Circuit Reliability Simulator for Interconnect, Via, and Contact Electromigration

Boon-Khin Liew, Student Member, IEEE, Peng Fang, Nathan W. Cheung, Member, IEEE, and Chenming Hu, Fellow, IEEE

Abstract—A model for predicting Al interconnect and intermetallic contact/via electromigration time-to-failure under arbitrary current waveform is incorporated in a circuit electromigration reliability simulator. The simulator can 1) generate layout advisory for width and length of each interconnect, and the number of contacts and vias at each node in a circuit and 2) estimate the overall circuit electromigration failure rate and/or cumulative percent failure as functions of time, temperature, voltage, frequency, and previous stress (e.g., burn-in).

I. INTRODUCTION

IN THE EFFORT to assure reliable design in VLSI systems, a circuit reliability simulator that will provide layout guidelines and highlight potential reliability hazards in the circuit can be useful. Designers and layout engineers can use this tool to adjust the layout and/or circuit design so that the reliability requirement is met.

In our previous work [1], [2], we described and experimentally verified the vacancy recombination model for interconnect and intermetallic contact electromigration lifetime. The model is similar to the models proposed by Maiz [3] and Hatanaka et al. [4] but contains more features. Using the model, the electromigration median time-to-failure (MTF) under time-varying current conditions can be calculated for arbitrary waveforms. In this paper, we will describe a circuit electromigration reliability simulator which incorporates this model.

Electromigration reliability simulator for circuits is not a new idea. At Clemson University, the RELIANT simulator was developed to predict the failure rate of circuits due to electromigration [5]. At the University of Illinois and Texas Instruments, a program CREST was developed to compute the probability current waveform for a set of possible input waveforms for rapid computation of electromigration reliability [6]. The present work differs from the above by integrating electromigration simulation with a simulator that also simulates time-dependent oxide breakdown, and hot-carrier-induced degradation in both CMOS and bipolar transistors [7]–[10]. It also includes a more up-to-date model that explicitly considers the effect of arbitrary current waveforms and reduces to the simpler models in special cases.

II. OVERVIEW OF THE SIMULATOR

The circuit electromigration simulator (EM simulator) [7] is part of the Berkeley Reliability Tool (BERT) package which also contains hot-electron aging of MOS and bipolar circuits, and time-dependent oxide breakdown simulators [8]–[10] and is available upon request [11]. The EM simulator is organized in a pre- and post-processor configuration linked to SPICE, a general-purpose circuit simulation program. We have coded the EM simulator to be compatible with SPICE2 and SPICE3 analysis to allow circuit designers to perform functional verification concurrently with circuit electromigration simulation and the other reliability simulations.

To use the EM simulator, the circuit designer prepares the SPICE input deck the usual way or uses the layout extractor contained in the EM package to produce the SPICE input deck. Before invoking SPICE to perform circuit analysis, the EM simulator pre-processor is invoked. The pre-processor adds a dummy voltage source in the original SPICE input deck for every terminal of the circuit element and requests SPICE to output the current in all the dummy voltage sources (this is necessary because SPICE can only print current flowing in voltage sources). In addition, a database of the dummy voltage sources and their associated circuit elements are generated by the pre-processor. This information will be passed automatically to the post-processor. The modified input deck is then entered to the SPICE simulator.

Output from SPICE simulation is filtered by the post-processor which reads the current waveforms and performs the required electromigration calculations. At the same time, the post-processor will remove from the SPICE output the voltage sources and printout lines added by the pre-processor. In this manner, the operation of the simulator is transparent to the user.
Equation (1) which gives MTF for the general case is best understood in special cases:

1) Consider the dc current case, where \( \bar{J} = |\bar{J}| = J_{dc} \), the dc current density. The top bar represents average value and the double vertical bar represents absolute value. Equation (1) reduces to

\[
MTF_{dc} = \frac{A_{dc} \exp(E_a/KT)}{J_{dc}^m}.
\]  

Equation (2) is the usual Black's formula. Failure time is customarily fitted to lognormal distribution and MTF is the median time-to-failure. \( m \) is approximately equal to two and may be given somewhat different values for different ranges to \( J \). \( A_{dc} \) is used to fit the measured MTF and the temperature dependence is often expressed in terms of \( \exp(E_a/KT) \) where \( E_a \) is an activation energy, roughly equal to 0.5 eV for Al interconnect. All the parameters in (2) are routinely collected for IC metal systems and are the core input parameters to the simulator.

