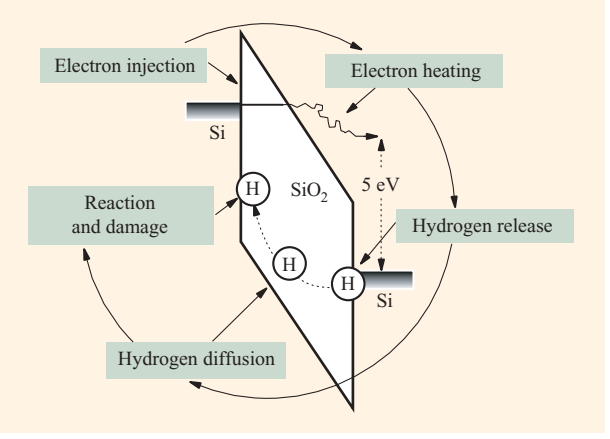


Figure 3

Schematic of the hot-electron-induced hydrogen-release mechanism of defect generation in SiO_2. Hydrogenous species (either protons or atomic hydrogen) are freed from the silicon layers by the energetic electrons, and then drift or diffuse through the oxide, reacting with the lattice to produce electron traps and other defects. Adapted with permission from [114].

measured as a function of incident electron energy [56] and showed a dependence remarkably similar to the voltage dependence of the trap-generation process [33, 36]. However, a major perceived stumbling block to the general acceptance of the "hydrogen model" for breakdown has been the apparent lack of any isotope effect for the breakdown process [57] compared to the large effect observed for hydrogen/deuterium desorption and for hot-electron-induced channel interface degradation [58]. This may recently have been resolved by the observation [59, 60] of a significant isotope effect on the stress-induced flat-band voltage shift and stress-induced leakage current (SILC), which is a measure [50, 61, 62] of the bulk traps that ultimately relate to breakdown. A significant isotope effect on trap generation and oxide breakdown in deuterated oxide was also reported earlier [63].

Two other physical models for breakdown have been widely discussed in the literature. The first is the anode-hole injection (AHI) model, which asserts that breakdown is caused by holes that are injected from the anode contact [3, 64]. One of the main attributes of this model is a constant value of the hole fluence to breakdown independent of oxide field [3], $Q_p \approx 0.1$ C/cm^2, decreasing for t_{ox} less than about 6 nm [64]. According to earlier calculations [65], the gate-voltage threshold for positive charge generation by hole trapping due to AHI is 7–8 V for FETs with n^+-poly gates, and this was confirmed experimentally [35]. Thus, this mechanism probably cannot account for defect generation and breakdown at lower voltage. A recent modification of the AHI model [14, 66] proposes that a weaker minority-carrier ionization process [67] is responsible for hole injection and defect generation at low voltages. This mechanism will be operative for electron injection into a p-type material or hole inversion layer, including n-FETs with low gate doping when the n-poly gate is inverted [68]. The modified model can successfully fit the measured slope of T_{BD} vs. voltage at high fields [14, 69] but cannot account for the absolute magnitude of the defect-generation rate. Since the hole current at low voltage (e.g., 2–3 V) is at least 12 orders of magnitude lower than the primary electron current [66], the defect-generation rate per hole is very much greater than the rate per injected electron. However, direct measurement of the rate of defect generation by transport of hot holes through an SiO_2 layer [70, 71] gives values comparable to the generation rate due to electrons and many orders of magnitude less than what is required by a minority-carrier mechanism. Charge-to-breakdown measurements using substrate hot-hole injection [72, 73] and on p-FETs at low bias where hole tunneling dominates the total gate current (**Figure 4**) [71] have also shown that the hole fluence to breakdown, Q_p, is as large as eight orders of magnitude greater than the value used

in the AHI model, consistent with the measured defect-generation rate. The data in Figure 4 indicate that p-FET breakdown may be the limiting factor at low voltage, contrary to the usual idea [74] that n-FETs should represent the worst case.

The other widely cited breakdown model is the "thermochemical" model, or "*E*-model," in which *E* refers to the electric field across the oxide. This model proposes that defect generation is a field-driven process, and that the current flowing through the oxide plays at most a secondary role. The development of this model from its origin in the mid-1980s through the late 1990s has been reviewed by McPherson et al. [75, 76]. The *E*-model has attained widespread acceptance, largely on the basis that, empirically, the time-to-breakdown (T_{BD}) data appear to follow an exponential dependence on oxide field [77–80], including an experiment of three years' duration at oxide fields down to 5.3 MV/cm on 9-nm films [81]. However, it must be pointed out that the exponential dependence on field is not proof of the validity of the particular physical model. Probably the strongest evidence against the thermochemical model comes from substrate-hot-electron (SHE) injection experiments [82, 83]. In this experiment, using a specially designed n-FET structure [84], a forward-biased n^+ diffusion in the p-type substrate is used to inject electrons toward the inversion layer at the Si/SiO_2 interface. The current through the oxide can be controlled independently of both the substrate and gate bias by varying the magnitude of the forward bias applied across the injector junction. The substrate voltage controls the maximum electron energy incident on the Si/SiO_2 interface, and for substrate bias high enough that some electrons are injected over the energy barrier into the oxide, the gate voltage additionally controls the energy of the electrons incident on the gate/oxide interface. In these experiments it was found that the charge to breakdown (Q_{BD}) is strongly dependent on the substrate bias, even though the oxide field is held fixed [83]. Therefore, Q_{BD} correlates with the electron energy, not the oxide field. This was also demonstrated for conventional Fowler–Nordheim stress by varying the doping of the anode [68, 85]. The SHE experiments also showed that T_{BD} is inversely related to the current density [83], again showing that breakdown is dominated by the effect of the energetic electrons and not the field in the oxide.

Independently of the assumed physical mechanism of defect creation, it is an experimental observation that in a large variety of oxide thicknesses stressed over a wide range of voltages, the charge to breakdown (Q_{BD}) is inversely related to the initial rate of defect generation for most stress conditions [27, 79, 86–88]. Defining the defect generation probability per injected electron density as $P_g \equiv qdN/dQ$, DiMaria proposed [4, 27, 34, 89] the simple relationship

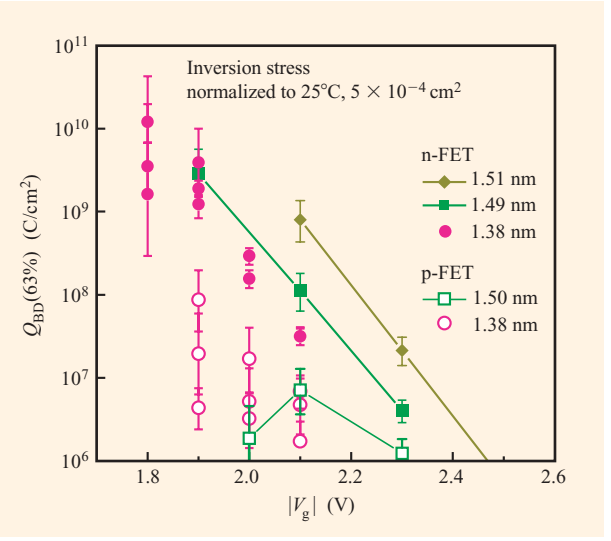


Figure 4

Charge-to-breakdown (Q_{BD}) measurements on p-FETs and n-FETs under inversion conditions as a function of gate voltage. The p-FET Q_{BD} corresponds to a critical hole fluence for breakdown, Q_p, for tunneling holes. Reprinted with permission from [24], ©2001 IEEE, and [71].

$$Q_{BD} = qN^{BD}/P_g, \tag{1}$$

where q is the magnitude of the electron charge and N^{BD} is a proportionality constant, which we later see is a function of t_{ox}. More generally,

$$N^{BD} = \frac{1}{q}\int_0^{Q_{BD}} P_g\, dQ. \tag{2}$$

Equation (1) holds when the defect generation is first-order in the electron fluence. (It is commonly observed that the defect density N is a sublinear power-law function of the injected charge Q. However, observed over a sufficiently wide range of fluence, the full dependence is typically sigmoidal, with a linear region bracketed by sublinear portions at low and high fluence [4, 71]. If the low-fluence background is subtracted, a linear behavior is found independent of stress conditions.)

Voltage dependence of defect generation

Extensive work in the 1980s revealed the existence of the so-called "2-eV trap-creation threshold" (electron energy measured with respect to the bottom of the SiO_2 conduction band) for defect creation by hot electrons in SiO_2 [33]. For thick oxides, electrons in the conduction band obtain this average energy for oxide fields greater than about 2–4 MV/cm. In the thin oxides of current interest, for which electron transport is ballistic or quasi-

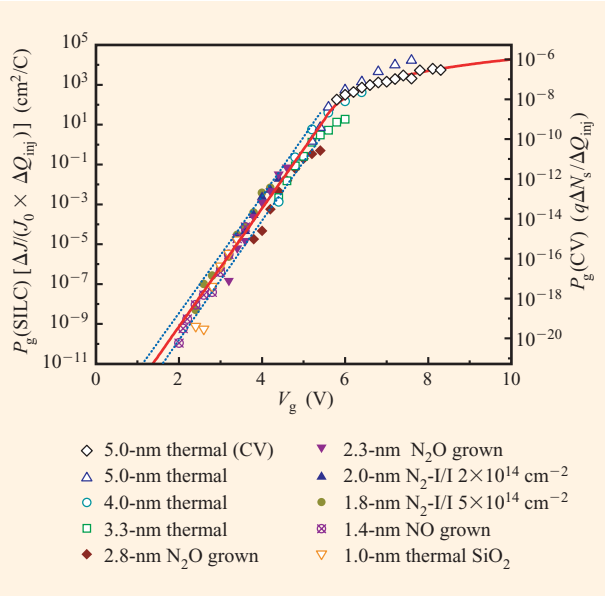


Figure 5

Defect-generation probability measured from SILC [P_g(SILC)] and CV stretch-out [P_g(CV)] as a function of stress voltage for various oxides. The solid red curve is the model based on an exponential fit to the data below 5.6 V and a power-law fit to the CV data. The dotted blue curves are the prediction band for 95% confidence. The data near threshold at 5 nm are used to match the CV and SILC values. Reprinted with permission from [4]; ©1998 IEEE.

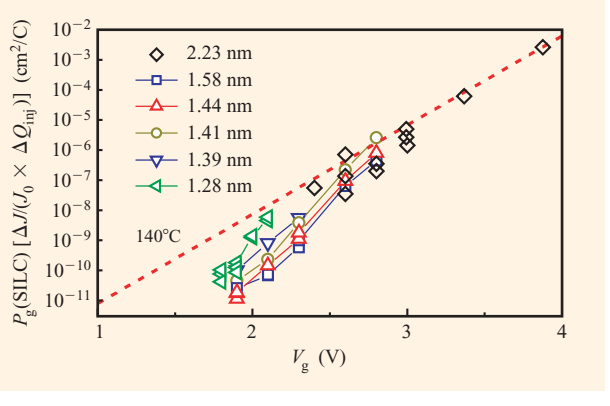


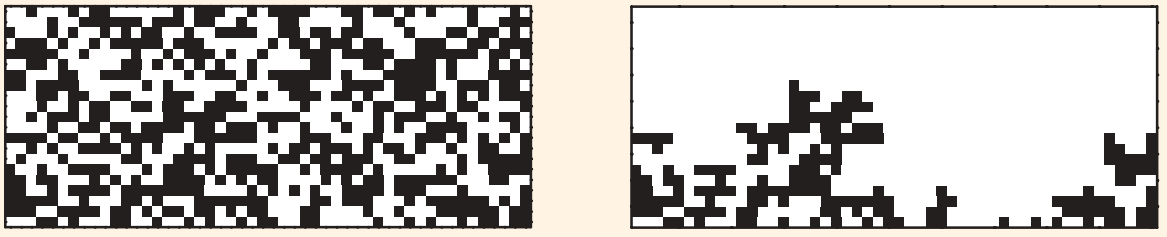
Figure 6

Detail of P_g data from long-term stress experiments, showing deviation from purely exponential behavior below 3 V. Adapted with permission from [24]; ©2001 IEEE.

ballistic [33], the threshold corresponds to a gate voltage of about 5 V for Fowler–Nordheim tunneling through the 3-eV potential barrier at the Si/SiO_2 interface. For voltages below this threshold, it was shown [34] that the defect-generation rate depends only on the absolute value of the gate voltage (V_g), independent of substrate or gate doping or polarity. Recent work has given the somewhat surprising result that 2 eV is not a hard threshold [36, 89]. Instead, there is a subthreshold trap-generation process that decreases exponentially below 5 V, shown in **Figure 5**. The data are capacitance voltage (CV) measurements of the interface state density and stress-induced leakage-current (SILC) data. In SILC measurements the relative increase in current in the direct tunneling range below 3 V is expressed as $\Delta J/J_0$, where J_0 is the initial current density in the as-fabricated device. This quantity is proportional to the density of generated neutral electron traps [50]. The left- and right-hand vertical scales in Figure 5 have been adjusted to match the values obtained using the two measurement techniques on the 5-nm sample, from which we obtain [90] the relation N_s (cm^{-2}) $\approx 3 \times 10^8\ \Delta J/J_0$. This is consistent with the value found [50] for the bulk neutral traps, where N_n (cm^{-2}) $\approx 1.25 \times 10^8\ \Delta J/J_0$.

