Available online at www.sciencedirect.com
SCIENCE DIRECT®
Microelectronics Reliability 45 (2005) 419–426
MICROELECTRONICS RELIABILITY
www.elsevier.com/locate/microrel

Introductory Invited Paper

Detailed study and projection of hard breakdown evolution in ultra-thin gate oxides

J.S. Suehle [a,*], B. Zhu [b], Y. Chen [c], J.B. Bernstein [b]

[a] Semiconductor Electronics Division, NIST, Gaithersburg, MD 20899, United States
[b] Department of Mechanical Engineering, University of Maryland, College Park, MD 20742, United States
[c] Electronics Parts Engineering Office, Jet Propulsion Laboratory, Pasadena, CA 91109, United States

Received 7 October 2004
Available online 15 December 2004

Abstract

The mechanism responsible for post-soft breakdown leakage current increase in ultra-thin oxides depends on the nature of the conducting filament formed at the instant of dielectric breakdown. The conductance of the filament formed during soft breakdown has been observed to be either stable until hard breakdown occurs or to increase continually with time. The acceleration factors for predicting hard breakdown are different in each case. Recent experimental results suggest that the "hardness" of the first breakdown influences the type of conducting filament formed during the soft breakdown event the time in which hard breakdown subsequently occurs. Electron current-induced defect formation appears to be the driving force for the eventual hard breakdown event.
© 2004 Elsevier Ltd. All rights reserved.

1. Introduction

There has been considerable interest in studying the time-dependent increase of post-soft breakdown conduction in ultra-thin gate oxides after it had been observed that some circuits remain functional even after a gate oxide failure [1–5]. Since some circuit architectures can tolerate the oxide leakage current immediately after soft breakdown, the end of circuit life can be redefined as the point in time when the leakage current through the oxide increases to an unacceptable level. In this case, gate oxide breakdown does not necessarily imply the end of product life. The accurate projection of circuit lifetime will require the correct physical model and acceleration parameters that describe the process by which soft breakdown (tolerable gate leakage current) evolves into hard breakdown (catastrophic level of leakage current).

The gate leakage current following dielectric breakdown is due to a permanent filament caused by a sudden local increase in temperature during the breakdown event. Suñé et al. [6] proposed that the filament resembled a quantum point contact (QPC). Indeed, the post-breakdown current vs. voltage curves have been successfully fitted to a model using two parameters: ϕ, the energy barrier height of the constriction, and α, which is correlated to the shape or thickness of the constriction [7]. High resolution TEM images have revealed that the breakdown filament can be the result of dielectric breakdown-induced epitaxial growth of Si from the cathode to the anode [8].

* Corresponding author. Tel.: +1 301 975 2247; fax: +1 301 975 5668.
E-mail address: john.suehle@nist.gov (J.S. Suehle).

0026-2714/$ - see front matter © 2004 Elsevier Ltd. All rights reserved.
doi:10.1016/j.microrel.2004.10.018

The progression of hard breakdown from soft breakdown has been observed to occur in two different modes. The conduction of the breakdown filament in the first mode remains stable until a more damaging hard breakdown occurs. The hard breakdown is uncorrelated temporally and spatially to the first soft breakdown filament. A technique known as the hard breakdown prevalence method [9,10] has been used to predict the hard breakdown failure distribution successfully by shifting the time distribution of the *first* occurrence of breakdown by $\ln(\alpha_{HBD})$ [10]. α_{HBD} is defined as the cumulative hard breakdown prevalence ratio. A second methodology was developed to model the failure distributions of statistically independent successive soft breakdowns. This approach assumes that several statistically independent soft breakdowns can occur before the total leakage current in the gate oxide is large enough to cause circuit failure and the breakdown filaments must remain stable with subsequent stress time [11,12]. It has been shown that the prevalence method and the successive breakdown approach are statistically compatible and can be related [13]. The right plot in Fig. 1A illustrates an example of the leakage current vs. time characteristic through an oxide that exhibits multiple statistically independent soft breakdowns. The lifetime of the circuit can be defined as when the total leakage current exceeds some critical value. The schematic on the left of Fig. 1A is a representation of how the total leakage current through an oxide is the sum of the leakage through individual breakdown filaments.