2) For the case of unidirectional current (\( \bar{J} = \bar{J} \)) and \( m = 2 \), (1) reduces to

\[
MTF_{dc} = \frac{A_{dc} \exp(E_a/KT)}{J^2}. 
\]  

Fig. 2 verifies (3).

3) If current is purely ac without a dc component such as the case in signal lines in CMOS circuits, \( \bar{J} = 0 \), (1) reduces to

\[
MTF_{ac} = \frac{A_{ac} \exp(E_a/KT)}{|\bar{J}|^m}. 
\]  

We allow the user to enter parameter \( A_{ac} \) to represent the fact that measured MTF_{ac} may be much larger than MTF_{dc} [1]-[4]. However, MTF_{ac} is not yet commonly measured. Default value of \( A_{ac} \) is \( A_{dc} \), in which case MTF_{ac} \approx MTF_{dc}.

We have shown that the model (1) is valid for interconnect as well as intermetal contacts/vias [1],[2] and believe that the model can be used for metal–semiconductor contacts as well. The necessary parameters in (1): \( m \), \( A_{dc} \), \( A_{ac} \), and \( E_a \) are experimentally determined constants. In addition, the coefficient \( \sigma \) of the lognormal time-to-failure distribution is determined from dc tests (\( \sigma \) is found to be independent of the current waveforms [1],[2]). The user can provide one set of all five parameters for each type of interconnect (up to three levels of interconnect are supported), contact, and via.

The dependence of MTF on interconnect length is treated as follows: the simulator uses the independent element model for length dependence described in [12],[13]. The user is requested to enter experimental time-to-failure data of a long interconnect, preferably over topographies that are representative of typical circuits. Failure statistics are calculated for shorter lines by the following assumption: long interconnect is equivalent to a series of shorter segments (see Fig. 3) and the failure probability of each segment is independent of each other. Therefore, the fail-
the length from the failure probability of the long interconnect \( F(t) \): at any time \( t \) the 2.25-cm line has a failure probability \( G_5(t) \) equal to (because the long interconnect will fail if either or both of the two segments fail)

\[
1 - G_5(t)^2 = [1 - F(t)]^2
\]

\[ (5) \]

\[
G_5(t) = 1 - [1 - F(t)]^{1/2}
\]

\[ (6) \]

the probability that the long interconnect will not fail is the product of the probability of the two shorter segments are good. For example, since \( F(7.5) = 50\% \), the failure probability for the 2.25-cm-long interconnect is \( G_5(7.5) = 0.29 \). In general, for an interconnect that is \( 1/x \) of the long interconnect line \( G_x(t) \) is

\[
G_x(t) = 1 - [1 - F(t)]^{1/x}.
\]

\[ (7) \]

The cumulative percent failure for shorter interconnect \( G_x(t) \) is calculated from the failure distribution of 4.5-cm-long interconnect and plotted in Fig. 4. The median-time-to-failure of interconnect as a function of length is indicated in the inset. The result shows MTF increases monotonically with decreasing length. Although several experimental studies [14], [15] have cast doubt on the absolute accuracy of the above model, no better model has been found for predicting the failure rate as a function of the line length. We intend to update the simulator as better models become available.