Figure 5 includes data from a variety of oxide thicknesses and processes, including thermal SiO_2 grown in O_2, oxides grown in N_2O, and oxides grown on nitrogen-ion-implanted (N_2–I/I) substrates, indicating the relative insensitivity of P_g to the oxidation process. P_g is also observed to be independent of stress voltage polarity. For each oxide thickness, it is possible to measure P_g over only a limited range of V_g while keeping the measurements within an experimentally accessible time scale. Plotting overlapping ranges of V_g obtained using oxides of different thicknesses shows a universal exponential behavior of P_g as a function of V_g, independent of t_{ox}. Using substrate hot-electron (SHE) and channel hot-electron (CHE) stress, it is possible to obtain much greater hot-electron flux at the interface, permitting measurements of P_g at low electron energy for even thick samples. Using these techniques, it has been shown [36, 91, 92] that defect generation by hot electrons impinging on the substrate/oxide interface follows the same dependence on energy as that from Fowler–Nordheim injection through the oxide, and that the exponential V_g dependence extends at least as low as 1.2 V [92].

A remarkable and unexpected feature of the data of Figure 5 is that there is no further threshold voltage below which the generation rate drops faster than the exponential trend. In particular, there is no large discontinuity or change in slope at the transition from Fowler–Nordheim to direct tunneling at about 3 V. This is contrary to earlier suggestions [93, 94] that breakdown would be strongly suppressed in the direct-tunneling regime, and strongly suggests that the relevant energy scale is the electron energy in the anode, which is the



(a) Occupied sites ($p_0 = 0.5$)

(b) Cluster No.1, connects to bottom face

Figure 7

Schematics of the percolation model computer simulation. (a) Two-dimensional simple cubic lattice with uniform 50% occupancy. "Defects" are represented by dark squares. (b) Starting from the same defect distribution as in (a), only the occupied sites that connect to the bottom face via nearest neighbors are kept. Note the use of periodic boundary conditions in the horizontal dimension. For this sample size and defect density, there is no breakdown path connecting the bottom face to the top face, but connecting paths would exist if the sample were made thinner or wider. This schematic shows a small (21×50) sample in two dimensions. The actual calculations were performed in three dimensions using a larger sample. Reprinted with permission from [24], ©2001 IEEE, and [114].

polysilicon gate or the silicon substrate depending on injection polarity. In fact, detailed measurements [26, 92] have found a sigmoidal inflection in the voltage dependence of P_g between 2 and 3 V. This is shown in **Figure 6**.

Critical defect density for breakdown

The idea of damage building up to a critical level [95, 96] has been a key insight in leading to a predictive model of oxide reliability. The concept does not depend in any way on the physics of defect generation: Like the hydrogen-induced-defect-creation model, the field-driven model [75, 76, 97] (E-model) also assumes a critical density of broken bonds in order to induce electric breakdown and thermal runaway, and the most recent version of the AHI model [14, 98] likewise adopts the viewpoint that holes create some (unspecified) form of defect which eventually leads to a critical conduction path. Two important papers in the 1990s have led to quantitative application of this concept. Suñé et al. [99] showed that the assumption of a critical defect density (treating N^{BD} as a fitting parameter) leads to the correct statistical behavior describing breakdown distributions. Degraeve et al. [100] formulated the percolation model, in which breakdown is envisioned as the formation of a connecting path of defects, as a result of random defect generation throughout the insulating film.

The percolation concept (and the origin of the thickness dependence of N^{BD}) is schematically illustrated by the computer simulation [90] in **Figure 7**. According to this model, breakdown can occur only when a connecting path of traps is formed across the gate oxide, forming a conducting path from the substrate to the gate. The probability of forming such a connecting path ("percolation path") with randomly generated defects throughout the oxide bulk is computed as a function of defect density and oxide thickness. For a given defect density, the formation of a percolation path is more likely for thinner oxides. Conversely, as the oxide is made thinner, a percolation path can form with some probability at a lower average defect density than is necessary in a thick oxide. In the simulation, we begin by placing defects (black squares) randomly throughout the sample [Figure 7(a)], and then discard all defects which are not part of a cluster that is attached (via nearest neighbors) to one face of the sample [Figure 7(b)].

The face at which we begin when the percolation cluster is computed does not matter, because if a site is part of a cluster that connects the two faces of the sample, it will be counted no matter where we begin. It would be computationally redundant to separately examine the opposite cluster extending from the other face of the sample, because this cluster cannot touch the first face unless the first cluster also spans the sample. For the particular defect density illustrated in Figure 7, there is no percolation path connecting the top to the bottom of the sample, and therefore we have not reached breakdown. However, breakdown would have occurred at this density if the sample were thinner. As the defect density is increased, the percolating cluster extends deeper into the sample, until at some critical density it touches the opposite face. In addition, even for the same defect density and thickness, if the sample were made wider (corresponding to a larger device area), there is a finite probability of locally finding a cluster that does span the thickness. Thus, this rather simple model accounts nicely for all of the

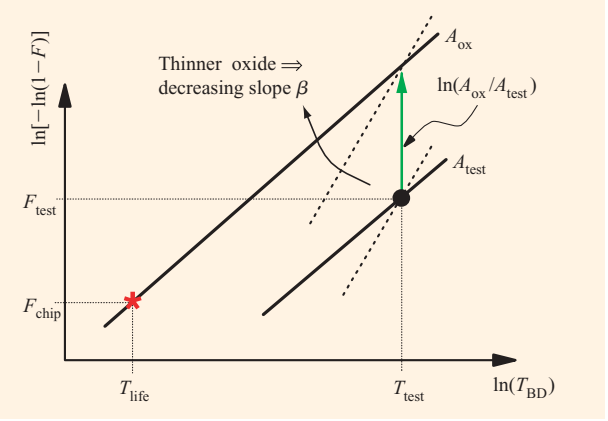


Figure 8

Schematic illustration of a Weibull plot of the cumulative distribution of breakdown times. F is the cumulative fraction of samples failed at time T. Thinner oxides result in broader distributions (smaller slope). This figure illustrates the procedure for scaling measured data (F_{test} at time T_{test}, using test structures with area A_{test}) to the expected product failure rate F_{chip} for chips with total gate area A_{ox}, at end of life, T_{life}. Reprinted with permission from [24]; ©2001 IEEE.

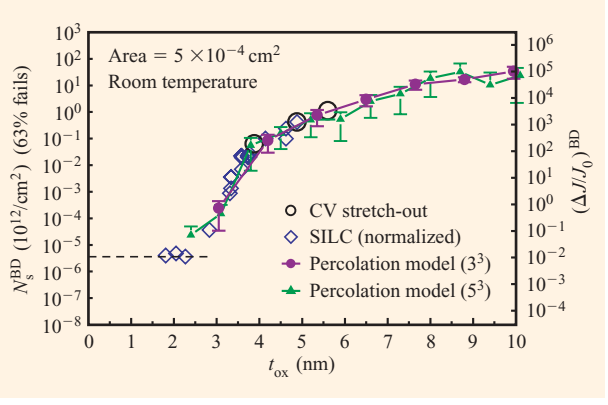


Figure 9

Critical defect density at breakdown (N^{BD}) as a function of oxide thickness. A strong decrease in N^{BD} for thickness less than about 6 nm, followed by a plateau below 3 nm, is explained by the percolation model, as indicated by the filled symbols. The CV and SILC data are normalized to their average values at 5 nm. Reprinted with permission from [24], ©2001 IEEE, and [114].

statistical features of oxide breakdown, including thickness and area dependence.

One important prediction of the percolation model is that the breakdown distributions become broader as t_{ox} is decreased [34, 90, 100, 101]. This is illustrated schematically in **Figure 8**. The shape of the failure distribution is an important parameter used in reliability estimations, as is discussed in the next section of this paper.

In **Figure 9** we plot the measured N^{BD} value (averaged over a range of stress voltage) as a function of t_{ox}. N^{BD} drops by a factor of $\sim 10^5$ from 6 to 3 nm, and then reaches a plateau below 3 nm. These features are fully explained by the percolation model calculations [90] shown by the filled symbols. The fit is described by two parameters, which are adjusted to give a consistent fit to both the data above 3 nm and the plateau below this thickness. The two parameters are the defect diameter (hopping distance) a_0 and the probability (f) that a percolation path triggers destructive breakdown. For $t_{ox} < a_0$, the percolation model predicts that the thickness dependence will vanish, because only one "defect" is required to form a connecting path across the oxide. Therefore, the plateau region is $t_{ox} < a_0$, and N^{BD} in the plateau corresponds to one active defect in the area of the sample. In terms of absolute defect density, normalizing the SILC to interface state density as discussed earlier, the plateau value below 3 nm corresponds to a constant number equal to 2000 defects in the 5×10^{-4}-cm^2 area of each sample, suggesting that each percolation path has a probability of about 10^{-3} of initiating a destructive breakdown event. In other words, only a fraction $f \sim 10^{-3}$ of the defects are "active" or "effective" in causing breakdown. This might result from the different energy levels of defects, or from their different physical or chemical nature. The idea that there exists some special subset of defects which trigger breakdown was independently suggested by the electron microscopy observation of breakdown patterns [102], where it was found that the density of "weak spots" is ~1% of the defect density at breakdown. The value of the defect "size" obtained from the N^{BD} data is $a_0 \approx 3$ nm. This is believed to correspond physically to the electrical sphere of influence of a point defect, e.g., a tunneling distance or trapping cross section. Other authors obtained smaller values for the defect diameter (1.6–1.8 nm) based on fitting to the distribution of breakdown times [100] or resistances [103].

One component of the percolation model is that N^{BD} is relatively independent of the way in which the oxide is stressed, since it is a zero-field model in the formulation described above. This has been the assumption whenever N^{BD} measurements at elevated voltage have been used to project oxide reliability down to operating conditions [4, 15, 104–106]. Several experiments have supported this assumption, using various measurements of the defect density including interface states, trapped charge, and stress-induced leakage [34, 83, 101, 107–111]. On the other hand, several groups have reported

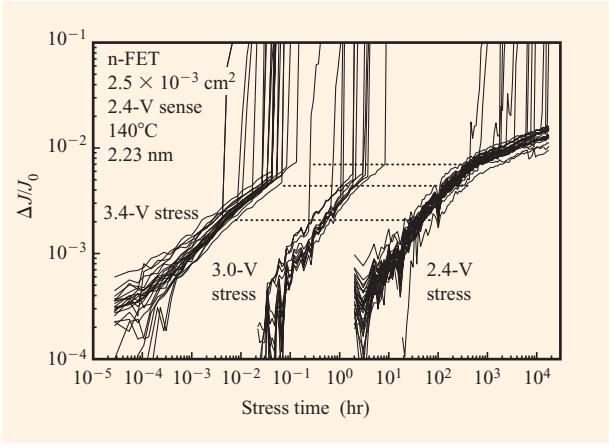


Figure 10

Long-term stress data, showing stress-induced leakage current data to breakdown for three different stress voltages out to two years. Dotted curves are 10%, median, and 90% values of $\Delta J/J_0$ at breakdown for the higher stress voltages. The 2.4-V stress shows breakdown events occurring at statistically higher values of $\Delta J/J_0$. Reprinted with permission from [24]; ©2001 IEEE.

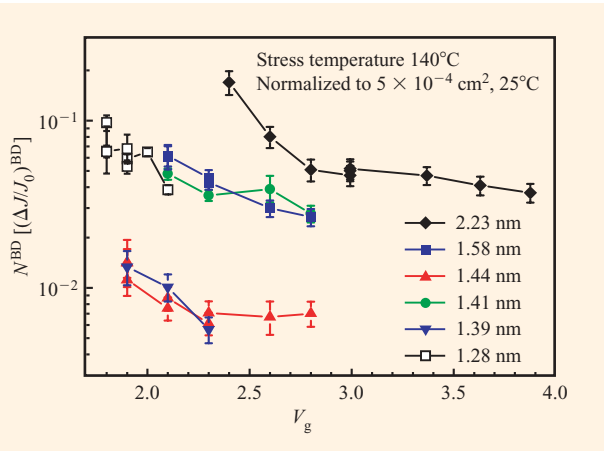


Figure 11

N^{BD} (from SILC data) vs. stress voltage for various ultrathin oxides. Adapted with permission from [24]; ©2001 IEEE.

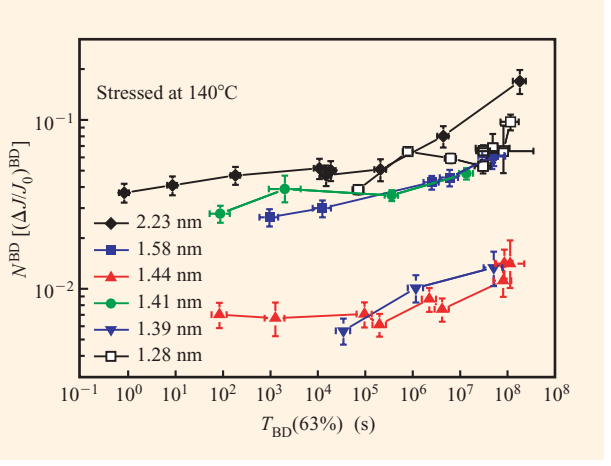


Figure 12

N^{BD} (from SILC data) vs. stress time for various ultrathin oxides, indicating a universal increase in critical defect density at longer times. Adapted with permission from [24]; ©2001 IEEE.

measurements showing a decrease in N^{BD} as the stress voltage is reduced using 3–5-nm oxides [89, 112–114]. This observation could have a significant effect on the reliability projections for such oxides.

In order to measure N^{BD} at lower voltage, i.e., closer to actual operating conditions, it has been necessary to perform long-term reliability experiments on bonded chips. **Figure 10** shows SILC measurements performed for two years at 2.4 V using n-FETs with t_{ox} = 2.2 nm. The horizontal dotted lines in this figure indicate roughly the 10%, median, and 90% values of $\Delta J/J_0$ at breakdown for the higher stress voltages. The 2.4-V stress shows breakdown events occurring at statistically significant higher values of N^{BD}.