A second mode has been observed when the conducting filament formed during the first soft breakdown does not remain stable and continues to degrade as stress continues [3–5]. This mode is known as "progressive breakdown" [4]. Accurately projecting circuit life in this case requires two sets of acceleration factors: one set to predict oxide wear-out (or the onset of the first soft breakdown event) and one set to predict the evolution into hard breakdown. The right plot in Fig. 1B illustrates the leakage current vs. time characteristic observed for unstable breakdown filaments that continually degrade with time, and the schematic to the left is a representation of a progressively growing breakdown filament. Wu et al. [14] recently showed that post-breakdown behavior can exhibit several competing modes requiring different analysis procedures.

This paper describes a detailed study of two competing post-soft breakdown modes observed for ultra-thin oxides during constant voltage stress. A brief description of the experimental details will be given in Section 2. It will be shown in Section 3 that the degree of the first breakdown "hardness" will influence the evolution and progression of the subsequent hard breakdown event. Section 4 will show that the acceleration factors are different for each mode, indicating different physical mechanisms are involved in the evolution and formation of the final hard breakdown event. Unstable filaments that result from the first soft breakdown progressively degrade and structurally change until their leakage current becomes unacceptably large. A set of voltage and temperature acceleration parameters different from oxide wear-out is necessary to project the progression of oxide leakage current with time. Section 5 will present a discussion of the results, and Section 6 will provide conclusions.

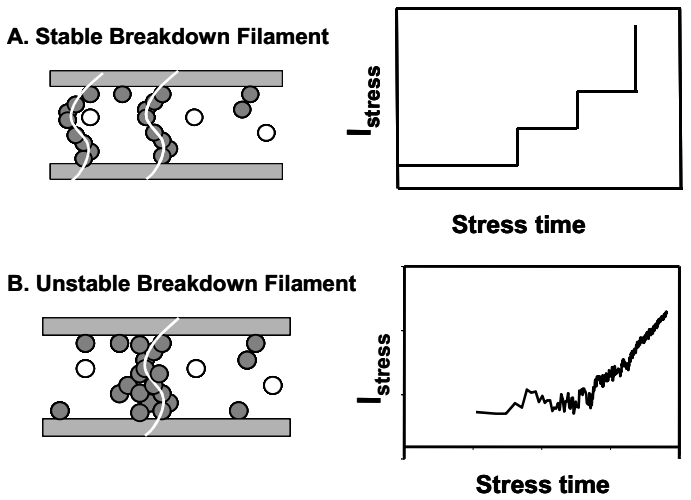

Fig. 1. Typical gate current vs. stress time illustrating: (A) stable breakdown filament formation following soft breakdown where each new filament formation results in a current step and, (B) non-stable filament formation resulting in progressive breakdown.

2. Experimental details

N-channel MOSFETs with a gate oxide thickness ranging from 1.6 nm to 2.2 nm and an active area of $2.5 \times 10^{-8}\,cm^2$ ($L = 0.25\,\mu m$, $W = 10\,\mu m$) were used in this study. All stresses were performed with constant voltage with the channel biased in accumulation. The stress was interrupted when the first breakdown was detected, which was defined as a 100% increase in the stress current. The first breakdown occurs when a percolation filament is formed by defects that are generated randomly throughout the volume of the oxide film. The filament connects the anode and cathode, causing a sudden surge of current through the filament. The current surge can locally produce a large amount of heat, causing permanent structural damage along the path of the filament.

Discrete resistors of various values were inserted between the gate and the power supply to limit the power dissipation (and heat) available during the first breakdown event. This allowed the evolution of hard breakdown to be studied for different degrees of "hardness" induced during the first breakdown. The effect of different resistor values on the conductance of the post-soft breakdown filament is shown in Fig. 2 for values of 10 kΩ 30 kΩ, and 500 kΩ. The filament leakage current shown in the plot was measured by applying 1.5 V on the gate. All devices were stressed at 150 °C with 3.9 V applied to the gate until the first breakdown was detected. Note the magnitude of the leakage current increases as the resistance value decreases, indicating that the severity of the breakdown event was larger. The distribution of the leakage current values can extend over three orders of magnitude for a given resistance value as shown in the figure.

