NSF Final Report

Project #0901576

This work performed under this project can be broken down into three parts. First, we discuss the methodology, then the data collection and analysis, and finally the full chip simulation capability.

i. Parameter Extraction Methodology

Choice of Failure Rate and Spatial Distributions

It is common to describe dielectric breakdown with a Weibull distribution having two parameters,

$$P(t) = 1 - exp\left(-(t/\eta)^{\beta}\right) \quad (1)$$

where *P,* is the probability of failure, t denotes time-to-failure, η is the characteristic lifetime, and β is the shape parameter.

It is most common to assume a Poisson spatial distribution of defects. For the Poisson distribution, the probability of failure is

$$P = 1 - \exp(-\lambda(t)A) \quad (2)$$

where $\lambda(t)$ is the defect density, which is a function of time, and *A* is the area. The Poisson distribution assumes that defects are uniformly distributed throughout an area. However, clustering of defects has been observed often. If there is defect clustering, then the appropriate distribution is the Negative Binomial distribution, where

$$P = (1 + A\lambda(t)/\alpha)^{-\alpha} \quad (3)$$

where α is the clustering parameter, and where the negative binomial distribution approaches the Poisson distribution at α approaches infinity.

Model of Time Dependent Dielectric Breakdown

The unit of analysis for backend dielectric breakdown is the dielectric segment. Each dielectric segment is characterized by a vulnerable length, L, (the length of a block of dielectric that is stressed) and a linespace, S (the distance between lines to which a voltage can be applied). The total area of a dielectric segment is the product of the vulnerable length and linespace.

The characteristic lifetime of a dielectric segment with vulnerable length and linespace S is

$$\eta = Aexp(-\gamma E^{m} - E_a/kT) \quad (4)$$

where *A* is a technology-dependent constant, γ is the field extraction factor, *m* is one for the *E* model and ½ for the $\sqrt{E}$ model, E_a is the activation energy, *k* is the Boltzmann constant, and *T* is temperature. The electric field, *E=V/S,* is a function of the applied voltage, *V,* and the linespace. Equation (4) indicates the difference between a measured lifetime at high temperature and high voltage and the expected lifetime of a dielectric segment in a chip under use conditions. We have shown that scaling to use conditions is a function of linespace, S [1]-[4]. Hence, all features in a chip do not scale uniformly to use conditions. Equation (4) can also be used to estimate the lifetime of a test structure based on data from another test structure with a different linespace, provided that the vulnerable length is the same.

Area Scaling

Test structures are used to collect data. Area scaling allows multiple datasets with different areas to be merged to extract parameters so that the failure distribution parameters, η, β, and α, in equations (1) and (3) can be extracted with tighter confidence bounds.

Let's suppose that a reference test structure has a reference vulnerable length of L_t, with Weibull parameters η_t and β. Then, for each test structure with a different vulnerable length, L_i , (and same linespace) we

collect a set of failure times, ordered by probability point, to generate the set of pairs (t_i, P_i). If we assume a Poisson spatial distribution of defects, combining equations (1) and (2), for each vulnerable length we plot

$$ln\left(-\frac{L_t}{L_i}\ln(1-P_i)\right) = \beta(lnt_i - ln\eta_t) \quad (5)$$

where the x-axis is lnt_i and the y-axis is the left hand side of equation (5). If we assume a Negative Binomial spatial distribution of defects, combining equations (1) and (3), for each vulnerable length we plot

$$ln\left(\alpha\left(1+\frac{L_t}{L_i}\left((1-P_i)^{-1/\alpha}-1\right)\right)\right) = \beta(lnt_i - ln\eta_t) \quad (6)$$

where the x-axis is lnt_i and the y-axis is the left hand side of equation (6). Then we can merge multiple datasets with the same linespace, S, to extract η_t (the x-intercept) and β (the slope). However, note that there is an extra parameter, α, for the Negative Binomial distribution, which we determine by optimization to find the best fit [5].

We have extracted the parameters for a dataset with multiple areas. The extracted the best fit was $\alpha = 4$, which is a very low level of clustering [5]. Hence, we have assumed the Poisson spatial distribution model in the rest of our work in this project.