Figure 11 plots the characteristic (i.e., the 63rd percentile) values of N^{BD} vs. the gate voltage during stress (V_g), for the same sample as in Figure 10 and other samples with thinner oxide. The oxide thickness is estimated from the accumulation capacitance, including quantum-mechanical corrections, but it must be admitted that the quoted values are uncertain to at least ±0.1 nm. Error bars in N^{BD} reflect the statistical uncertainty; in some cases, at the lowest voltages only ~10% of the samples had failed, and therefore N^{BD} is estimated using the initial fail data and the statistical distributions from higher voltages. In this figure, the N^{BD} values have been normalized to a reference area of 5×10^{-4} cm^2 using Weibull statistics. It can be observed that N^{BD} is only weakly dependent on t_{ox} in this range, and tends to *increase* with decreasing V_g. The observed increase in N^{BD} correlates with an increase in charge to breakdown at low voltage [26], which is shown in the next section.

Although the data in Figure 11 show a trend for N^{BD} to increase as V_g is lowered, there does not appear to be a universal behavior as a function of V_g. For example, for the 2.23-nm oxide, N^{BD} starts to increase below 2.8 V, while for the 1.44-nm oxide the increase in N^{BD} is not seen until V_g is reduced below ~2.2 V. In **Figure 12** the same data have been replotted as a function of the

breakdown times [$T_{BD}(63\%)$] instead of the stress voltage. T_{BD} is a function of V_g, t_{ox}, and device area. In this figure T_{BD} is the projected value to 63% failure for cases in which not all samples have reached breakdown. Again, the error bars reflect the statistical uncertainty. Note that in this figure, N^{BD} is again normalized to a constant reference area, but T_{BD} is *not* normalized; i.e., the x-axis values reflect the actual time under stress, projected to 63% failure but not normalized by area. Here the data show a universal trend, wherein N^{BD} begins to increase for stress times longer than about 10^6 seconds (~10 days). The long stress times necessary to see this effect would preclude its observation in most experiments. For the 1.28-nm oxide at 1.8 V (open squares), devices with two different areas were tested, which when plotted against V_g result in different values of N^{BD} even after area normalization (Figure 11). However, smaller-area devices have a longer lifetime, which accounts for the higher value of N^{BD} as shown in Figure 12.

A voltage-dependent N^{BD} could arise from several sources. The percolation path, which has been modeled in zero field [90, 100] to obtain the fit shown in Figure 9, could be weakly field-dependent. The formation of the percolation path, i.e., the generation of new defects, could depend on the local field produced by the other defects [115, 116]. This would lead to more directed paths at higher voltage, so that the average defect density to form a connecting path across the sample would be reduced. However, according to the idea that the stress time, not the voltage, is responsible for the increase in N^{BD}, the data may be interpreted in terms of a reduction in the defect "effectiveness" (the fraction f of "active" defects), which was introduced [90, 102, 113] to describe the probability that a defect, or percolation path, can trigger breakdown. For long stress times it requires a greater defect density to trigger breakdown, contrary to the assumption that N^{BD} should be independent of the stress condition. The observation of reduced defect effectiveness for stress experiments of very long duration may imply that defects undergo a slow relaxation that reduces their ability to participate in breakdown. It must be emphasized, however, that the final runaway stage of destructive breakdown, whereby a percolation path leads to catastrophic failure, is still not fully understood.

Charge to breakdown (Q_{BD}) and oxide lifetime

Q_{BD} in Equation (1), defined as the time-integrated current density that flows through the oxide until breakdown occurs, is a *physically* meaningful quantity, but the quantity of *practical* interest for an electronic component is the failure rate, which can be derived from the lifetime or time to breakdown, T_{BD}. For constant-voltage stress this is related to Q_{BD} by the relation

$$\int_0^{T_{BD}} J\,dt = Q_{BD}, \tag{3}$$

where J is the instantaneous value of the current density. For thin oxides the current is nearly constant until breakdown (in marked contrast to thicker oxides, in which electron trapping and/or hole trapping cause large changes in the current during stress); therefore,

$$T_{BD} = Q_{BD}/J. \tag{4}$$

The complete measurements of P_g and N^{BD} described in the previous section can explain in detail the voltage and thickness dependence of Q_{BD} in ultrathin gate oxides, shown in **Figure 13(a)**. Because of the exponentially increasing tunnel-current density with decreasing t_{ox} [4], T_{BD} decreases exponentially with decreasing $t_{ox} \lesssim 3$ nm, even though Q_{BD} is only slightly thickness-dependent in this range because of the weak thickness dependence of N^{BD}. The measured T_{BD} data are shown in **Figure 13(b)**. This implies a rapidly diminishing margin for reliability as device dimensions are scaled. Figure 13(b) also shows that data from different laboratories [26, 117–119] are in reasonable agreement, taking into account the differences in oxide thickness, indicating little dependence on processing for state-of-the-art facilities.

It is commonplace in the microelectronics industry to specify an operating life of ten years, i.e., to guarantee a specified (usually small) failure rate over a ten-year period. First, the T_{BD} data such as those in Figure 13(b) must be projected from the high percentiles where experimental data are collected to the low percentiles desired for the product failure rate. Second, a lifetime correction must be made from the small-area test structures to the total gate-oxide area of a chip. Both of these projections depend on the shape of the breakdown distribution, which must be known for reliability estimates to be accurate.

The statistics of gate-oxide breakdown are described using the Weibull distribution [100, 102–124],

$$F(x) = 1 - \exp[-(x/\alpha)^{\beta}], \tag{5}$$

which is an extreme-value distribution in $\ln(x)$ and is appropriate for a "weakest-link" type of problem. Here F is the cumulative failure probability, i.e., the population fraction failed by age x, where x can be either charge or time. The characteristic life α is percentile 63.2, and β is called the slope parameter or Weibull slope. Plotting

$$W \equiv \ln[-\ln(1 - F)] \tag{6}$$

against $\ln(x)$ yields a straight line with slope β. Equation (6) has the useful property that if the area of the samples tested is increased by a factor N, the curve shifts vertically by $\ln(N)$ [90]. Figure 8 illustrates this effect schematically.

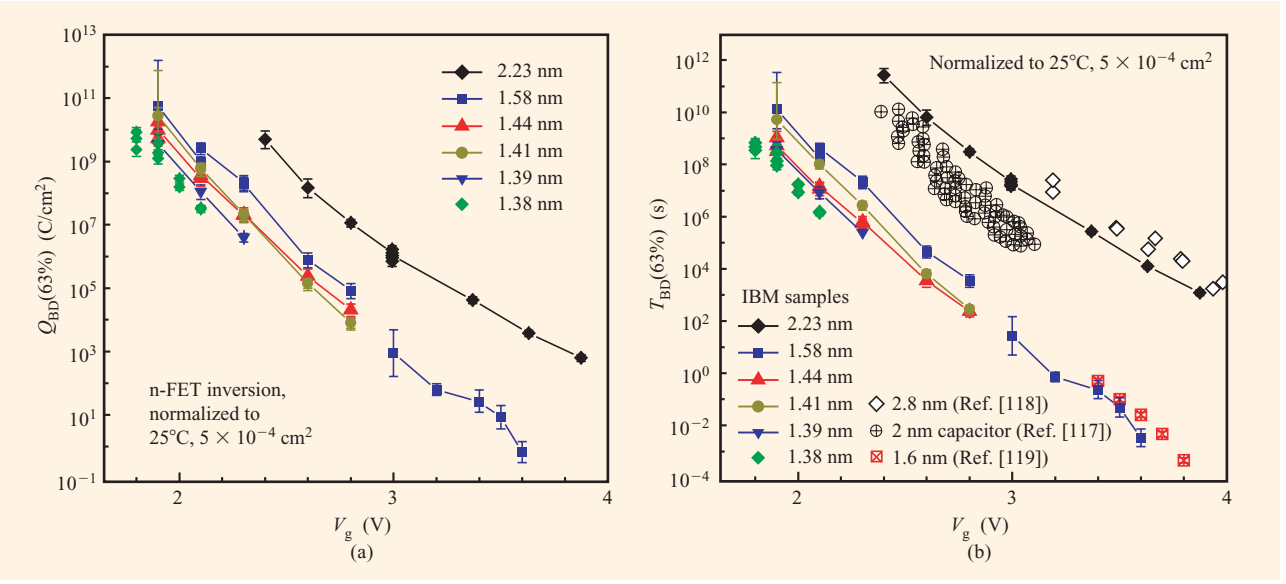


Figure 13

(a) Q_{BD} (charge-to-breakdown) and (b) T_{BD} (time-to-breakdown) data from long-term stress experiments on ultrathin oxides, as a function of gate voltage, showing published data from other laboratories for comparison with IBM data. All data have been scaled to a reference area using Weibull statistics, and to room temperature. Adapted with permission from [24]; ©2001 IEEE.

If the desired low failure rate is F_{chip} over the product lifetime T_{life} for the total gate area A_{ox} on the chip (point indicated by the asterisk in Figure 8), this is equivalent to a higher failure rate F_{test} in time T_{test} on the test structures with area A_{test}, where from Figure 8 we can obtain

$$\frac{T_{life}}{T_{test}} \approx \left(\frac{A_{test}}{A_{ox}} \frac{F_{chip}}{F_{test}} \right)^{1/\beta} . \quad (7)$$

This equation is used to scale measured breakdown times to the expected product lifetime, or equivalently to estimate the chip failure rate from test-structure measurements. Since $F_{chip} < F_{test}$ and typically $A_{test} < A_{ox}$, then $T_{test} > T_{life}$ by this equation; therefore, it is always necessary to measure the test structure under accelerated stress conditions (voltage and temperature). Understanding the voltage dependence is the major reason for investing the time to obtain long-term stress data, as shown in Figures 10 and 13, and it is the reason why so much attention is paid to the physical models of trap generation and breakdown, as discussed above [125, 126]. The uncertainty in the extrapolation to operating voltage is the major contributor to our present uncertainty in the reliability predictions. We now see that neither T_{BD} nor Q_{BD} will obey any simple "law" such as exponential dependence on E, $1/E$, or V_g, as has been commonly assumed in reliability extrapolations. Without the simplifying assumptions of a voltage-independent N^{BD} and a purely exponential voltage dependence of P_g, it becomes more difficult to extrapolate reliability to operating voltage. The steeper V_g dependence of P_g between 2 and 3 V can easily lead to an overly optimistic T_{BD} projection if data are limited to this range. Similarly, projection of data from higher voltages without knowing the complete V_g dependence may lead to more pessimistic or otherwise erroneous projections. Moreover, according to the idea that N^{BD} increases for long stress duration, breakdown is a time-dependent as well as voltage-dependent phenomenon. Care must be taken to separate the voltage- and time-dependent effects when interpreting Q_{BD} or T_{BD} data.

There is no guarantee in practice that actual failure distributions will be perfectly linear on a Weibull plot, and other distributions could in principle be more appropriate [123]. Nonetheless, a study of 900 samples [125] clearly supported the Weibull distribution down to a failure rate of 10^{-3}. (See also [127].) In the simplest case, the presence of multiple failure modes such as extrinsic failure caused by processing defects will lead to bi- or multi-modal distributions with varying slope. A subtler problem can arise from nonuniformity of t_{ox}, which will lead to variations in Q_{BD} because of the thickness dependence of the critical defect density, and will cause further variation in T_{BD} because of the thickness dependence of the tunnel current. This will lead to distortion and curvature of the

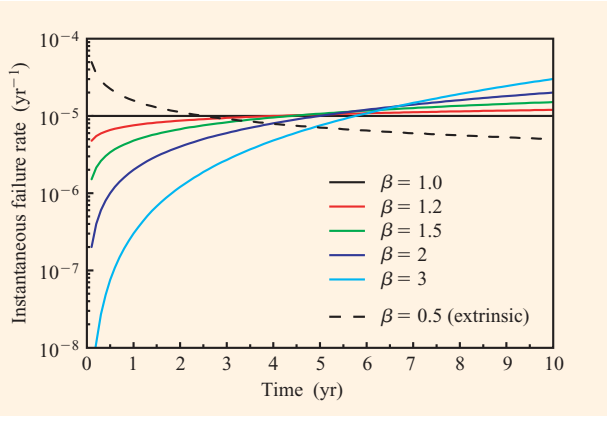


Figure 14

Instantaneous failure rate (per year) as a function of time, with Weibull slope β as a parameter. The failure distributions are normalized to a cumulative failure rate of 100 ppm over ten years. For large β, most failures occur at end of life; however, for β close to 1 the failure rate becomes nearly constant.

Weibull plot [119, 128–130]. Such issues can seriously complicate the otherwise rather straightforward projections. Curvature in the Weibull plot usually indicates a problem such as extrinsic failure modes or nonuniformities that must be dealt with before the data can be properly interpreted.

From the previous discussion, it can be seen that an important parameter for reliability projections is the Weibull slope β. A key advance was the realization that β is a function of t_{ox}, becoming smaller as t_{ox} decreases and approaching $\beta = 1$ for $t_{ox} \lesssim 2$ nm [34, 90, 100]. The thickness variation of the Weibull slope of the Q_{BD} distribution stems from the properties of N^{BD} and is a statistical property of the percolation model for N^{BD}. It is an intrinsic property of the breakdown of ultrathin oxides, and does not imply that the breakdown mechanism is changing. A smaller β means greater sensitivity to the area and failure-rate extrapolations and gives more pessimistic projections via Equation (7), so this is a crucial issue for predicting breakdown of ultrathin oxides. It is often difficult to obtain an accurate value of β from a direct measurement of the failure distribution because of thickness variations and statistical uncertainty [131, 132]. In order to circumvent these difficulties, Wu et al. [125, 133] adopted the area-scaling relation of Equation (7) in order to obtain accurate values.