The evolution of hard breakdown was studied by removing the gate resistor and continuing the stress at

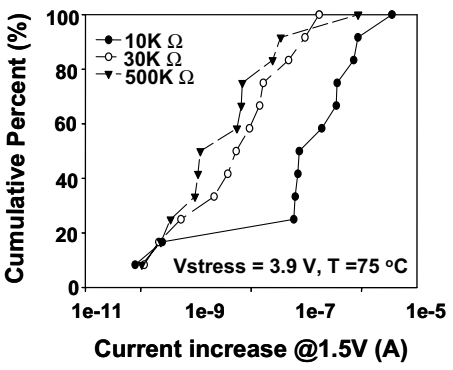

Fig. 2. Breakdown filament conduction distribution measured at a gate voltage of 1.5 V after the first breakdown for three different gate resistor values. The 10 kΩ resistor results in the largest post-breakdown conduction values.

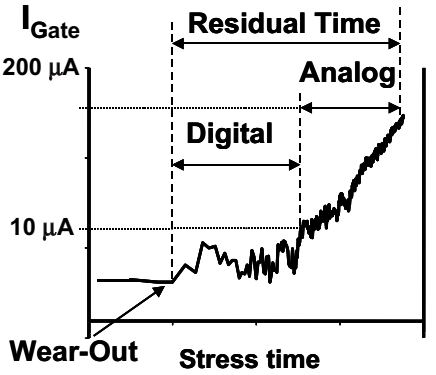

Fig. 3. Typical gate current vs. stress time for non-stable breakdown filaments illustrating "digital and "analog" phases of the residual time to hard breakdown.

voltages and temperatures that may be different than those used during the first breakdown in a manner similar to Linder's study [3]. The second stress was continued until hard breakdown occurred (defined when the gate current exceeded 200 μA).

Fig. 3 shows a typical current vs. time characteristic for a *non-stable* breakdown filament exhibited by the majority of the devices used in this study. Here various phases during the evolution of hard breakdown are identified. "Wear-out" is when a percolation path is first formed through the gate dielectric and causes the first breakdown event. We define a "digital" phase as the region in time when the current increases from 1 μA to 10 μA and the "analog" phase as the region in time when the current increases from 10 μA to 200 μA similar to the study in [3]. The digital phase is usually accompanied by noisy fluctuations in current. Sakura et al. also observed similar phases in their study [15]. The current degradation rate is defined as the slope of the "analog" phase. The "residual time" is the time between "wear-out" and the hard breakdown event. It will be shown later that the length of both the "digital" and "analog" phases exhibited a temperature and voltage dependence.

An example of *stable* filament behavior is shown in Fig. 4. Note that unlike the case of the unstable filament, there is very small current noise until hard breakdown occurs. A "digital" and an "analog" phase can still be defined. However, the length of the "analog" phase was the same for all stress conditions and did not exhibit either a temperature or a voltage dependence as observed in the *unstable* filament case. The length of the "analog" phase in stable filaments is most likely due to the behavior of the hard breakdown current due to the series resistance of the test structure and other factors.

A technique developed by Degraeve et al. [16] was used to map the breakdown location in the MOSFET

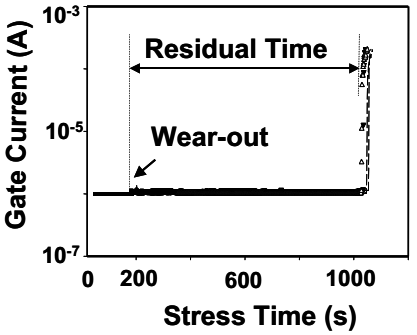

Fig. 4. Typical current vs. voltage characteristics following the first soft breakdown for devices that exhibit stable filament formation. Note that the digital phase is less noisy than shown in Fig. 3, and the analog phase is more abrupt than that shown for non-stable filaments.