We have surveyed the present available literature and could not find a suitable theoretical model for the width dependence of interconnect MTF. Therefore, the simulator accepts measured MTF as a function of metal width provided by the user. The MTF data are represented by piecemeal fit of two second-order polynomial functions (Fig. 5), which would allow an increase in MTF as metal width is decreased below and above the average grain size [16], [17]. The four parameters that represent the width dependence of MTF are: \( A_w \), \( B_w \), \( C_w \), and \( D_w \). \( B_w \) is the linewidth at the minimum MTF versus width plot. It is approximately equal to the average grain size in the interconnect. If \( A_w \) and \( C_w \) are set to zero, MTF will be a constant independent of width.

**B. Statistical Models**

The user is given the choice of using either log-normal or Weibull distribution for electromigration lifetime in the simulation. The expression for lognormal distribution is

\[
F(t) = \frac{1}{\sqrt{2\pi} a t} \int_0^t \exp \left[ \frac{\log (u) - \log (MTF)^2}{2a^2} \right] du
\]

\[ (8) \]

with MTF provided by (1). The Weibull distribution function is

\[
F(t) = 1 - \exp \left[ -0.693 \frac{t^\beta}{MTF} \right]
\]

\[ (9) \]

where \( F(t) \) is the cumulative percent failure at time \( t \). In order to specify log-normal distribution, the median-time-
to-failure MTF and the coefficient $\sigma$ are entered. Both parameters are found by plotting the experimental time-to-failure data in a lognormal plot. Similarly, the coefficient $\beta$ (which is extracted from Weibull plot) is entered to specify Weibull distribution.

If log-normal distribution is chosen, the length dependence model adopted in the simulator will result in distributions for shorter interconnects that are not log-normal. This is illustrated in Fig. 3. The long interconnect time-to-failure data are represented by a log-normal distribution (a straight line in log-normal plot) and it is evident that the cumulative failure distribution of the shorter interconnects constructed using (7) and (8) are not log-normal. If Weibull distribution is chosen, the inconsistency is removed, i.e., failure distributions for all line lengths are Weibull.

C. Via and Contact Electromigration Model

Electromigration lifetimes of vias and contacts are modeled in a similar way as that of interconnect (i.e., (1)). This model assumes the use of barrier metal technology and does not consider failure due to Si migration at the Al to diffusion contact. The parameters in (1) are extracted from lifetime experiment using via or contact chain structures. The area of each via or contact opening in the test structure should be the same as the via or contact opening that user designs in the circuit layout. The area of all contacts/vias in a particular technology is usually fixed to guarantee uniform clearing of contacts/vias in the etch process. Current-crowding effect and step-coverage issue are already factored in the parameters $A_{dc}$ and $A_{ac}$ since the same via or contact size is used in the test structures and circuit layout. For the same reason, the simulator does not require a model to calculate the dependence of lifetime on via or contact size. In the layout advisory table, the simulator calculates the number of via or contact openings (each with the same area) needed for reliability. Current density is assumed to divide evenly among via or contact openings at a particular connection.

IV. User Input Parameters

The parameters in (1) and the coefficients for lifetime distribution in (8) and (9) of a long test line or chain for each level of metal, or via/contact chain provide the input regarding the technology, materials, and defects. One set of $m, A_{dc}, A_{ac}$ (optional), $E_a$, $\sigma$, or $\beta$, the length and width of the interconnect (from which the parameters are extracted) are needed for each layer of interconnect. The parameters are entered by the user in the EM design rule file which will be read by the post-processor. These parameters can be extracted from the MTF versus current density plot and the Arrhenius plot of MTF versus temperature. If MTF is a function of width (the simulator defaults to constant width dependence), the width parameters have to be extracted from additional experiments using long lines of different widths (see Fig. 4). Either log-normal or Weibull distribution is chosen to represent the time-to-failure distribution and the appropriate parameters are entered.

The parameter sets for contact and via are set up similarly. The test data for contact or via electromigration are obtained from contact/via chain test structures. User enters the $A_{dc}, A_{ac}, m, E_a$. The area at each contact/via opening in the test structure and the total number of contacts/vias in the chain are also entered.