Another implication of the reduction in β for ultrathin oxides is a change in the character of the statistical failure rate. For larger values of β, the failure rate is initially small, and most of the fails occur near the end of life (see **Figure 14**). However, for β approaching unity, Equation (5) reduces to a simple exponential, and the failure rate becomes constant. This means that a certain number of early fails are to be expected from good parts that pass all screening tests. In this context, it should be noted that intrinsic breakdown, as discussed in this paper, is a purely statistical phenomenon, so that it is not possible to pre-select good samples on the basis of any prior measurement, as far as is currently known. Similarly, it is not possible to screen out the early fails using burn-in. Burn-in, where chips are operated for a short time at elevated voltage and temperature, is designed to eliminate bad parts with an early failure distribution (modeled by the dashed curve with $\beta = 0.5$ in Figure 14). This will also remove those chips in the early part of the intrinsic distribution, but will not reduce (in fact will slightly increase) the intrinsic fail rate during operation.

Reliability projections

The common notion that oxide breakdown is controlled by the magnitude of the electric field E across the insulator is in accord with traditional scaling laws [1], even if in actual practice the voltage has not been reduced as quickly as the oxide thickness. The popular anode hole-injection (AHI) model [3, 134] predicted a lifetime exponentially dependent on $1/E$, and it was claimed, on the basis of extrapolations of data from 2.4-nm and thicker oxides [135], that an oxide field of up to 8 MV/cm is compatible with a 20-year lifetime. Thus, the projected maximum operating voltage for a 2-nm oxide should be 1.6 V. The E-model, as implied by the name, predicts that T_{BD} scales as $\exp(-\gamma E)$, where γ is a constant [75, 76] and the observation that wearout and breakdown for ultrathin oxides is controlled by gate voltage rather than oxide field suggests a similar relationship of the form $\exp(-\gamma V_g)$, where γ is a constant known as the acceleration parameter [4, 28, 34, 105]. As we have seen from the previous section, none of these simple extrapolation laws are exact. **Figure 15** shows various oxide-reliability projections from different research groups [4, 21, 127, 136]. Plotted in this figure is the maximum allowable voltage that can be applied to the total gate area on a chip, such that no more than a specified failure rate will result. The failure rate in this case is defined as the fraction of chips which will experience one or more oxide breakdown events. As can be ascertained from this figure, there remains considerable disagreement among laboratories regarding this reliability limit for SiO_2.

The various projections in this figure are compared to the latest international roadmap for the operating voltage and physical oxide thickness [16]. Note that various laboratories may use different methods to determine the oxide thickness, which can produce results varying by as

much as 0.4 nm [4]. As a rule of thumb, a factor of 10 change in any reliability criterion (e.g., lifetime, failure rate, device area) results in ~0.1 nm change in the minimum t_{ox} for a given supply voltage.

The three most conservative projections date from several years ago, and were based on fitting the average P_g and N^{BD} behavior using breakdown data from relatively short-term experiments ($T_{BD} \lesssim 10^5$ s). The earliest curve in this figure is from data on 2.5-nm and thicker oxides measured over a range of voltages down to about 2.5 V. These data were then fitted to the AHI model [135, 136]. Since the data fitted were the breakdown times of individual small structures, this curve is a voltage projection only, not a full projection to low failure rate and chip area, since the appropriate β values were not known at that time. It is probably fortuitous, therefore, that the extrapolation of this line appears in good agreement with some of the other projections in this figure.

The projections from the more recent version of the AHI model [14] are shown by the open and closed circles in Figure 15 [21]. In this case the model was fitted to data on oxides ranging in thickness from 1.4 nm to 2.4 nm and measured between 3.2 and 4.9 V [137]. An example of such data is shown in Figure 13(b). In this voltage range, the field in the oxide is ~20–28 MV/cm, which is in the range of the intrinsic breakdown strength of SiO_2. Such high field data may therefore be influenced by other bond-breaking mechanisms, in addition to wearout, and probably should not be used to extrapolate to long-term-use conditions. These are the most optimistic of any published oxide-reliability projections of which we are aware.

Using the recent long-term stress data [26], a range of possible oxide reliability projections results depending on how the data are extrapolated. The lower range results from taking into account the complete behavior of the defect-generation rate (P_g) and the critical defect density (N^{BD}) as described previously, using the N^{BD} values projected to approximately ten-year lifetimes and the sigmoidal voltage dependence of P_g to extrapolate from the lowest-voltage data points. In this case, the roadmap runs into problems very quickly below 2 nm, which corresponds to the technology being manufactured today. The upper range of projected oxide reliability comes from a straight-line [$\exp(-\gamma V_g)$] projection of the T_{BD} data [Figure 13(b)] in the region between 2 and 3 V. In this voltage range T_{BD} shows the strongest voltage dependence, because of the behavior of P_g (decreasing relative to the average exponential trend) and N^{BD} (increasing with decreasing voltage and increasing stress time) for these measurement conditions. In this case the results suggest a concern regarding the ITRS roadmap below about 1.6 nm for 1.2-V operation. If the voltage acceleration parameter

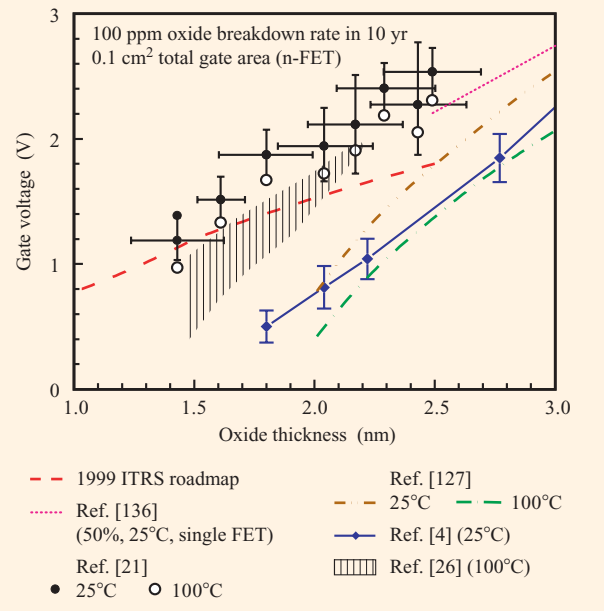


Figure 15

Oxide reliability projections from different research groups. Plotted in this figure is the maximum allowable voltage which can be applied to the total gate area on a chip, such that no more than a specified failure rate will result. The failure rate in this case is defined as the fraction of chips which will experience one or more oxide breakdown events.

γ should continue to increase at lower V_g, as some have suggested [28, 29], a more favorable projection will result.

It is clear that there is no agreement at the present time regarding the correct oxide-reliability extrapolation. While all of the estimates for the minimum t_{ox} are rather closely clustered in the 1.5–2.5-nm range, this is precisely the range of interest for the industry roadmap, and it represents a large variation in device performance and circuit density. Thus, the situation is complicated by the fact that we cannot safely adopt either a worst-case or a best-case outlook: The worst-case scenario appears to be unacceptable from a technological and economic perspective, and the best-case scenario is risky because even in this case we are left with no margin for error. Thus, it appears that the traditional method of assessing oxide reliability, by observing the breakdown time during an accelerated stress experiment and extrapolating this to operation conditions using physical or empirical models, may no longer be adequate to convincingly ensure the long-term reliability of products based on CMOS circuits. It has therefore become necessary to look in more detail at the nature of the breakdown event and the behavior of devices and circuits after oxide failure.

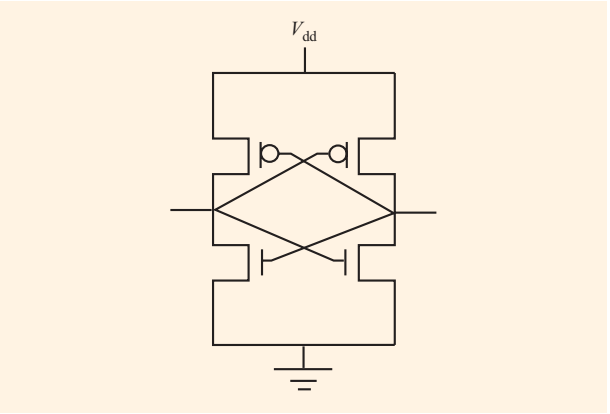


Figure 16

Basic SRAM storage cell. FET gates are connected to other FETs in series with the supply voltage. Reprinted with permission from [24]; ©2001 IEEE.

Device and circuit reliability

The signature of oxide breakdown, which we have been discussing up to now, is the first discrete jump in leakage current, and it has been implicitly assumed that such a breakdown event, affecting even a single transistor in a circuit, can potentially cause a circuit to fail. Some researchers have pointed out, however, that the magnitude of the current increase after breakdown is not always large enough to completely destroy the functionality of a transistor, much less that of an entire circuit [30, 138, 139]. On the other hand, some have argued that breakdown will cause a significant increase in the transistor off current if the breakdown spot is in the drain region [119, 125, 137, 140, 141], which is increasingly likely in short-channel devices [142]. It is also claimed that gate-oxide breakdown can significantly affect rf performance [143]. Current research is therefore aimed at better understanding the nature of the conduction after breakdown [144–155] and the impact of this conduction on device and circuit functionality.

Oxide breakdown events are typically classified as "soft" or "hard" depending on the magnitude of the post-breakdown conduction. There is probably some confusion in the literature over the precise characterization of breakdown modes because of the lack of a precise definition of the terms and because for some experimental conditions the detection of one or the other breakdown mode may be difficult. For example, when testing a large-area structure or a very thin oxide where the initial current is larger than the breakdown current, a "soft" event could be missed, or a "hard" event could be interpreted as soft.

"Soft" breakdown (SBD) occurs in a localized spot, and therefore the current after breakdown is independent of device area [148, 153, 156–158]. The size of this spot is estimated to be 10^{-14} to 10^{-12} cm^2 [153]. Several soft-breakdown events, occurring in different spots, may sometimes be seen in large-area devices prior to a hard breakdown [148, 150, 153, 156, 159, 160]. In addition, the SBD current–voltage (I–V) characteristic is independent of the oxide thickness, at least down to 3 nm, within a band of observed curves [148, 150, 153, 161]. In contrast to the "hard" breakdown (HBD) which shows a roughly linear (ohmic) I–V characteristic of resistance of about 10 kΩ if the damage region remains localized and does not propagate [102, 153], the SBD I–V characteristic is a power law with an exponent of 3–6 [144, 145, 153, 162], although it may be better described by an exponential voltage dependence [154]. The SBD voltage dependence can be explained by a quantum point-contact model [146, 148–151, 153, 154]. The point-contact model can account for both SBD and HBD within a single framework as two limiting cases, depending only on the lateral size of the breakdown spot, which determines the energies of the subbands in the conduction path [153].

Simultaneous with the jump in channel-to-gate current after breakdown is a large increase in the substrate-to-gate current [157, 163–166] and in gate and substrate noise [147, 157, 163, 164, 167–170]. These may be important concerns in specific applications, e.g., analog circuits, but their impact has not been investigated in sufficient detail.

Besides understanding the nature of the conduction through a breakdown spot, a complete assessment of circuit reliability will require that we understand the factors that control the "hardness" of the breakdown event, i.e., the magnitude of the post-breakdown conductivity. HBD is probably a result of thermal damage when sufficient energy is deposited during the breakdown transient [102, 171–181]. This is influenced by the impedance of the circuitry driving the gate in a particular application. For example, in a CMOS circuit the stress conditions on a transistor gate are not the same as the typical experimental conditions used to study reliability. A circuit does not usually subject a gate to either a constant-voltage or a constant-current stress, which are the two types of stress typically employed, but rather to a current-limited stress in which the current through a breakdown spot is limited by the saturation on-current (I_{dsat}) of a complementary transistor in series [180]. **Figure 16** shows the situation for an SRAM cell consisting of two cross-coupled inverters. Not all circuits will fall into this category (for example, some gates will be driven by a low-impedance clock), but SRAM (cache) occupies a large fraction of the area of many chips, and therefore it is a useful place to begin an investigation of circuit reliability. In the SRAM cell, the gate oxides are not connected

directly to the power supply. The n-FET gates are connected via p-FETs, and the p-FET gates via n-FETs. Compared to a typical constant-voltage source, the small transistor in series has a much higher impedance, a much lower current capability, and a lower capacitive loading. It has recently been shown [180, 182] that limiting the drive current to low values (corresponding to very small p-FET drivers) will halt the breakdown event such that the transistors may still be operational after breakdown. Similarly, increasing the circuit time constant by inserting an inductive impedance in series with the device will reduce the severity of the breakdown [181]. Breakdown also becomes softer at low inversion-layer density [155]. In all of these examples, the time to breakdown is not affected, only the post-breakdown conduction. This shows that the nature of the breakdown spot is influenced by the circuit environment of the device, which in turn may affect the overall projected reliability of a circuit.

Besides considering the magnitude of the perturbation of the FET electrical properties after oxide breakdown, we must also consider possible voltage droop caused by the increased current after breakdown [24, 182]. The increased gate current will cause a voltage drop across the transistor which is driving the gate, and this voltage drop will increase as transistors become smaller [182]. Moreover, because there will be a distribution of post-breakdown leakage levels [103, 151, 153, 155, 162, 178, 179], it is likely that some circuit failures will occur as a result of oxide breakdown, but it is difficult to make quantitative predictions. The overall effect of SBD on circuit performance is still an open question, since many different circuit elements are used in practice and some may be more sensitive than others to noise and voltage margins.

After a "hard" breakdown, a MOSFET device is clearly nonfunctional by any ordinary criterion, exhibiting a negative drain current when the gate-to-drain leakage exceeds the normal transistor-on current [140]. However, even a hard breakdown may not completely destroy circuit functionality: It was recently reported [183] that in some cases a circuit may be able to survive an oxide breakdown that previously would have been assumed to be catastrophic. In this experiment it was observed that a ring oscillator continued to function even after multiple transistors suffered oxide breakdown, with breakdown resistances of about 3 kΩ. This means that a ring oscillator/inverter is not highly susceptible to breakdown, and is a poor test vehicle for oxide reliability. Indeed, simple circuit analysis shows that shorting one transistor in an inverter with a resistance of this magnitude produces a circuit which still functions as an inverter, albeit one with poorer performance and higher power consumption.