channel spatially for the first and final breakdown event. The location of a breakdown event can be located by measuring the source, drain, and gate leakage currents after the first soft breakdown. The ratio, $I_D/(I_D + I_S)$, is calculated. A ratio near 0 indicates the breakdown occurred near the source region, a ratio near 1 indicates the breakdown is near the drain region, and a ration close to 0.5 indicates the breakdown is located near the center of the channel. Fig. 5 shows the ratio obtained after the first breakdown plotted against the ratio obtained after the final breakdown. The plot on the left is for breakdowns that were identified as non-stable filaments (i.e., the final breakdown was caused by a progressive deterioration of the initial filament), and the right plot is for filaments identified as stable. Note that the ratios are generally correlated and suggest that the first and final breakdown positions are possibly related.

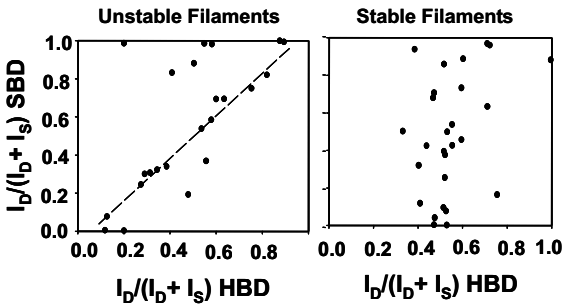

Fig. 5. The ratio, $I_D/(I_D + I_S)$, determined after the first breakdown plotted against the value determined after the final hard breakdown. The left plot is for devices that exhibited non-stable filament formation. Note that the position of the breakdown in the channel shows a degree of correlation. The right plot is for devices that exhibited stable breakdown filaments. In this case there is little correlation in the position of the breakdown for the first and final hard breakdown.

If the ratios are not correlated between the first and final breakdown, it is assumed that the first and final breakdown locations are independent. The right plot shows small correlation, suggesting spatially independent breakdown sites.

3. Effect of initial breakdown "hardness"

Fig. 6 shows the effect of varying the gate series resistance value on the residual time as defined in Figs. 3 and 4. The residual time is the time until hard breakdown occurs (in our case 200 μA) after the first soft breakdown. Four distributions are shown with the stress conditions for the first breakdown shown in the inset. The second stress condition used for obtaining hard breakdown was the same for all of the distributions (V_{stress} = 3.0 V, T = 175 °C). The distributions appear bimodal, especially the distribution obtained with the lowest resistor value (10 kΩ). Wu also reported bimodal residual time distributions and explained the dual distribution as a mixing of stable and unstable breakdown modes [14]. The devices in the distribution shown for the case where the gate series resistor was 10 kΩ experienced the "hardest" first breakdown and exhibited the largest post-breakdown leakage currents (shown in Fig. 2). Note that this distribution also appears to be more bimodal than the other distributions for devices that experienced "softer" first breakdowns. Fig. 6 also shows a trend that devices that had harder first breakdowns generally exhibit shorter residual times.

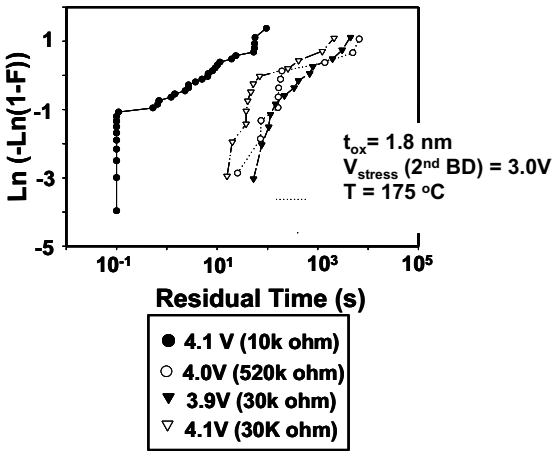

Fig. 6. Residual time distributions for the final hard breakdown (defined as 200 μA) for different levels of filament conduction (hardness) induced by the first breakdown. The devices that exhibited the largest conduction following the first breakdown showed the largest dispersion in the residual time distribution.