Let's suppose that Weibull parameters, η_t, and β have been determined by measuring data from a collection of test structures with vulnerable length, L_t and linespace, *S.* Suppose that a chip has a dielectric segment with vulnerable length, L_i, and linespace, S. Then, we have shown, using equation (5), that the characteristic lifetime of the dielectric segment, η_i, (for the same applied voltage and temperature) is [1]-[10]:

$$\eta_i = \eta_t(L_t/L_i)^{1/\beta}. \quad (7)$$

Parameter Extraction in the Presence of Die-to-Die Variation

Die-to-die variation has become significant. We have shown that in the presence of die-to-die variation, Weibull distributions are not linear when plotted on the Weibull scale, i.e. lnt vs. $\ln(-\ln(1-P))$ [1],[2],[11]-[13]. As a result, it is not possible to extract β by finding the slope of the Weibull plot. Instead, we use area scaling to extract β. Using equation (5) at the 67% probability point and using a reference test structure with vulnerable length, L_t, and characteristic lifetime, η_t, we plot the characteristic lifetime vs. the area ratio for all other test structures, i.e., $(ln\eta_i - ln\eta_t)$ vs. $ln(L_t/L_i)$, and extract the slope, which is β [1],[2],[11]-[13]. This method works because the characteristic lifetime is not sensitive to die-to-die variation. However, unlike the standard method, test structures with a variety of vulnerable lengths are required.

Parameter Extraction in the Presence of Multiple Features

A test structure can contain multiple features for which we aim to extract a failure rate separately. In particular, field enhancement occurs at the tips of combs or at the tips of lines that run in parallel. It is desirable to separately determine the failure rate for the vulnerable length and the failure rate at the tips due to field enhancement. We have developed a methodology to separate these effects [11]-[13].

We solve this problem by noting that at any time point, *t,* for all mechanisms, *i,* using equation (2),

$$\sum_i \lambda_i(t)A_i = -\ln(1-P). \quad (8)$$

A_i is known for each failure mechanism. It is the vulnerable length and/or the number of tips. The probability of failure, *P,* at each time point is also known from the data collected for each test structure. Then, if we have different test structures with different values of A_i for each mechanism, we solve for $\lambda_i(t)$. If we compute $\lambda_i(t)$ at each time point, we can reconstruct a Weibull curve for each failure mechanism separately by applying equation (2) to find the appropriate probability points.

In order to separate the impact of tips from vulnerable length, we have developed a variety to test structures with different vulnerable lengths and number of tips, i.e. the following vulnerable length, tip pairs (1X, 1X), (3X, 1X), (9X, 1X), (3X, 3X), (4,5X, 9X), and (9X, 9X). After analyzing the data, we determined that field enhancement at tips did not increase the failure rate [11]-[13].

Parameter Extraction for Multiple Failure Mechanisms

Some of the Weibull plots from our test structures appeared to be bimodal. The Weibull shape parameter in the upper percentile range is smaller than in the lower percentile regime. This indicates that there are two distinct failure modes. It is likely that failures in the upper percentile range are due to breakdown paths through the bulk dielectric, while failures in the lower percentiles may be due to the development of traps at the top passivating dielectric layer interface, where lattice mismatch creates a large number of dangling bonds. This bimodal failure rate distribution also indicates that the failure paths are associated with different defect generation rates.

We can directly extract the parameters for a bimodal failure rate, where p denotes the probability of failure due to the early failure mechanisms (along the interface). The overall probability of failure is

$$P = p\left(1 - exp\left(-(t/\eta_1)^{\beta_1}\right)\right) + (1-p)\left(1 - exp\left(-(t/\eta_2)^{\beta_2}\right)\right) \quad (9)$$

where η_1 and β_1 are the Weibull parameters of the early failure mechanism and η_2 and β_2 are the Weibull parameters of the bulk failure mechanism. The Weibull parameters cannot be extracted directly by finding the slope and x-intercept of the Weibull curve. They are found by maximum likelihood estimation. In fact, the composite Weibull failure rate distribution for two failure mechanisms will appear to have a very small Weibull shape parameter for the bulk failure rate, even if the underlying Weibull shape parameter for bulk failures is not small. This methodology has been used to analyze data for different linewidths [14].

ii. Data Collection and Analysis

Impact of Area

The impact of vulnerable length and the corresponding Weibull parameters were extracted based on test structures with four different areas: 1X, 3X, 4.5X, and 9X [1]-[3],[11]-[13],[15]. The parameter extraction methodology took into account die-to-die variation in the extraction of β [1],[2],[11]-[13].

Impact of Line Edge Roughness

Line edge roughness impact both the circuit and the test structures, and therefore is taken into account. Our samples show line edge roughness as expected. The standard deviation of the line edge roughness is similar to what is reported in the literature [14]. Line edge roughness changes the local linespace between interconnect lines and leads to a lower characteristic lifetime. We have approximated line edge roughness as a normal distribution and estimated the impact on characteristic lifetime. We found that its impact is much more significant for narrow linespaces [14].