Even if devices survive after an initial breakdown event, the subsequent stress on the damaged oxide can lead to erratic behavior and a progressive degradation of the device characteristics [125]. Thus, it would be premature to disregard oxide breakdown as a factor in circuit reliability. Much more research will be needed in order to formulate a complete methodology for the quantitative prediction of device and circuit reliability.

Future outlook

Since different circuits will have various degrees of sensitivity to the erosion of noise and voltage margins resulting from oxide breakdown, more research is needed in order to develop a quantitative methodology for predicting the reliability of circuits. The present oxide-reliability methodology will have to evolve from characterizing oxide degradation and breakdown, to the more complex problem of characterizing the response of circuits to oxide breakdown [24].

In the meantime, what avenues are open for avoiding the potential scaling limit imposed by gate-oxide reliability? Although the intrinsic breakdown does imply a fundamental physical limit to the oxide thickness, oxide breakdown is a statistical process, so this limit is dependent on what is deemed to be an acceptable failure rate. Laptop and palmtop reliability of three to five years may be practical. However, because of the uncertainty in prediction, most manufacturers have a minimum lifetime goal of ten years [25], and some key technologies (e.g., telecommunications and networking infrastructure) may require even greater longevity, e.g., 25 to 30 years. However, adjusting the reliability target does not have a great effect on the oxide thickness limit. As pointed out earlier, a factor of 10 change in the reliability target reduces the minimum t_{ox} by only about 0.1 nm. This allows some leeway, but it is not a real solution.

Improved control over the physical properties of the Si/SiO_2 interface has been argued by some to be key for improved performance and reliability [119, 184–189], and to some extent a reduction in the strain at the Si/SiO_2 interface may account for the excellent reliability of present state-of-the-art oxides [188, 189]. Continued development of pre-oxidation cleaning and oxide growth technologies are likely to provide incremental improvements in gate leakage, device performance, and yield, but will not alter the fundamental physics limiting the long-term reliability of ultrathin SiO_2 films [129, 130, 190].

Another avenue is to more aggressively reduce the power-supply voltage. For example, 0.5-V operation of a microprocessor built with 1.5-nm oxide has been proposed [191]. As can be seen from Figure 15, this would nearly satisfy even the most pessimistic reliability projection based on our current data. Of course, this approach has its own problems, including tight control of threshold voltage.

Some of the power-management schemes used for low-power applications can also have a side benefit of improving reliability if sections of a chip are turned off when not in use. It is important to understand, however, that standby power, not just active power, is responsible for the degradation that leads to oxide breakdown [71, 92]. It is therefore not sufficient to use clock gating, which reduces active power but not standby power. Instead, active-power-supply management may be required.

Finally, any failure mechanism, provided that it is not too severe, can be fixed by using error correction and redundancy—features already present in many systems.

Acknowledgments

This paper owes a great debt to D. J. DiMaria, who took many of the key data and provided key insights into the physics of defect generation in SiO_2. A. Vayshenker took the long-term stress data, and I thank B. P. Linder for extremely helpful discussions about soft breakdown.

References

1. R. H. Dennard, F. H. Gaensslen, H.-N. Yu, V. L. Rideout, E. Bassous, and A. R. LeBlanc, "Design of Ion-Implanted MOSFETs with Very Small Physical Dimensions," *IEEE J. Solid State Circuits* **SC-9,** 256–268 (1974).
2. M. Lenzlinger and E. H. Snow, "Fowler–Nordheim Tunneling into Thermally Grown SiO_2," *J. Appl. Phys.* **40,** 278–283 (1969).
3. I. C. Chen, S. E. Holland, K. K. Young, C. Chang, and C. Hu, "Substrate Hole Current and Oxide Breakdown," *Appl. Phys. Lett.* **49,** 669–671 (1986).
4. J. H. Stathis and D. J. DiMaria, "Reliability Projection for Ultra-Thin Oxides at Low Voltage," *IEDM Tech. Digest*, pp. 167–170 (1998).
5. B. Hoeneisen and C. A. Mead, "Fundamental Limitations in Microelectronics—I. MOS Technology," *Solid-State Electron.* **15,** 819–829 (1972).
6. K. Nagai and Y. Hayashi, "Static Characteristics of 2.3-nm Gate-Oxide MOSFETs," *IEEE Trans. Electron Devices* **35,** 1145–1147 (1988).
7. H. S. Momose, M. Ono, T. Yoshitomi, T. Ohguro, S. Nakamura, M. Saito, and H. Iwai, "Tunneling Gate Oxide Approach to Ultra-High Current Drive in Small Geometry MOSFETs," *IEDM Tech. Digest*, pp. 593–596 (1994).
8. T. H. Ning, "Silicon Technology Directions in the New Millennium," *Proceedings of the International Reliability Physics Symposium*, 2000, pp. 1–6.
9. T. Ghani, K. Mistry, P. Packan, S. Thompson, M. Stettler, S. Tyagi, and M. Bohr, "Scaling Challenges and Device Design Requirements for High Performance Sub-50 nm Gate Length Planar CMOS Transistors," *Symposium on VLSI Technology*, *Digest of Technical Papers*, 2000, pp. 174–175.
10. M. Hirose, M. Koh, W. Mizubayashi, H. Murakami, K. Shibahara, and S. Miyazaki, "Fundamental Limit of Gate Oxide Thickness Scaling in Advanced MOSFETs," *Semicond. Sci. Technol.* **15,** 485–490 (2000).
11. D. A. Muller, T. Sorsch, S. Moccio, F. H. Baumann, K. Evans-Lutterodt, and G. Timp, "The Electronic Structure at the Atomic Scale of Ultrathin Gate Oxides," *Nature* **399,** 758–761 (1999).
12. G. Timp, K. K. Bourdelle, J. E. Bower, F. H. Baumann, T. Boone, R. Cirelli, K. Evans-Lutterodt, J. Garno, A. Ghetti, H. Gossmann, M. Green, D. Jacobson, Y. Kim, R. Kleiman, F. Klemens, A. Kornblit, C. Lochstampfor, W. Mansfield, S. Moccio, D. A. Muller, L. E. Ocola, M. L. O'Malley, J. Rosamilia, J. Sapjeta, P. Silverman, T. Sorsch, D. M. Tennant, W. Timp, and B. Weir, "Progress Toward 10nm CMOS Devices," *IEDM Tech. Digest*, pp. 615–618 (1998).
13. M. V. Fischetti and S. E. Laux, "Performance Degradation of Small Silicon Devices Caused by Long-Range Coulomb Interactions," *Appl. Phys. Lett.* **76,** 2277–2279 (2000).
14. M. A. Alam, B. E. Weir, P. J. Silverman, J. D. Bude, A. Ghetti, Y. Ma, M. M. Brown, D. Hwang, and A. Hamad, "Physics and Prospects of Sub-2nm Oxides," *The Physics and Chemistry of SiO_2 and the Si–SiO_2 Interface–4*, Vol. 2000-2, H. Z. Massoud, I. J. R. Baumvol, M. Hirose, and E. H. Poindexter, Eds., The Electrochemical Society, Pennington, NJ, 2000, pp. 365–376.
15. B. Kaczer, R. Degraeve, N. Pangon, and G. Groeseneken, "The Influence of Elevated Temperature on Degradation and Lifetime Prediction of Thin Silicon-Dioxide Films," *IEEE Trans. Electron Devices* **47,** 1514–1521 (2000).
16. *International Technology Roadmap for Semiconductors* (ITRS), 2000 update, International Sematech, Austin, TX; see *http://public.itrs.net/*.
17. G. D. Wilk, R. M. Wallace, and J. M. Anthony, "High-κ Gate Dielectrics: Current Status and Materials Properties Considerations," *J. Appl. Phys*. **89,** 5243–5275 (2001).
18. E. P. Gusev, E. Cartier, D. A. Buchanan, M. A. Gribelyuk, M. Copel, H. Okorn-Schmidt, and C. D'Emic, "Ultrathin High-*K* Metal Oxides on Silicon: Processing, Characterization and Integration Issues," *Microelectron. Eng*. **59,** 341–349 (2001).
19. L.-K. Han, S. Crowder, M. Hargrove, E. Wu, S.-H. Lo, F. Guarin, E. Crabbé, and L. Su, "Electrical Characteristics and Reliability of Sub-3 nm Gate Oxides Grown on Nitrogen Implanted Silicon Substrates," *IEDM Tech. Digest*, pp. 643–646 (1997).
20. P. E. Nicollian, G. C. Baldwin, K. N. Eason, D. T. Grider, S. V. Hattangady, J. C. Hu, W. R. Hunter, M. Rodder, and A. L. P. Rotondaro, "Extending the Reliability Scaling Limit of SiO_2 Through Plasma Nitridation," *IEDM Tech. Digest*, pp. 545–548 (2000).
21. M. L. Green, E. P. Gusev, R. Degraeve, and E. L. Garfunkel, "Ultrathin (<4nm) SiO_2 and Si-O-N Gate Dielectric Layers for Silicon Microelectronics: Understanding the Processing, Structure, and Physical and Electrical Limits," *J. Appl. Phys.* **90,** 2057–2121 (2001).
22. C. J. Frosch and L. Derick, "Surface Protection and Selective Masking During Diffusion in Silicon," *Proc. Electrochem. Soc*., p. 547 (1957).
23. D. Khang, "Silicon–Silicon Dioxide Surface Device," (Bell Laboratories technical memorandum issued on January 16, 1961), *Semiconductor Devices: Pioneering Papers*, S. M. Sze, Ed., World Scientific Press, Singapore, 1991, pp. 583–596.
24. J. H. Stathis, "Physical and Predictive Models of Ultra Thin Oxide Reliability in CMOS Devices and Circuits," *IEEE Trans. Device & Mater. Reliabil.* **1,** 43–59 (2001).
25. R. W. Thomas, "An Analysis of the Quality and Reliability Supplement to the SIA Roadmap," *Microelectron. Reliabil.* **38,** 861–868 (1988).
26. J. H. Stathis, A. Vayshenker, P. R. Varekamp, E. Y. Wu, C. Montrose, J. McKenna, D. J. DiMaria, L.-K. Han, E. Cartier, R. A. Wachnik, and B. P. Linder, "Breakdown Measurements of Ultra-Thin SiO_2 at Low Voltage," *Symposium on VLSI Technology*, *Digest of Technical Papers*, 2000, pp. 94–95.