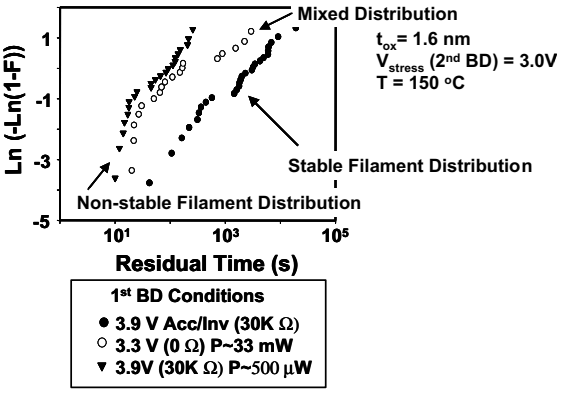

Fig. 7. Residual time distributions for 1.6 nm thick gate oxide for three different first breakdown conditions. The figure shows a distribution comprised of non-stable breakdown filaments (left), a distribution comprised of stable filaments (right), and a distribution comprised of a mixture (center).

Another set of residual times distributions is shown in Fig. 7. In this case all of the post-soft breakdown stresses were performed at 3.3 V and 150 °C. Different stress conditions were used to induce the first breakdown such that residual time distributions of only stable filaments, non-stable filaments, and a combination on both types could be produced. A technique was demonstrated in [17] where the initial breakdown filament could be made stable during the second stress (to induce hard breakdown) by *reversing* the polarity of the stress used to induce the first breakdown. The left most distribution in Fig. 7 is composed of only non-stable breakdown filaments, and the right most distribution is composed of only stable breakdown filaments using the technique presented in [17]. The center distribution is a combination of both types of filaments. It is important to note that the prevalence ratio method (α_{HBD}) can be used to estimate the upper half of a mixed mode bimodal residual time distribution.

4. Acceleration parameters for projecting residual time

This section will show that the voltage and temperature acceleration factors used to predict the residual time will depend on the type of breakdown filament formed after the first breakdown event. Fig. 8 shows the voltage and temperature acceleration factors for the evolution of hard breakdown for non-stable breakdown filaments. The left plot in the figure shows the voltage acceleration for temperatures in range from 25 °C to 275 °C. The slopes of the plots are nearly independent of temperature and have a value of about 3.2 dec/V. Linder et al. [3] observed a somewhat large voltage acceleration of 5 dec/V. It was observed that the temperature dependence was not Arrhenius and exhibited a much better fit when plotted to $T_{BD} \sim \exp(AT)$ ($A = 1 \times 10^{-2}$) as shown in the right plot of Fig. 8. Note that the temperature acceleration is not voltage dependent. The exact mechanism by which the non-stable breakdown filament continues to grow and result in a hard breakdown event is not currently known and will dictate the functional form of the acceleration parameters.

For comparison, the acceleration factors for stable and non-stable filaments are shown plotted together in Fig. 9. As shown in both plots, the voltage and temperature acceleration for the evolution of hard breakdown depends on the type of filament formed during the first soft breakdown event. The left plot also shows the voltage acceleration factor for the first breakdown event (the conventional voltage acceleration for dielectric wear-out when a percolation path is initially formed by defect generation) is identical to the voltage acceleration when a stable filament is formed. Similarly, the temperature acceleration of conventional dielectric wear-out is also identical to that observed when stable breakdown filaments are formed during the first breakdown event.

The similarity of the acceleration parameters between conventional dielectric wear-out and the evolution of

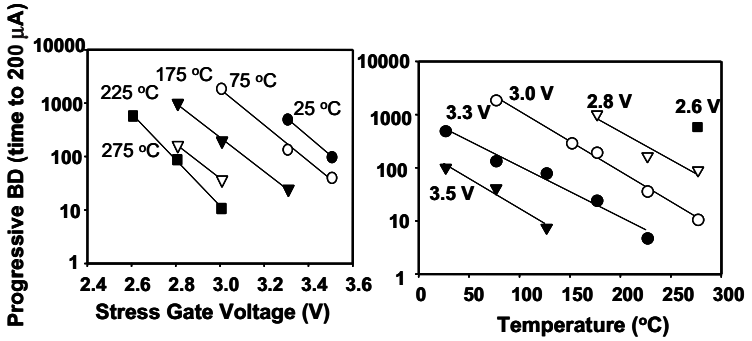

Fig. 8. Voltage (left plot) and temperature (right plot) acceleration for the residual time to hard breakdown for non-stable filaments. The temperature acceleration is not Arrhenius and is independent of voltage, and the voltage acceleration factor is independent of temperature.