Impact of Linewidth

Test structures with width ratio of 1X, 3X, and 5X have been implemented. Test results show a strong impact of linewidth, when the linespace remains constant [1],[2],[6],[14],[15]. Note that standard test structures for the impact of linewidth simultaneously vary density and linewidth. To distinguish between the density and linewidth effect, a test structure where the width of the lines for one comb is 1X and the other is 5X is included. The density of this test structure matches that of the 3X width test structure. In this way, we can vary linewidth independently of density.

It was found the linewidth, rather than density, determines the lifetime. Aspect-ratio-dependent-etching, where narrow trenches suffer from greater lateral etch near the critical chemical mechanical polishing (CMP) interface and a less vertical sidewall profile, can explain the trend [6],[14],[15].

Scanning Electron Microscopy (SEM) data were used to determine the difference between the actual linewidth and the drawn linewidth. This translates into a shift in linespace. A model was created for the shift in linespace by fitting the SEM data via regression [1],[2],[4],[6],[14].

The analysis of the linewidth data also took into account the fact that the Weibull distributions were bimodal by extracting the early failure and bulk failure rate parameters separately. This is because the wafers did not come from a mature process for 45nm technology. We found that both η and β for both the bulk and early failure populations are a function of the worst case aspect ratio (among the two lines bordering the dielectric segment) [14]. The probability, *p,* of early failure is best modeled as a function of the characteristic lifetime for bulk failures, which in turn is a function of linewidth [14]. In fact, the narrow lines, with higher electric fields and aspect ratios, have higher probabilities of early failures. The probability of early failures increases as the bulk characteristic lifetime degrades for narrow lines.

Impact of Irregular Features

Test structures to determine any impact of field enhancement because of irregular geometries have been designed. These test structures emphasize (a) the electric field between parallel routing tracks that end at the same point, (b) the electric field between line ends and perpendicular lines, (c) the electric fields between line ends in the same track that abut, and (d) the electric fields between line ends that terminate next to each other in adjacent tracks. All features were found to have a significant impact on lifetime [1],[2],[7],[8]. Hence, lifetime estimates need to take into account these features.

iii. Full Chip Simulator

Layout Extraction of Vulnerable Area and Irregular Features

We have developed a layout extraction tool using the standard object oriented programming language C++ [1]-[4],[9],[10]. We have used it to extract all of the dielectric segments and their features for a microprocessor system containing 310k nets which form around 31 million dielectric segments [1],[9].

Vulnerable length and features are extracted by comparing pairs of lines in a layout. Since tens of millions of lines exist in each metal layer in a layout, it is necessary to find the adjacent lines that border a dielectric segment or form a critical feature quickly. Our algorithm involves bucket sort to order the lines by x- and y-coordinates. Then, the first line in the first bucket is compared with the second line in the first bucket and the lines in the second bucket. The first line is deleted after extracting the critical features and vulnerable length, and the algorithm moves on to the next line.

Extraction of Activity

Because backend dielectric breakdown is activity and temperature dependent, our methodology includes determining the temperature and stress for each dielectric segment while running benchmarks [1],[9],[16]-[18]. For activity tracking, the hardware RTL/netlist is synthesized for emulation on an FPGA and counters are placed at the I/O ports, which track the state probabilities and toggle rates of the ports during application runtime. A standard set of benchmarks are used as the applications for analysis.

The I/O activities and the gate-level netlist are then used for activity propagation to each net in the design for a complete stress/transition probability profile of the internal nodes of a circuit under study. This provides the probability of a transition occurring at any node and the probability of each state, i.e., the probability at logic "1". It is this probability at logic "1" and logic "0" that is needed to compute the probability that each dielectric segment is under stress. The probabilities of dielectric stress of each dielectric segment is determined by

$$\alpha = \alpha_1(1-\alpha_2) + \alpha_2(1-\alpha_1) \tag{10}$$

where α_1 and α_2 are the probabilities that each net in the pair of nets that border the dielectric segment is at logic "1". If η_{dc} is the characteristic lifetime under dc stress, then the characteristic lifetime under ac stress, η_{ac}, is $\eta_{ac} = \eta_{dc}/\alpha$.

The netlist is also used for layout generation. The RC information from the layout, together with the net activity, is used for extraction of the power profile and the consequent thermal profile, through the power simulator and the thermal simulator.

We have tested our methodology using the well-known open-source LEON3 IP-core processor, where we have computed the average temperature distribution and the stress distribution [1],[9],[16]-[18].