27. D. J. DiMaria, E. Cartier, and D. Arnold, "Impact Ionization, Trap Creation, Degradation, and Breakdown in Silicon Dioxide Films on Silicon," *J. Appl. Phys.* **73,** 3367–3384 (1993).
28. E. Y. Wu, J. Aitken, E. Nowak, A. Vayshenker, P. Varekamp, G. Hueckel, J. McKenna, D. Harmon, L.-K. Han, C. Montrose, and R. Dufresne, "Voltage-Dependent Voltage-Acceleration of Oxide Breakdown for Ultra-Thin Oxides," *IEDM Tech. Digest*, pp. 541–544 (2000).
29. M. Takayanagi, S. Takagi, and Y. Toyoshima, "Experimental Study of Gate Voltage Scaling for TDDB Under Direct Tunneling Regime," *Proceedings of the International Reliability Physics Symposium*, 2001, pp. 380–385.
30. B. E. Weir, P. J. Silverman, D. Monroe, K. S. Krisch, M. A. Alam, G. B. Alers, T. W. Sorsch, G. L. Timp, F. Baumann, C. T. Liu, Y. Ma, and D. Huang, "Ultra-Thin Gate Dielectrics: They Break Down, But Do They Fail?," *IEDM Tech. Digest*, pp. 73–76 (1997).
31. S. Tyagi, M. Alavi, R. Bigwood, T. Bramblett, J. Brandenburg, W. Chen, B. Crew, M. Hussein, P. Jacob, C. Kenyon, C. Lo, B. Mcintyre, Z. Ma, P. Moon, P. Nguyen, L. Rumaner, R. Schweinfurth, S. Sivakumar, M. Stettler, S. Thompson, B. Tufts, J. Xu, S. Yang, and M. Bohr, "A 130 nm Generation Logic Technology Featuring 70nm Transistors, Dual Vt Transistors and 6 Layers of Cu Interconnects," *IEDM Tech. Digest*, pp. 567–570 (2000).
32. P. Smeys, V. McGahay, I. Yang, J. Adkisson, K. Beyer, O. Bula, Z. Chen, B. Chu, J. Culp, S. Das, A. Eckert, L. Hadel, M. Hargrove, J. Herman, L. Lin, R. Mann, E. Maciejewski, S. Narasimha, P. O'Neil, S. Rauch, D. Ryan, J. Toomey, L. Tsou, P. Varekamp, R. Wachnik, T. Wagner, S. Wu, C. Yu, P. Agnello, J. Connolly, S. Crowder, C. Davis, R. Ferguson, A. Sekiguchi, L. Su, R. Goldblatt, and T. C. Chen, "A High Performance 0.13 μm SOI CMOS Technology with Cu Interconnects and Low-*k* BEOL Dielectric," *Symposium on VLSI Technology*, *Digest of Technical Papers*, 2000, pp. 184–185.
33. D. J. DiMaria and J. W. Stasiak, "Trap Creation in Silicon Dioxide Produced by Hot Electrons," *J. Appl. Phys.* **65,** 2342–2356 (1989).
34. D. J. DiMaria and J. H. Stathis, "Explanation for the Oxide Thickness Dependence of Breakdown Characteristics of Metal-Oxide-Semiconductor Structures," *Appl. Phys. Lett.* **70,** 2708–2710 (1997).
35. D. J. DiMaria, E. Cartier, and D. A. Buchanan, "Anode Hole Injection and Trapping in Silicon Dioxide," *J. Appl. Phys.* **80,** 304–317 (1996).
36. D. J. DiMaria, "Electron Energy Dependence of Metal-Oxide Semiconductor Degradation," *Appl. Phys. Lett.* **75,** 2427–2428 (1999).
37. J. H. Stathis and E. Cartier, "The Role of Atomic Hydrogen in Degradation and Breakdown of SiO_2 Films," *Extended Abstracts*, *International Conference on Solid State Devices and Materials*, 1996, pp. 791–793.
38. C.-T. Sah, J. Y.-C. Sun, and J. J.-T. Tzou, "Deactivation of the Boron Acceptor in Silicon by Hydrogen," *Appl. Phys. Lett.* **43,** 204–206 (1983).
39. R. Gale, F. J. Feigl, C. W. Magee, and D. R. Young, "Hydrogen Migration Under Avalanche Injection of Electrons in Si Metal-Oxide Semiconductor Capacitors," *J. Appl. Phys.* **54,** 6938–6942 (1983).
40. D. A. Buchanan, A. D. Marwick, D. J. DiMaria, and L. Dori, "Hot-Electron Induced Hydrogen Redistribution in SiO_2," *The Physics and Chemistry of SiO_2 and the Si–SiO_2 Interface–2*, C. R. Helms and B. E. Deal, Eds., Plenum Press, New York, 1993, pp. 481–488.
41. I. P. Lisovskii, V. G. Litovchenko, G. P. Romanova, P. I. Didenko, and E. G. Schmidt, "Proton Release and Defect Creation in MOS Structures Due to High Electric Fields," *Phys. Status Solidi (A)* **142,** 107–116 (1994).
42. Y. Nissan-Cohen and T. Gorczyca, "The Effect of Hydrogen on Trap Generation, Positive Charge Trapping, and Time-Dependent Dielectric Breakdown of Gate Oxides," *IEEE Electron Device Lett.* **9,** 287–289 (1988).
43. L. do Thanh and P. Balk, "Elimination and Generation of Si–SiO_2 Interface Traps by Low Temperature Hydrogen Annealing," *J. Electrochem. Soc.* **135,** 1797–1801 (1988).
44. E. Cartier, J. H. Stathis, and D. A. Buchanan, "Passivation and Depassivation of Silicon Dangling Bonds at the Si/SiO_2 Interface by Atomic Hydrogen," *Appl. Phys. Lett.* **63,** 1510–1512 (1993).
45. J. H. Stathis and E. Cartier, "Atomic Hydrogen Reactions with P_b Centers at the (100) Si/SiO_2 Interface," *Phys. Rev. Lett.* **72,** 2745–2748 (1994).
46. R. E. Stahlbush and E. Cartier, "Interface Defect Formation in MOSFETs by Atomic Hydrogen Exposure," *IEEE Trans. Nucl. Sci.* **41,** 1844–1853 (1994).
47. E. Cartier, D. A. Buchanan, J. H. Stathis, and D. J. DiMaria, "Atomic Hydrogen-Induced Degradation of Thin SiO_2 Gate Dielectrics," *J. Non-Cryst. Sol.* **187,** 244–247 (1995).
48. K. G. Druijf, J. M. M. de Nijs, E. v.d. Drift, V. V. Afanas'ev, E. H. A. Granneman, and P. Balk, "The Microscopic Nature of Donor-Type Si–SiO_2 Interface States," *J. Non-Cryst. Sol.* **187,** 206–210 (1995).
49. E. Cartier and J. H. Stathis, "Atomic Hydrogen-Induced Degradation of the Si/SiO_2 Structure," *Microelectron. Eng.* **28,** 3–10 (1995).
50. D. J. DiMaria and E. Cartier, "Mechanism for Stress-Induced Leakage Currents in Thin Silicon Dioxide Films," *J. Appl. Phys.* **78,** 3883–3894 (1995).
51. E. Cartier and J. H. Stathis, "Hot-Electron Induced Passivation of Silicon Dangling Bonds at the $Si(111)/SiO_2$ Interface," *Appl. Phys. Lett.* **69,** 103–105 (1996).
52. R. E. Stahlbush, "Slow and Fast State Formation Caused by Hydrogen," *The Physics and Chemistry of SiO_2 and the Si–SiO_2 Interface–3*, Vol. 96-1, H. Z. Massoud, C. R. Helms, and E. H. Poindexter, Eds., The Electrochemical Society, Pennington, NJ, 1996, pp. 525–537.
53. A. Szekeres and S. Alexandra, "Low-Temperature Treatment of Si/SiO_2 Structures in an RF Hydrogen Plasma," *Vacuum* **47,** 1483–1486 (1996).
54. J. M. M. de Nijs, K. G. Druijf, and V. V. Afanas'ev, "Hydrogen-Induced Donor States in the MOS System: Hole Traps, Slow States and Interface States," *Fundamental Aspects of Ultrathin Dielectrics on Si-Based Devices*, *NATO Science Series*, E. Garfunkel, E. Gusev, and A. Vul', Eds., 3/47 Ed., Kluwer Academic Publishers, Dordrecht, Netherlands, 1998, pp. 425–430.
55. J. H. Stathis, "Electrical Defects at the SiO_2/Si Interface Studied by EPR," *Fundamental Aspects of Ultrathin Dielectrics on Si-Based Devices*, *NATO Science Series*, E. Garfunkel, E. Gusev, and A. Vul', Eds., 3/47 Ed., Kluwer Academic Publishers, Dordrecht, Netherlands, 1998, pp. 325–333.
56. T.-C. Shen, C. Wang, G. C. Abeln, J. R. Tucker, J. W. Lyding, P. Avouris, and R. E. Walkup, "Atomic-Scale Desorption Through Electronic and Vibrational Excitation Mechanisms," *Science* **268,** 1590–1592 (1995).
57. J. Wu, E. Rosenbaum, B. MacDonald, E. Li, J. Tao, B. Tracy, and P. Fang, "Anode Hole Injection Versus Hydrogen Release: The Mechanism for Gate Oxide Breakdown," *Proceedings of the International Reliability Physics Symposium*, 2000, pp. 27–32.
58. K. Hess, I. C. Kizilyalli, and J. W. Lyding, "Giant Isotope Effect in Hot Electron Degradation of Metal

Oxide Silicon Devices," *IEEE Trans. Electron Devices* **45,** 406–416 (1998).
59. Y. Mitani, H. Satake, H. Ito, and A. Toriumi, "A Study of the Effect of Deuterium on Stress Induced Leakage," *Jpn. J. Appl. Phys. Part 2* **39,** 564–566 (2000).
60. Y. Mitani, H. Satake, H. Ito, and A. Toriumi, "Highly Reliable Gate Oxide Under Fowler–Nordheim Electron Injection by Deuterium Pyrogenic Oxidation and Deuterated Poly-Si Deposition," *IEDM Tech. Digest*, pp. 343–346 (2000).
61. D. J. Dumin and J. R. Maddux, "Correlation of Stress-Induced Leakage Current in Thin Oxides with Trap Generation Inside the Oxides," *IEEE Trans. Electron Devices* **40,** 986–993 (1993).
62. M. Kimura and H. Koyama, "Stress-Induced Low-Level Leakage Mechanism of Ultrathin Silicon Dioxide Films Caused by Neutral Oxide Trap Generation," *Proceedings of the International Reliability Physics Symposium*, 1994, pp. 167–172.
63. H. Kim and H. Hwang, "High-Quality Ultrathin Gate Oxide Prepared by Oxidation in D_2O," *Appl. Phys. Lett.* **74,** 709–710 (1999).
64. C. Schuegraf and C. Hu, "Metal-Oxide Semiconductor Field-Effect-Transistor Substrate Current During Fowler–Nordheim Tunneling Stress and Silicon Dioxide Reliability," *J. Appl. Phys.* **76,** 3695–3700 (1994).
65. M. V. Fischetti, "Model for the Generation of Positive Charge at the Si–SiO_2 Interface Based on Hot-Hole Injection from the Interface," *Phys. Rev. B* **31,** 2099–2113 (1985).
66. J. D. Bude, B. E. Weir, and P. J. Silverman, "Explanation of Stress-Induced Damage in Thin Oxides," *IEDM Tech. Digest*, pp. 179–182 (1998).
67. Y. Lu and C.-T. Sah, "Two Pathways of Positive Oxide-Charge Buildup During Electron Tunneling into Silicon Dioxide Film," *J. Appl. Phys.* **76,** 4724–4727 (1994).
68. P. E. Nicollian, W. R. Hunter, and J. C. Hu, "Experimental Evidence for Voltage Driven Breakdown Models in Ultrathin Gate Oxides," *Proceedings of the International Reliability Physics Symposium*, 2000, pp. 7–15.
69. M. A. Alam, J. D. Bude, and A. Ghetti, "Field Acceleration for Oxide Breakdown—Can an Accurate Anode Hole Injection Model Resolve the E vs. $1/E$ Controversy?," *Proceedings of the International Reliability Physics Symposium*, 2000, pp. 21–26.
70. D. J. DiMaria, "Defect Generation in Ultrathin Silicon Dioxide Films Produced by Anode Hole Injection," *Appl. Phys. Lett.* **77,** 2716–2718 (2000).
71. D. J. DiMaria and J. H. Stathis, "Anode Hole Injection, Defect Generation, and Breakdown in Ultrathin Silicon Dioxide Films," *J. Appl. Phys.* **89,** 5015–5024 (2001).
72. T. Tomita, H. Utsunomiya, Y. Kamakura, and K. Taniguchi, "Hot Hole Induced Breakdown of Thin Silicon Dioxide Films," *Appl. Phys. Lett.* **71,** 3664–3666 (1997).
73. E. M. Vogel, M. D. Edelstein, and J. S. Suehle, "Reliability of Ultra-Thin Silicon Dioxide Under Substrate Hot-Electron, Substrate Hot-Hole and Tunneling Stress," *Microelectron. Reliabil.* **59,** 73–83 (2001).
74. Y. Shi, T. P. Ma, S. Prasad, and S. Dhanda, "Polarity-Dependent Tunneling Current and Oxide Breakdown in Dual-Gate CMOSFETs," *IEEE Electron Device Lett.* **19,** 391–393 (1998).
75. J. W. McPherson and H. C. Mogul, "Underlying Physics of the Thermochemical E Model in Describing Low-Field Time-Dependent Breakdown in SiO_2 Thin Films," *J. Appl. Phys.* **84,** 1513–1523 (1998).
76. J. W. McPherson and R. B. Khamankar, "Molecular Model for Intrinsic Time-Dependent Dielectric Breakdown in SiO_2 Dielectrics and the Reliability Implications for Hyper-Thin Gate Oxide," *Semicond. Sci. Technol.* **15,** 462–470 (2000).
77. G. A. Swartz, "Gate Oxide Integrity of NMOS Transistor Arrays," *IEEE Trans. Electron Devices* **ED-33,** 1826–1829 (1986).
78. J. S. Suehle, P. Chaparala, C. Messick, W. M. Miller, and K. C. Boyko, "Field and Temperature Acceleration of Time-Dependent Dielectric Breakdown in Intrinsic Thin SiO_2," *Proceedings of the International Reliability Physics Symposium*, 1994, pp. 120–125.
79. D. J. Dumin, S. K. Mopuri, S. Vanchinathan, R. S. Scott, R. Subramoniam, and T. G. Lewis, "High Field Related Thin Oxide Wearout and Breakdown," *IEEE Trans. Electron Devices* **42,** 760–772 (1995).
80. M. Kimura, "Field and Temperature Acceleration Model for Time-Dependent Dielectric Breakdown," *IEEE Trans. Electron Devices* **46,** 220–229 (1999).
81. J. McPherson, V. Reddy, K. Banerjee, and H. Le, "Comparison of E and $1/E$ TDDB Models for SiO_2 Under Long-Term/Low-Field Test Conditions," *IEDM Tech. Digest*, pp. 171–174 (1998).
82. K. Umeda, T. Tomita, and K. Taniguchi, "Silicon Dioxide Breakdown Induced by SHE (Substrate Hot Electron) Injection," *Electron. & Commun. Jpn. Part 2* **80,** 11–19 (1997).
83. E. M. Vogel, J. S. Suehle, M. D. Edelstein, B. Wang, Y. Chen, and J. B. Bernstein, "Reliability of Ultrathin Silicon Dioxide Under Combined Substrate Hot-Electron and Constant Voltage Stress," *IEEE Trans. Electron Devices* **47,** 1183–1191 (2000).
84. J. F. Verwey, "Nonavalanche Injection of Hot Carriers into SiO_2," *J. Appl. Phys.* **44,** 2681–2687 (1973).
85. D. J. DiMaria, "Explanation for the Polarity Dependence of Breakdown in Ultrathin Silicon Dioxide Films," *Appl. Phys. Lett.* **68,** 3004–3006 (1996).
86. P. P. Apte and K. C. Saraswat, "Correlation of Trap Generation to Charge-to-Breakdown (Q_{bd}): A Physical-Damage Model of Dielectric Breakdown," *IEEE Trans. Electron Devices* **41,** 1595–1602 (1994).
87. D. J. Dumin, J. R. Maddux, R. S. Scott, and R. Subramoniam, "A Model Relating Wearout to Breakdown in Thin Oxides," *IEEE Trans. Electron Devices* **41,** 1570–1580 (1994).
88. D. J. DiMaria, "Defect Production, Degradation, and Breakdown of Silicon Dioxide Films," *Solid-State Electron.* **41,** 957–965 (1997).
89. D. J. DiMaria and J. H. Stathis, "Ultimate Limit for Defect Generation in Ultra-Thin Silicon Dioxide," *Appl. Phys. Lett.* **71,** 3230–3232 (1997).
90. J. H. Stathis, "Percolation Models for Gate Oxide Breakdown," *J. Appl. Phys.* **86,** 5757–5766 (1999).
91. D. J. DiMaria, "Defect Generation Under Substrate-Hot-Electron Injection into Ultrathin Silicon Dioxide Layers," *J. Appl. Phys.* **86,** 2100–2109 (1999).
92. D. J. DiMaria, "Defect Generation in Field-Effect Transistors Under Channel-Hot-Electron Stress," *J. Appl. Phys.* **87,** 8707–8715 (2000).
93. M. Depas, B. Vermeire, P. W. Mertens, M. Meuris, and M. M. Heyns, "Wear-Out of Ultra-Thin Gate Oxides During High-Field Electron Tunneling," *Semicond. Sci. Technol.* **10,** 753–758 (1995).
94. M. Depas, B. Vermeire, and M. M. Heyns, "Breakdown and Defect Generation in Ultrathin Gate Oxide," *J. Appl. Phys.* **80,** 382–387 (1996).
95. B. Ricco, M. Y. Azbel, and M. H. Brodsky, "Novel Mechanism for Tunneling and Breakdown in Thin SiO_2 Films," *Phys. Rev. Lett.* **51,** 1795–1798 (1983).
96. Y. Nissan-Cohen, J. Shappir, and D. Frohman-Bentchkowsky, "Trap Generation and Occupation