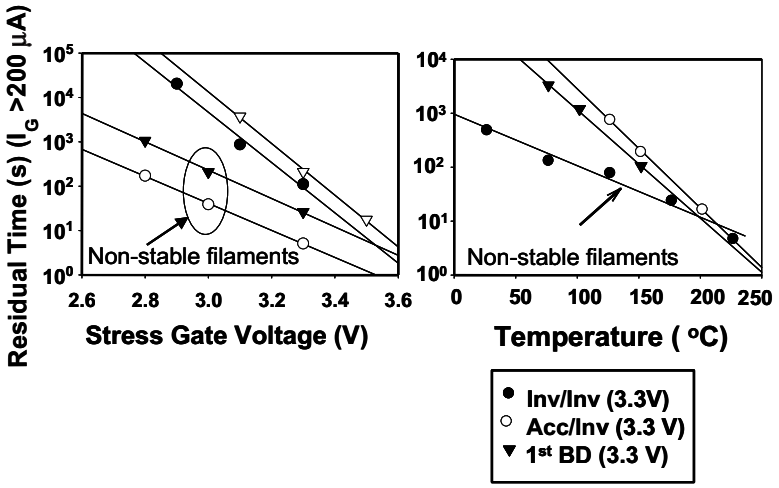

Fig. 9. Comparison of voltage (left plot) and temperature (right plot) dependence of the residual time to hard breakdown for unstable filaments, stable filaments, and device wear-out. Note that the acceleration parameters for stable filaments and oxide wear-out are nearly identical.

hard breakdown from a stable filament indicates that the first soft breakdown and final hard breakdown share a common physical process. Tunneling electrons continue to generate defects in the non-damaged region of the oxide after the first breakdown filament is formed. Eventually, a critical number of defects are created and trigger the formation of a second percolation filament leading to hard breakdown.

5. Discussion

Additional measurements were conducted to study the physical mechanism responsible for non-stable filament growth after the first breakdown event. Substrate hot-electron injection (SHEI) experiments were conducted following the formation of the first filament to investigate if tunneling current is the driving force for subsequent filament growth. SHEI allows additional electrons to be injected into the gate dielectric independent of the applied gate voltage. This is accomplished by biasing the *n*-MOSFET device into inversion and applying a gate voltage relative to the grounded source and drain regions. A separate diffusion is used as an injector for additional carriers which are subsequently swept into the channel region by an applied substrate potential.

Fig. 10 shows the effect of SHEI on the residual time distributions of non-stable breakdown filaments. During the second stress to induce hard breakdown, approximately 1 μA of additional current was injected by SHEI while maintaining the same gate voltage stress. The residual time of the "analog" phase and the "digital" phase was separated to determine if the phases had the same physical origin. The figure shows that the residual time of both phases was significantly shortened by the injection of additional electrons into the gate. Electron current driven defect generation appears to accelerate further degradation of the non-stable breakdown filament. Even though the "digital" and "analog" phases of filament growth exhibit different current vs. time characteristics during the evolution of hard breakdown, both phases are affected similarly by additional electron current.

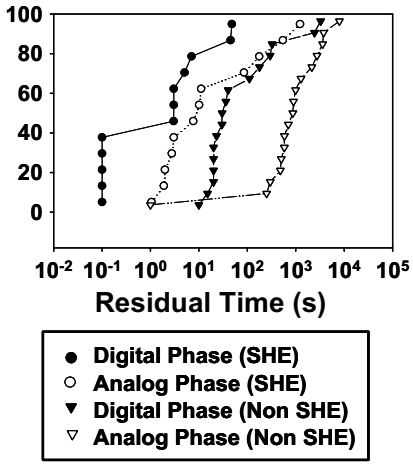

Fig. 10. Residual time distributions of the "digital" and "analog" phases of the residual time to hard breakdown before and after substrate hot electron injection for non-stable filaments. Both phases are shortened by the additional injection of electron current, indicating that electrons play a role in subsequent filament degradation.