Estimation of Lifetime and Probability of 1st Fail

Given a collection of *n* independent wearout mechanisms modeled with Weibull distributions, having parameters, $\eta_i, i = 1, \ldots, n$ and $\beta_i, i = 1, \ldots, n$, then the characteristic lifetime of the system, η_{chip}, i.e. the time when 63% of the population has failed from any mechanism, is the solution of [1]-[4],[6]-[10],[14],[16]-[18]

$$1 = \sum_{i=1}^{n} \left(\eta_{chip}/\eta_i\right)^{\beta_i}. \quad (11)$$

Similarly, the shape parameter of the system, β_{chip}, is [1],[2],[6]-[9],[14],[16]-[18]

$$\beta_{chip} = \sum_{i=1}^{n} \beta_i \left(\eta_{chip}/\eta_i\right)^{\beta_i}. \quad (12)$$

The components in equations (11) and (12) could be different wearout mechanisms, different layers of a chip, different geometries within geometries within a layer, or different geometries within a layer at different temperatures. Hence, a reliability simulator has to (a) determine the characteristic lifetime and shape parameters for all the underlying wearout mechanisms and geometries, after all components are scaled for temperature and to use conditions with equation (4), and (b) apply equations (11) and (12) to solve for η_{chip} and β_{chip}.

Note that if the dielectric segments fail based on a Weibull distribution, the system may not fail according to a Weibull distribution if the parameters of the underlying components vary widely. Equation (11) provides the lifetime of the system when 63% have failed, the Weibull characteristic lifetime. For an arbitrary probability of failure, *P,* the time-to-failure, *t,* is the solution of the following [1],[6]

$$-\ln(1-P) = \sum_{i=1}^{n} (t/\eta_i)^{\beta_i}. \quad (13)$$

We have also proposed equations to find identify the block or mechanism that will fail first in a large system [1],[18]. If β is constant, then the probability that the i^{th} unit fails first is

$$P_i = \left(\eta_{chip}/\eta_i\right)^{\beta}. \quad (14)$$

Otherwise, if β is not constant, then

$$P_i = \int_0^{\infty} e^{\alpha} du \quad (15)$$

where

$$\alpha = -u - \sum_{j \neq i} \left(\eta_i/\eta_j\right)^{\beta_j} u^{\beta_j/\beta_i}. \quad (16)$$

Several circuits have been analyzed to estimate lifetime and determine the critical features for lifetime. First, several versions of a radix-2, 256-point and 512-point FFT circuit was synthesized and implemented with the NCSU 45nm technology library [1],[3],[4],[6]-[8]. The 256-point circuit has 324k gates and 329k nets, and the 512-point circuit has 324k gates and 329k nets. The number of layers used in routing varied from five to eight. Using more routing layers results in shorter wirelength and better timing performance. Timing was optimized using buffer insertion and gate sizing. We also considered the LEON3 IP core processor, which includes a 32-bit general purpose integer unit, a 32-bit multiplier, a 32-bit divider, a memory management unit, a window-base register file, separate data and instruction caches, and cache tag storage units [1],[2],[9],16]-[18].

Our results show that (a) all linespaces need to be taken into account when computing lifetime, not just the minimum linespace [1]-[4], (b) increasing linewidth to improve feature-level lifetime does not improve the chip lifetime because the overall area of the chip increases [1]-[3],[6], (c) the impact of irregular features in significant

at the chip-level and most significant for metal 1 because it has the most irregular geometries and fewer routing restrictions [1],[2],[7],[8], (d) the irregular geometry with the greatest impact is when two lines in adjacent tracks terminate next to each other [1].[2].[8], (e) total wirelength in a layout correlates with lifetime [1]-[4], (f) no relation was found between lifetime and timing performance [1]-[4], (g) lifetime correlates more strongly with temperature than activity [1],[9],[16],[17], and (h) the probability that any of the units, except the two caches, fail first is negligible [1],[18].

Analysis of Redundancy

On-line reconfiguration through redundancy allocation involves detecting failing memory cells and reconfiguration of memory blocks using redundant spare rows and columns. In this way, memory blocks can tolerate a small number of faults.

If there is redundancy, we can solve for the lifetime in the presence of redundancy. Let $J_k, k = 1, \dots, K$ be disjoint sets of components. For example, J_i could consist of all dielectric segments in the metal 1 layer of the data cache. Any such set of components, J_k, can tolerate up to q_k defects. For example, if there are two redundant columns, then $q_k = 2$. Then the characteristic lifetime under redundancy, η_{chip}, is the solution of the following [1]:

$$1 = \sum_i \left(\frac{\eta_{chip}}{\eta_i}\right)^{\beta_i} - \sum_{k=1}^{K} ln\left(\sum_{x=0}^{q_k} \frac{\left(\sum_{i \in J_k} (\eta_{chip}/\eta_i)^{\beta_i}\right)^x}{x!}\right). \quad (17)$$

References

[1] M. Bashir, C.-C. Chen, L. Milor, D.H. Kim, and S.K. Lim, "Backend Dielectric Reliability Full Chip Simulator, *IEEE Trans. VLSI,* in press.