Dynamics in SiO_2 Under Charge Injection Stress," *J. Appl. Phys.* **60,** 2024–2035 (1986).
97. J. W. McPherson, R. B. Khamankar, and A. Shanware, "A Complementary Molecular-Model (Including Field and Current) for TDDB in SiO_2 Dielectrics," *Microelectron. Reliabil.* **40,** 1591–1597 (2000).
98. M. A. Alam, J. D. Bude, B. Weir, P. Silverman, A. Ghetti, D. Monroe, K. P. Cheung, and S. Moccio, "An Anode Hole Injection Percolation Model for Oxide Breakdown—the 'Doom's Day' Scenario Revisited," *IEDM Tech. Digest*, pp. 715–718 (1999).
99. J. Suñé, I. Placencia, N. Barniol, E. Farrés, F. Martín, and X. Aymerich, "On the Breakdown Statistics of Very Thin SiO_2 Films," *Thin Solid Films* **185,** 347–362 (1990).
100. R. Degraeve, G. Groeseneken, R. Bellens, M. Depas, and H. E. Maes, "A Consistent Model for the Thickness Dependence of Intrinsic Breakdown in Ultra-Thin Oxides," *IEDM Tech. Digest*, p. 866 (1995).
101. R. Degraeve, G. Groeseneken, R. Bellens, J. L. Ogier, M. Depas, P. J. Roussel, and H. E. Maes, "New Insights in the Relation Between Electron Trap Generation and the Statistical Properties of Oxide Breakdown," *IEEE Trans. Electron Devices* **45,** 904–911 (1998).
102. S. Lombardo, A. La Magna, C. Spinella, C. Gerardi, and F. Crupi, "Degradation and Hard Breakdown Transient of Thin Gate Oxides in Metal-SiO_2-Si Capacitors: Dependence on Oxide Thickness," *J. Appl. Phys.* **86,** 6382–6391 (1999).
103. M. A. Alam, B. E. Weir, P. J. Silverman, Y. Ma, and D. Hwang, "The Statistical Distribution of Percolation Resistance as a Probe into the Mechanics of Ultra-Thin Oxide Breakdown," *IEDM Tech. Digest*, pp. 529–532 (2000).
104. K. Okada, H. Kubo, A. Ishinaga, and K. Yoneda, "A New Prediction Method for Oxide Lifetime and Its Application to Study Dielectric Breakdown Mechanism," *Symposium on VLSI Technology*, *Digest of Technical Papers*, 1998, pp. 158–159.
105. T. Nigam, R. Degraeve, G. Groeseneken, M. M. Heyns, and H. E. Maes, "A Fast and Simple Methodology for Lifetime Projection of Ultra-Thin Oxides," *Proceedings of the International Reliability Physics Symposium*, 1999, pp. 381–388.
106. A. Ghetti, J. D. Bude, and G. Weber, "TBD Prediction from Measurements at Low Field and Room Temperature Using a New Estimator," *Symposium on VLSI Technology*, *Digest of Technical Papers*, 2000, pp. 218–219.
107. Y. Kamakura, H. Utsunomiya, T. Tomita, K. Umeda, and K. Taniguchi, "Investigations of Hot-Carrier-Induced Breakdown of Thin Oxides," *IEDM Tech. Digest*, pp. 81–84 (1997).
108. K. Umeda and K. Taniguchi, "Hot-Electron-Induced Quasibreakdown of Thin Gate Oxides," *J. Appl. Phys.* **82,** 297–302 (1997).
109. E. Vincent, S. Bruyere, C. Papadas, and P. Mortini, "Dielectric Reliability in Deep Submicron Technologies: From Thin to Ultrathin Oxides," *Microelectron. Reliabil.* **37,** 1499–1506 (1997).
110. E. Vincent, C. Papadas, and G. Ghibaudo, "Temperature Dependence of Charge Build Up Mechanisms and Breakdown Phenomena in Thin Oxides Under Fowler Nordheim Injection," *Solid-State Electron.* **41,** 1001 (1997).
111. M. Houssa, P. W. Mertens, and M. M. Heyns, "Relation Between Stress-Induced Leakage Current and Time-Dependent Dielectric Breakdown in Ultra-Thin Gate Oxides," *Semicond. Sci. Technol.* **14,** 892–896 (1999).
112. K. P. Cheung, C. T. Liu, C.-P. Chang, J. I. Colonell, W. Y.-C. Lai, R. Liu, J. F. Miner, C. S. Pai, H. Vaidya, J. T. Clemens, and E. Hasegawa, "Field Dependent Critical Trap Density for Thin Gate Oxide Breakdown," *Proceedings of the International Reliability Physics Symposium*, 1999, pp. 52–56.
113. K. Okada, "Analysis of the Relationship Between Defect Site Generation and Dielectric Breakdown Utilizing A-Mode Stress Induced Leakage Current," *IEEE Trans. Electron Devices* **47,** 1225–1230 (2000).
114. J. H. Stathis and D. J. DiMaria, "Defect Generation and Reliability of Ultra-Thin SiO_2 at Low Voltage," *The Physics and Chemistry of SiO_2 and the Si–SiO_2 Interface–4*, Vol. 2000-2, H. Z. Massoud, I. J. R. Baumvol, M. Hirose, and E. H. Poindexter, Eds., The Electrochemical Society, Pennington, NJ, 2000, pp. 33–44.
115. L. Niemeyer, L. Pietronero, and H. J. Wiesmann, "Fractal Dimension of Dielectric Breakdown," *Phys. Rev. Lett.* **52,** 1033–1036 (1984).
116. T. Tanamoto and A. Toriumi, "A Percolation Approach to Dielectric Breakdown Statistics," *Jpn. J. Appl. Phys.* **36,** 1439–1442 (1997).
117. T. Ghani, S. Ahmed, P. Aminzadeh, J. Bielefeld, P. Charvat, C. Chu, M. Harper, P. Jacob, C. Jan, J. Kavalieros, C. Kenyon, R. Nagisetty, P. Packan, J. Sebastian, M. Taylor, J. Tsai, S. Tyagi, S. Yang, and M. Bohr, "100 nm Gate Length High Performance/Low Power CMOS Transistor Structure," *IEDM Tech. Digest*, pp. 415–418 (1999).
118. R. Degraeve, N. Pangon, B. Kaczer, T. Nigam, G. Groeseneken, and A. Naem, "Temperature Acceleration of Oxide Breakdown and Its Impact on Ultra-Thin Gate Oxide Reliability," *Symposium on VLSI Technology*, *Digest of Technical Papers*, 1999, pp. 59–60.
119. B. E. Weir, P. J. Silverman, M. A. Alam, F. Baumann, D. Monroe, A. Ghetti, J. D. Bude, G. L. Timp, A. Hamad, T. M. Oberdick, N. X. Zhao, Y. Ma, M. M. Brown, D. Hwang, T. Sorsch, and J. Madic, "Gate Oxides in 50 nm Devices: Thickness Uniformity Improves Projected Temperature," *IEDM Tech. Digest*, pp. 437–440 (1999).
120. E. J. Gumbel, *Statistical Theory of Extreme Values and Some Practical Applications*, National Bureau of Standards Applied Mathematics Series, Vol. 33, U.S. Government Printing Office, Washington, D.C., 1954.
121. D. R. Wolters, "Breakdown and Wearout Phenomena in SiO_2," *Insulating Films on Semiconductors*, M. Schulz and G. Pensl, Eds., Springer-Verlag, Berlin, 1981, pp. 180–194.
122. D. R. Wolters and J. F. Verwey, "Breakdown and Wear-Out Phenomena in SiO_2 Films," *Instabilities in Silicon Devices*, G. Barbottin and A. Vapaille, Eds., North-Holland, Amsterdam, 1986, pp. 315–362.
123. W. G. Kleppmann, "Reliability Predictions for Integrated Circuits from Test Structure Data," *Quality & Reliabil. Eng. Intl.* **5,** 155–163 (1989).
124. W. Nelson, *Accelerated Testing*, John Wiley, New York, 1990.
125. E. Y. Wu, J. H. Stathis, and L.-K. Han, "Ultra-Thin Oxide Reliability for ULSI Applications," *Semicond. Sci. Technol.* **15,** 425–435 (2000).
126. E. Y. Wu, D. L. Harmon, and L.-K. Han, "Interrelationship of Voltage and Temperature Dependence of Oxide Breakdown for Ultrathin Oxides," *IEEE Electron Device Lett.* **21,** 362–364 (2000).
127. R. Degraeve, B. Kaczer, and G. Groeseneken, "Reliability: A Possible Showstopper for Oxide Thickness Scaling?," *Semicond. Sci. Technol.* **15,** 436–444 (2000).
128. E. Y. Wu, E. Nowak, L.-K. Han, R. Dufresne, and W. W. Abadeer, "Nonlinear Characteristics of Weibull Breakdown Distributions and Its Impact on Reliability Projection for Ultra-Thin Oxides," *IEDM Tech. Digest*, pp. 441–444 (1999).