If the simulator is requested to generate layout guidelines, the user is required to provide the reliability specification, for example, 0.1 FIT (0.01% failure per million device operating hours) at $10^4$ h (1.1 year) for each interconnect, contact, and via between different levels of conductors.
V. EXAMPLES OF SIMULATION

A. 21-Stage BiCMOS Inverter Chain

In this example, the simulator is used to generate layout guidelines and to calculate the failure rate of a 21-stage BiCMOS inverter chain. The circuit schematic of one inverter is shown in Fig. 5.

Because the circuit has 21 identical inverter cells, and only the current waveforms in one of the cells are needed, we use a subcircuit element for all but one of the inverter cells. This implicitly instructs the EM simulator not to perform failure calculation for the metal connections in the subcircuit element. As a result, this will reduce the amount of CPU time used in the EM and SPICE analyses. Note that because the power bus feeds to all 21 cells, the current density in the power line increases from $I \times J$ at the furthest cell to $21 \times J$ near the $V_{cc}$ source. The simulator can be flagged to treat this "stacked" connection, thereby eliminating the need for user to check each segment of the power bus.

The layout advisory for failure rate to be 1 FIT at $10^4$ h (1.1 year) for each interconnect and contact are generated. Table I lists the maximum length $L$ allowed for each interconnect in the circuit. The guidelines for the gate connections of MOS transistors and base connections of bipolar transistors are not listed because the simulator finds that they are the least likely to cause electromigration failures. The most constraining layout design from Table I is the $V_{cc}/$ground lines. Because the simulator treats the power line as one "stacked" connection, the length printed in the layout advisory is the maximum allowable length for the power line connected to one cell. Designer can use this table to choose the metal width which satisfies the length requirement dictated by routing and at the same time meets the reliability requirement.

The output in the layout advisory table can also be used to estimate the failure rate of each interconnect for any chosen length and width. The failure rate at $10^4$ h is approximately

$$FIT = \frac{\text{Interconnect Length}}{L \text{ from Table I}} \times 1 \text{ FIT}$$

1 FIT is used on the right-hand side of (10) because we have requested the layout advisory table to meet failure rate of 1 FIT at $10^4$ h in the present example.

Table II lists the layout advisory for metal to silicon contact. The simulator calculates the safety factor ($S$) for the number of contacts $N$ at each terminal (the size of each contact opening is taken to be fixed). The safety factor $S$ can be more easily interpreted as a measure of safety margin. For example, if a single 0.5-$\mu m^2$ contact is used at the emitter of $Q1$ (see Table II), the FIT is 1.9 times lower than 1 FIT, or 1800 times lower than 1 FIT if two contact openings are used. Thus a larger $S$ means more safety margin in meeting the reliability requirement. If a safety factor number of less than 1.0 is printed for a particular connection, the number of contact openings at that connection is not sufficient to meet the reliability specification and more contact openings are needed. For example, the designer might decide that a safety margin of 1.9 is not enough with one contact opening at the emitter of $Q1$, two contact openings can be used to increase the safety margin to 1800. This would be a wise decision if this same cell appears many times in a large circuit. Another interpretation of $S = 1800$ is that if 1800 identical cells are present, the total failure rate due to that emitter contact would be 1 FIT. The failure rate of a circuit is approximately equal to the sum of the FIT's of all the interconnects, contacts, and vias.

The circuit failure rate calculation described by the last sentence of the previous paragraph can be more easily and more accurately performed using the simulator in the failure rate prediction mode. The geometry information can be supplied to the EM simulator through a user-entered geometry description file (if the CIF layout is not available) or more conveniently by using the layout extractor.
Fig. 6. The cumulative percent failure and failure rate calculated for the 21-stage BiCMOS inverter chain.