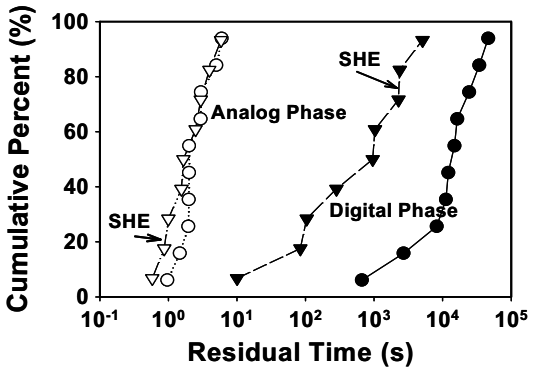

Fig. 11. Residual time distributions of the "digital" and "analog" phases of breakdown before and after substrate hot electron injection for stable filaments. In this case only the "digital" phase becomes shortened by the additional injection of carriers.

In a similar manner, SHEI was used during the second stress used to induce hard breakdown for filaments that were stable following the first breakdown. The results are shown in Fig. 11. In this case only the "digital" phase is affected by additional injection of electron current. This is expected since the length of the "analog" phase is not due to additional filament growth and probably is related to the behavior of the hard breakdown current due to the series resistance of the test structure as discussed in Section 2. The reduction of time observed for the "digital" phase is the SHEI acceleration of intrinsic oxide defect generation and wear-out, resulting in the increased probability of forming a second and hard breakdown filament. It has already been shown in several studies that SHEI accelerates intrinsic gate oxide wear-out [18].

The degree of "hardness" of the first breakdown was shown to affect the type of percolation filament formed and consequently the acceleration parameters for projecting the residual time. For softer breakdowns there is not enough power dissipation available during the breakdown transient to initiate dielectric-breakdown-induced epitaxy (DBIE) [19]. In this case a set of acceleration parameters different from those use to predict dielectric wear-out is required.

For higher levels power dissipation during the breakdown transient, the local temperature rise is sufficient to induce DBIE. The DBIE structure caps the current density due to negative feedback, and the filament remains stable. Other regions of the oxide continue to age until a second percolation path forms. Techniques such as the prevalence method or the successive breakdown method can be used to predict the formation of this second filament or subsequent filaments. In actual ULSI environments where the supply voltage is being scaled down and current available to the breakdown event is limited by series connected circuit elements, breakdown is expected to become softer. The results of this study suggest that evolution of non-stable breakdown filaments into hard breakdown will dominate the ultimate failure in circuits.

6. Conclusions

This study showed that two types of breakdown filament can be formed following soft breakdown in ultra-thin gate oxides. The type of filament depends on the amount of power available during the breakdown transient. Stable filaments are more likely to be formed when the power dissipation is large and the initiation of DBIE is favorable. The subsequent hard breakdown can be predicted by the prevalence method [9,10] or the successive breakdown theory [13] if multiple breakdowns can be tolerated in a circuit.

The acceleration factors are different for predicting the residual time to hard breakdown for non-stable breakdown filaments, indicating a physical mechanism different from oxide wear-out is involved in the evolution and formation of the final hard breakdown event.

Experiments using substrate hot-electron injection indicate that tunneling current appears to be the driving force for the progression of hard breakdown for both stable and unstable first breakdown filaments. This study also suggests that non-stable breakdown filament formation will most likely dominate in actual circuit applications where the power dissipation at breakdown is limited.

Acknowledgments

The author would like to thank the NIST Office of Microelectronics Programs for their support of this work.

References

[1] Kaczer B, Degraeve R, Groeseneken G, Rasras M, Kubicek S, Vandamme RE et al. Impact of MOSFET oxide breakdown on digital circuit operation and reliability. Tech Digest IEDM 2000:553.