129. E. Y. Wu, E. J. Nowak, and R. P. Vollertsen, "Can Macroscopic Oxide Thickness Uniformity Improve Oxide Reliability?," *IEEE Electron Device Lett.* **21,** 402–404 (2000).
130. E. Y. Wu, E. J. Nowak, R. P. Vollertsen, and L. K. Han, "Weibull Breakdown Characteristics and Oxide Thickness Uniformity," *IEEE Trans. Electron Devices* **47,** 2301–2309 (2000).
131. J. Jacquelin, "Inference of Sampling on Weibull Parameter Estimation," *IEEE Trans. Dielectrics & Electrical Insulation* **3,** 809–816 (1996).
132. B. P. Linder, J. H. Stathis, and D. J. Frank, "Calculating the Error in Long Term Oxide Reliability Estimates," *Proceedings of the International Reliability Physics Symposium*, 2001, pp. 168–171.
133. E. Y. Wu, W. W. Abadeer, L.-K. Han, S.-H. Lo, and G. Hueckel, "Challenges for Accurate Reliability Projections in the Ultra-Thin Oxide Regime," *Proceedings of the International Reliability Physics Symposium*, 1999, pp. 57–65.
134. J. C. Lee, I. C. Chen, and C. Hu, "Modeling and Characterization of Gate Oxide Reliability," *IEEE Trans. Electron Devices* **35,** 2268–2278 (1988).
135. C. Hu, "Gate Oxide Scaling Limits and Projection," *IEDM Tech. Digest*, pp. 319–322 (1996).
136. C. Schuegraf and C. Hu, "Hole Injection SiO_2 Breakdown Model for Very Low Voltage Lifetime Extrapolation," *IEEE Trans. Electron Devices* **41,** 761–767 (1994).
137. B. E. Weir, M. A. Alam, J. D. Bude, P. J. Silverman, A. Ghetti, F. Baumann, P. Diodato, D. Monroe, T. Sorsch, G. L. Timp, Y. Ma, M. M. Brown, A. Hamad, D. Hwang, and P. Mason, "Gate Oxide Reliability Projection to the Sub-2 nm Regime," *Semicond. Sci. Technol.* **15,** 455–461 (2000).
138. K. Okada, H. Kubo, A. Ishinaga, and K. Yoneda, "A Concept of Gate Oxide Lifetime Limited by 'B-Mode' Stress Induced Leakage Currents in Direct Tunneling Regime," *Symposium on VLSI Technology*, *Digest of Technical Papers*, 1999, pp. 57–58.
139. K. Okada, "The Gate Oxide Lifetime Limited by 'B-Mode' Stress Induced Leakage Current and the Scaling Limit of Silicon Dioxides in the Direct Tunneling Regime," *Semicond. Sci. Technol.* **15,** 478–484 (2000).
140. T. Pompl, H. Wurzer, M. Kerber, R. C. W. Wilkins, and I. Eisele, "Influence of Soft Breakdown on NMOSFET Device Characteristics," *Proceedings of the International Reliability Physics Symposium*, 1999, pp. 82–87.
141. W. K. Henson, N. Yang, and J. J. Wortmann, "Observation of Oxide Breakdown and Its Effects on the Characteristics of Ultra-Thin-Oxide nMOSFETs," *IEEE Electron Device Lett.* **20,** 605–607 (1999).
142. R. Degraeve, B. Kaczer, A. De Keersgeiter, and G. Groeseneken, "Relation Between Breakdown Mode and Breakdown Location in Short Channel nMOSFETs and Its Impact on Reliability Specifications," *Proceedings of the International Reliability Physics Symposium*, 2001, pp. 360–366.
143. L. Pantisano and K. P. Cheung, "The Impact of Post Breakdown Gate Leakage on MOSFET RF Performances," *IEEE Electron Device Lett.* **22,** 585–587 (2001).
144. M. Houssa, T. Nigam, P. W. Mertens, and M. M. Heyns, "Soft Breakdown in Ultrathin Gate Oxides: Correlation with the Percolation Theory of Nonlinear Conductors," *Appl. Phys. Lett.* **73,** 514–516 (1998).
145. M. Houssa, T. Nigam, P. W. Mertens, and M. M. Heyns, "Model for the Current-Voltage Characteristics of Ultrathin Gate Oxides After Soft Breakdown," *J. Appl. Phys.* **84,** 4351–4355 (1998).
146. J. Suñé, E. Miranda, M. Nafría, and X. Aymerich, "Point Contact Conduction at the Oxide Breakdown of MOS Devices," *IEDM Tech. Digest*, pp. 191–194 (1998).
147. N. Vandewalle, M. Ausloos, M. Houssa, P. W. Mertens, and M. M. Heyns, "Non-Gaussian Behavior and Anticorrelations in Ultrathin Gate Oxides After Soft Breakdown," *Appl. Phys. Lett.* **74,** 1579–1581 (1999).
148. E. Miranda, J. Suñé, R. Rodríguez, M. Nafría, F. Martín, and X. Aymerich, "Soft Breakdown in Ultrathin SiO_2 Layers: The Conduction Problem from a New Point of View," *Jpn. J. Appl. Phys.* **38,** 2223–2226 (1999).
149. J. Suñé, E. Miranda, M. Nafría, and X. Aymerich, "Modeling the Breakdown Spots in Silicon Dioxide Films as Point Contacts," *Appl. Phys. Lett.* **75,** 959–961 (1999).
150. E. Miranda, J. Suñé, R. Rodriguez, M. Nafría, X. Aymerich, L. Fonseca, and F. Campabadal, "Soft Breakdown in Ultrathin (3–5 nm) Gate Dielectrics," *IEEE Trans. Electron Devices* **47,** 82–89 (2000).
151. E. Miranda and J. Suñé, "Modeling the Conduction Characteristics of Broken Down Gate Oxides in MOS Structures," *Microelectron. Reliabil.* **40,** 1599–1603 (2000).
152. J. Suñé, M. Nafria, X. Oriols, R. Rodriguez, and X. Aymerich, "Failure Physics of Ultra-Thin SiO_2 Gate Oxides Near Their Scaling Limit," *Semicond. Sci. Technol.* **15,** 445–454 (2000).
153. J. Suñé and E. Miranda, "Soft and Hard Breakdown of Ultrathin SiO_2 Gate Oxides," *The Physics and Chemistry of SiO_2 and the Si–SiO_2 Interface–4*, Vol. 2000-2, H. Z. Massoud, I. J. R. Baumvol, M. Hirose, and E. H. Poindexter, Eds., The Electrochemical Society, Pennington, NJ, 2000, pp. 333–344.
154. J. Suñé and E. Miranda, "Post Soft Breakdown Conduction in SiO_2 Gate Oxides," *IEDM Tech. Digest*, pp. 533–536 (2000).
155. S. Lombardo, F. Crupi, and J. H. Stathis, "Softening of Breakdown in Ultra-Thin Gate Oxide nMOSFETs at Low Inversion Layer Density," *Proceedings of the International Reliability Physics Symposium*, 2001, pp. 163–167.
156. K. Okada, S. Kawasaki, and Y. Hirofuji, "New Experimental Findings on Stress Induced Leakage Current of Ultra Thin Silicon Dioxides," *Extended Abstracts of the 1994 International Conference on Solid State Devices and Materials*, 1994, pp. 565–567.
157. S.-H. Lee, B.-J. Cho, J.-C. Kim, and S.-H. Choi, "Quasi-Breakdown of Ultrathin Gate Oxide Under High Field Stress," *IEDM Tech. Digest*, pp. 605–607 (1994).
158. S. Bruyere, E. Vincent, and G. Ghibaudo, "Quasi-Breakdown in Ultra-Thin SiO_2 Films: Occurrence Characterization and Reliability Assessment Methodology," *Proceedings of the International Reliability Physics Symposium*, 2000, pp. 48–54.
159. K. Okada and S. Kawasaki, "New Dielectric Breakdown Model of Local Wearout in Ultra Thin Silicon Dioxides," *Extended Abstracts of the 1995 International Conference on Solid State Devices and Materials*, 1995, pp. 473–475.
160. K. Okada, "Extended Time Dependent Dielectric Breakdown Model Based on Anomalous Gate Area Dependence of Lifetime in Ultra Thin Silicon Dioxides," *Jpn. J. Appl. Phys.* **36,** 1443–1447 (1997).
161. E. Miranda, J. Suñé, R. Rodríguez, M. Nafría, and X. Aymerich, "Detection and Fitting of the Soft Breakdown Failure Mode in MOS Structures," *Solid-State Electron.* **43,** 1801–1805 (1999).
162. E. Miranda, J. Suñé, R. Rodríguez, M. Nafría, and X. Aymerich, "A Function-Fit Model for the Soft Breakdown Failure Mode," *IEEE Electron Device Lett.* **20,** 265–267 (1999).
163. F. Crupi, R. Degraeve, G. Groeseneken, T. Nigam, and H. Maes, "On the Properties of the Gate and Substrate Current After Soft Breakdown in Ultrathin Oxide

Layers," *IEEE Trans. Electron Devices* **45,** 2329–2334 (1998).
164. F. Crupi, R. Degraeve, G. Groeseneken, T. Nigam, and H. Maes, "Investigation and Comparison of the Noise in the Gate and Substrate Current After Soft-Breakdown," *Jpn. J. Appl. Phys.* **38,** 2219–2222 (1999).
165. Y. D. He, H. Guan, M. F. Li, B. J. Cho, and Z. Dong, "Conduction Mechanism Under Quasibreakdown of Ultrathin Gate Oxide," *Appl. Phys. Lett.* **75,** 2432–2434 (1999).
166. H. Guan, M. F. Li, Y. He, B. J. Cho, and Z. Dong, "A Thorough Study of Quasi-Breakdown Phenomenon of Thin Gate Oxide in Dual-Gate CMOSFETs," *IEEE Trans. Electron Devices* **47,** 1608–1615 (2000).
167. T. Sakura, H. Utsunomiya, Y. Kamakura, and K. Taniguchi, "A Detailed Study of Soft- and Pre-Soft-Breakdowns in Small Geometry MOS Structures," *IEDM Tech. Digest*, pp. 183–186 (1998).
168. T. Tomita, H. Utsunomiya, T. Sakura, Y. Kamakura, and K. Taniguchi, "A New Soft Breakdown Model for Thin Thermal SiO_2 Films Under Constant Current Stress," *IEEE Trans. Electron Devices* **46,** 159–164 (1999).
169. G. B. Alers, B. E. Weir, M. R. Frei, and D. Monroe, "J-Ramp on Sub-3nm Dielectrics: Noise as a Breakdown Criterion," *Proceedings of the International Reliability Physics Symposium*, 1999, pp. 410–413.
170. F. Crupi, B. Neri, and S. Lombardo, "Pre-Breakdown in Thin SiO_2 Films," *IEEE Electron Device Lett.* **21,** 319–321 (2000).
171. J. C. Jackson, T. Robinson, O. Oralkan, D. J. Dumin, and G. A. Brown, "Nonuniqueness of Time-Dependent-Dielectric-Breakdown Distributions," *Appl. Phys. Lett.* **71,** 3682–3684 (1997).
172. J. C. Jackson, T. Robinson, O. Oralkan, D. J. Dumin, and G. A. Brown, "Differentiation Between Electric Breakdowns and Dielectric Breakdown in Thin Silicon Oxides," *J. Electrochem. Soc.* **45,** 1033–1038 (1997).
173. S. Lombardo, F. Crupi, A. La Magna, C. Spinella, A. Terrasi, A. La Mantia, and B. Neri, "Electrical and Thermal Transient During Dielectric Breakdown of Thin Oxides in Metal–SiO_2 Capacitors," *J. Appl. Phys.* **84,** 472–479 (1998).
174. S. Lombardo, F. Crupi, C. Spinella, and B. Neri, "Transients During Pre-Breakdown and Hard Breakdown of Thin Gate Oxides in Metal–SiO_2–Si Capacitors," *Mater. Sci. Semicond. Process.* **2,** 359–367 (1999).
175. S. Lombardo, A. La Magna, C. Gerardi, M. Alessandri, and F. Crupi, "Soft Breakdown of Gate Oxides in Metal–SiO_2–Si Capacitors Under Stress with Hot Electrons," *Appl. Phys. Lett.* **75,** 1161–1163 (1999).
176. P. Blöchl and J. H. Stathis, "Aspects of Defects in Silica Related to Dielectric Breakdown of Gate Oxides in MOSFETs," *Physica B* **273–274,** 1022–1026 (1999).
177. M. A. Alam, B. Weir, J. Bude, P. Silverman, and D. Monroe, "Explanation of Soft and Hard Breakdown and Its Consequences for Area Scaling," *IEDM Tech. Digest*, pp. 449–452 (1999).
178. H. Satake and A. Toriumi, "Dielectric Breakdown Mechanism of Thin-SiO_2 Studied by the Post-Breakdown Resistance Statistics," *IEEE Trans. Electron Devices* **47,** 741–745 (2000).
179. H. Satake and A. Toriumi, "SiO_2 Dielectric Breakdown Mechanism Studied by the Post-Breakdown Resistance Statistics," *Semicond. Sci. Technol.* **15,** 471–477 (2000).
180. B. P. Linder, J. H. Stathis, R. A. Wachnik, E. Wu, S. A. Cohen, A. Ray, and A. Vayshenker, "Gate Oxide Breakdown Under Current Limited Constant Voltage Stress," *Symposium on VLSI Technology*, *Digest of Technical Papers*, 2000, pp. 214–215.
181. A. Toriumi, S. Takagi, and H. Satake, "Study of Soft Breakdown in Thin SiO_2 Films by Carrier-Separation Technique and Breakdown Transient Modulation," *The Physics and Chemistry of SiO_2 and the Si–SiO_2 Interface–4*, Vol. 2000-2, H. Z. Massoud, I. J. R. Baumvol, M. Hirose, and E. H. Poindexter, Eds., The Electrochemical Society, Pennington, NJ, 2000, pp. 399–407.
182. B. P. Linder, D. J. Frank, J. H. Stathis, and S. A. Cohen, "Transistor-Limited Constant Voltage Stress of Gate Dielectrics," *Symposium on VLSI Technology*, *Digest of Technical Papers*, 2001, pp. 93–94.
183. B. Kaczer, R. Degraeve, G. Groeseneken, M. Rasras, S. Kubicek, E. Vandamme, and G. Badenes, "Impact of MOSFET Oxide Breakdown on Digital Circuit Operation and Reliability," *IEDM Tech. Digest*, pp. 553–556 (2000).
184. M. Morita and T. Ohmi, "Pre-Gate Oxide Si Surface Control," *The Physics and Chemistry of SiO_2 and the Si–SiO_2 Interface–2*, C. R. Helms and B. E. Deal, Eds., Plenum Press, New York, 1993, pp. 199–206.
185. T. Hattori, H. Nohira, and K. Takahashi, "The Initial Growth Steps of Ultrathin Gate Oxides," *Microelectron. Eng.* **48,** 17–24 (1999).
186. J. Yugami, "Oxide Reliability Improvement Controlling Microstructures of Substrate/Oxide and Oxide/Gate Interfaces," *Superlatt. & Microstruct.* **27,** 395–404 (2000).
187. J. Sapjeta, M. L. Green, J. P. Chang, P. J. Silverman, T. W. Sorsch, B. E. Weir, W. Gladden, Y. Ma, C. Y. Sung, and W. N. Lennard, "Relationship Between Interfacial Roughness and Dielectric Reliability for Silicon Oxynitride Gate Dielectrics Processes with Nitric Oxide," *Ultrathin SiO_2 and High-K Materials for ULSI Gate Dielectrics*, Vol. 567, H. R. Huff, C. A. Richter, M. L. Green, G. Lucovsky, and T. Hattori, Eds., Materials Research Society, Warrendale, PA, 1999, pp. 289–294.
188. K. Eriguchi, Y. Harada, and M. Niwa, "Effects of Strained Layer Near SiO_2–Si Interface on Electrical Characteristics of Ultrathin Gate Oxides," *J. Appl. Phys.* **87,** 1990–1995 (2000).
189. Y. Harada, K. Eriguchi, M. Niwa, T. Watanabe, and I. Ohdomari, "Impact of Structural Strained Layer Near SiO_2/Si Interface on Activation Energy of Time-Dependent Dielectric Breakdown," *Jpn. J. Appl. Phys.* **39,** 4687–4691 (2000).
190. M. Houssa, T. Nigam, P. W. Mertens, and M. M. Heyns, "Effect of Extreme Surface Roughness on the Electrical Characteristics of Ultra-Thin Gate Oxides," *Solid-State Electron.* **43,** 159–167 (1999).
191. H. S. Momose, M. Ono, T. Yoshitomi, T. Ohguro, S. Nakamura, M. Saito, and H. Iwai, "Prospects for Low-Power, High-Speed MPUs Using 1.5nm Direct-Tunneling Gate Oxide MOSFETs," *Solid-State Electron.* **41,** 707–714 (1997).

Received September 16, 2001; accepted for publication December 28, 2001

James H. Stathis *IBM Research Division, Thomas J. Watson Research Center, P.O. Box 218, Yorktown Heights, New York 10598 (stathis@us.ibm.com).* Dr. Stathis is a Research Staff Member in the Silicon Technology Department of the IBM Research Division. He received a bachelor's degree in physics from Washington University in St. Louis in 1980, and a Ph.D. in physics from the Massachusetts Institute of Technology in 1986, joining IBM the same year. His work has mainly been on experimental studies of point defects in semiconductors and insulators, with emphasis on SiO_2 and the Si/SiO_2 interface. He is the author or coauthor of more than 80 research papers and has presented more than 20 invited talks at major technical conferences. Dr. Stathis has received an IBM Outstanding Technical Achievement Award for work on point defects in SiO_2.