[2] M. Bashir, C.-C. Chen, L. Milor, D.H. Kim, and S.K. Lim, "Backend Dielectric Reliability Full Chip Simulator," *Proc. Design of Circuits and Integrated Systems,* 2012.

[3] M. Bashir, D.H. Kim, S.K. Lim, and L. Milor, "TDDB Chip Reliability in Copper Interconnects," *Proc. Integrated Reliability Workshop,* 2010, pp. 121-124.

[4] M. Bashir, D.H. Kim, K. Athikulwongse, S.K. Lim, and L. Milor, "Backend Low-K TDDB Chip Reliability Simulator," *Proc. Int. Reliability Physics Symp.,* 2011, pp. 65-74.

[5] L. Milor and C. Hong, "Area Scaling for Backend Dielectric Breakdown," *IEEE Trans. Semiconductor Manufacturing,* vol. 23, no. 3, pp. 429-441, Aug. 2010.

[6] M. Bashir, L. Milor, D.H. Kim, and S.K. Lim, "Methodology to Determine the Impact of Linewidth Variation on Chip Scale Copper/Low-K Backend Dielectric Breakdown," *Microelectronics and Reliability,* vol. 50, nos. 9-11, pp. 1341-1346, Sept./Nov. 2010.

[7] M. Bashir, L. Milor, D.H. Kim, and S.K. Lim, "Impact of Irregular Geometries on Low-K Dielectric Breakdown," *Microelectronics and Reliability,* vol. 51, nos. 9-11, pp. 1582-1586, Sept./Nov. 2011.

[8] C.-C. Chen, M. Bashir, L. Milor, D.H. Kim, and S.K. Lim, "Backend Dielectric Chip Reliability Simulator for Complex Interconnect Geometries," *Proc. Int. Reliability Physics Symp.,* 2012, pp. BD.4.1-BD.4.8.

[9] C.-C. Chen, F. Ahmed, D.H. Kim, S.K. Lim, and L. Milor, "Backend Dielectric Reliability Simulator for Microprocessor System," *Microelectronics and Reliability,* vol. 52, nos. 9-11, pp. 1953-1959, Sept./Oct. 2012.

[10] M. Bashir and L. Milor, "Towards a Chip Level Reliability Simulator for Copper/Low-K Backend Processes," *Proc. Design Automation and Test in Europe,* 2010, pp. 279-282.

[11] M. Bashir and L. Milor, "A Methodology to Extract Failure Rates of Low-K Dielectric Breakdown with Multiple Geometries and in the Presence of Die-to-Die Linewidth Variation," *Microelectronics and Reliability,* vol. 49, no. 9-11, pp. 1096-1102, 2009.

[12] M. Bashir and L. Milor, "Modeling Low-K Dielectric Breakdown to Determine Lifetime Requirements," *IEEE Design and Test of Computers,* vol. 26, no. 6, pp. 18026, Nov./Dec. 2009.
[13] M. Bashir and L. Milor, "Determination of the Impact of Field Enhancement in Low-K Dielectric Breakdown," *Proc. Int. Interconnect Technology Conf.,* 2009, pp. 104-106.
[14] M.M. Bashir and L. Milor, "Impact of Linewidth on Backend TDDB and Incorporation of the Linewidth Effect in Full Chip Lifetime Analysis," submitted to *IEEE Trans. Semiconductor Manufacturing.*
[15] M. Bashir and L. Milor, "Analysis of the Impact of Metal Linewidth on Low-K Dielectric Breakdown," *Proc. Int. Reliability Physics Symp.,* 2010, pp. 895-902.
[16] C.-C. Chen, F. Ahmed, and L. Milor, "A Comparative Study of Wearout Mechanisms in State-of-Art Microprocessors," *Proc. Int. Conf. on Computer Design,* 2012, pp. 271-276.
[17] C.-C. Chen and L. Milor, "System-Level Modeling and Microprocessor Reliability Analysis for Backend Wearout Mechanisms," *Proc. Design Automation and Test in Europe,* 2013, pp. 1615-1620.
[18] C.-C. Chen and L. Milor, "System-Level Modeling and Reliability Analysis of Microprocessor Systems," *Proc. Int. Workshop on Advances in Sensors and Interfaces,* 2013, pp. 178-183.