[2] Rodriguez R, Stathis JH, Linder BP. Modeling and experimental verification of the effect of gate oxide breakdown on CMOS inverters. Proc IEEE International Rel Phy Symp, 2003. p. 11.

[3] Linder BP, Stathis JH, Frank DJ, Lombardo S, Vayshenker A. Growth and scaling of oxide conduction after breakdown. Proc IEEE International Rel Phy Symp, vol. 41, 2003. p. 402.

[4] Monsieur F, Vincent E, Roy D, Bruyere S, Vildeuil JC, Pananakakis G. A thorough investigation of progressive

breakdown in ultra-thin oxides, physical understanding and application for industrial assessment. Proc IEEE International Rel Phy Symp, vol. 40, 2002. p. 45.
[5] Monsieur F, Vincent E, Ribes G, Huard V, Bruyere S, Roy D. Evidence for defect-generation-driven wear-out of breakdown conduction path in ultra-thin oxides. Proc IEEE International Rel Phy Symp, vol. 41, 2003. p. 424.
[6] Suñé J, Miranda E, Nafria M, Aymerich X. Point contact conduction at the oxide breakdown of MOS devices. IEDM Tech Digest 1998:191–4.
[7] Suñé J, Miranda E. Postsoft breakdown conduction in SiO_2 gates oxides. Tech Digest IEDM 2000:533–6.
[8] Tung CH, Pey KL, Lin WH, Radhakrishnan MK. Polarity dependent dielectric breakdown-induced epitaxy (DBIE) in Si MOSFETs. IEEE Electron Dev Lett 2002;23(9): 526.
[9] Alam M, Smith RK, Weir B, Silverman P, Ma Y, Hwang D. The statistical distribution of percolation resistance as a probe into the mechanics of ultra-thin oxides breakdown. Tech Digest IEDM 2000:529.
[10] Suñé J, Wu E, Jiménez D, Vollersten RP, Miranda E. Understanding soft and hard breakdown statistics, prevalence ratios and energy dissipation during breakdown runaway. Tech Digest IEDM 2001:117.
[11] Suñé J, Wu E. Statistics of successive breakdown events for ultra-thin gate oxides. Tech Digest IEDM 2002:147.
[12] Alam M, Smith RK, Weir B, Silverman P. Statistically independent soft-breakdowns redefine oxide reliability specifications. IEDM Tech Digest 2002:151.
[13] Wu EY, Suñé J. Successive breakdown events and their relation with soft and hard breakdown modes. IEEE Electron Dev Lett 2003;24(11):692.
[14] Wu EY, Suñé J, Linder B, Stathis J, Lai W. Critical assessment of soft breakdown stability time and the implementation of new post-breakdown methodology for ultra-thin gate oxides. Tech Digest IEDM 2003:147.
[15] Sakura T, Utsunomiya H, Kamakura Y, Taniguchi K. A detailed study of soft and pre-soft breakdown in small geometry MOS structures. IEDM Tech Dig 1998:128.
[16] Degraeve R, Kaczer B, Keersgieter AD, Groeseneken G. Relation between breakdown mode and breakdown location in short channel NMOSFETs and its impact on reliability specifications. Proc International Reliability Physics Symposium, vol. 39, 2001. p. 360.
[17] Suehle JS, Zhu B, Chen Y, Bernstein JB. Acceleration factors and mechanistic study of progressive breakdown in small area ultra-thin gate oxides. Proc International Reliability Physics Symposium, vol. 42, 2004. p. 95.
[18] Vogel EM, Suehle JS, Edelstein MD, Wang B, Chen Y, Bernstein JB. Reliability of ultra-thin silicon dioxide under combined substrate hot-electron and constant voltage tunneling stress. IEEE Trans Electron Dev 2000;47: 1183–91.
[19] Tung C-H, Pey K-L, Tang L-J, Radhakrishnan MK, Lin W-H, Palumbo F et al. Percolation path and dielectric-breakdown-induced-epitaxy evolution during ultrathin gate dielectric breakdown transient. Appl Phys Lett 2003;83(11):2223.