Fig. 7. The failure rate for the 21-stage BiCMOS inverter chain increases as the input cycle time of the input waveform is reduced. As the MOS transistor switches on and off more often, the average current density in the interconnect and contact becomes larger and degrades electromigration reliability.

included in the simulator. Projected overall circuit failure rate and cumulative percent failure are plotted in Fig. 6. Note that the projected failure rate decreases after $10^4$ h. This is because the failure rate for a log-normal distribution can first increase, reach a maximum, and then decrease.

The circuit designer can use the simulator to perform tradeoff between circuit performance and reliability. One example of such tradeoff involves the circuit speed. In Fig. 7, the simulator is used to calculate the failure rate of the BiCMOS inverter chain as a function of the clock frequency of input signal. It is seen that by increasing the clock frequency, the circuit electromigration reliability degrades significantly (Fig. 7). The degradation for reliability with increasing clock frequency does not contradict the model prediction that MTF is independent of frequency for a given time-average current density $J$. In a CMOS circuit, because current flows only when the MOS transistor changes state, the average current density is larger when the MOS transistor switches on and off more often. Another design consideration that the simulator can be useful for is to evaluate the improvement in failure rate if supply voltage is reduced, for example from 6.5 to 3.5 V (Fig. 8).

B. CMOS Fulladder Circuit

In this example, the layout extractor is used to extract the dimensions of the MOS transistors, the connectivity, and geometry information of a CMOS fulladder circuit (Fig. 9). The SPICE input deck produced by the extractor is then combined with the input waveform, supply voltages, and the necessary commands for SPICE simulation. The EM simulator is invoked to calculate the failure rate of the circuit. In addition, the output from the simulator lists the worst 10 electromigration hazards in the circuit. This information can be superimposed on the original CIF file to identify the locations of electromigration trouble spots (see Fig. 9). Circuit designer can use this information to redesign the trouble spots in the circuit and improve reliability.

By using a command option in the layout extractor, one can scale down the geometry of interconnect, contact, and via in the CIF file and evaluate the effect of device downscaling on electromigration failure rate (Fig. 10). The precipitous rise in the failure rate as device dimensions become smaller is a major challenge to circuit scaling. The simulated results indicated in Fig. 10 are calculated assuming scaling in the lateral dimension only and have not taken into account scaling in the vertical dimension. There are two consequences which will further worsen the metallization reliability if vertical dimension is also scaled: by scaling the metal line thickness, the interconnect will carry larger current density, and by scaling gate oxide thickness, the current in the MOS transistor will increase. Because the total failure rate of the circuit is often dominated by the large failure rates of a few metal interconnects, contacts, or vias, the simulator can be useful in identifying such locations. The designer can then widen the interconnect, make more contact or via openings for the connections listed as worst hazards and therefore achieve optimal use of chip area.
Fig. 9. The layout for a CMOS fulladder circuit. The simulator highlights the worst trouble spots in the layout. Metal lines that pose reliability hazards are flagged and contacts are circled.

![Diagram of CMOS fulladder circuit]

Fig. 10. The effect of downsampling geometry on the failure rate of the CMOS fulladder circuit. The minimum interconnect width and contact area of the design are given in the inset.

![Graph showing effect of downsampling]

VI. SUMMARY

We have described the circuit electromigration simulator that incorporates a recent model of interconnect, via/contact electromigration for general current waveforms. The simulator can be used to provide layout advisory for circuit designer or to assess the failure rate/cumulative percent failure of a particular design. Examples have been presented on how the simulator can be used in a circuit design and technology development environment. The simulator described in this paper requires and allows considerable other user input. The width dependence of interconnect electromigration, and the statistical model (lognormal versus Weibull), for example, are left for the user to specify. If and when theoretical models become available, they can be added to the simulator easily.

REFERENCES


[11] Industrial Liaison Program Software Distribution, Department of Electrical Engineering and Computer Sciences, Cory Hall Rm 457, University of California, Berkeley, California 94720, Tel.: (510) 643-6687.


Boon-Khim Liew (S'90) was born in Kuala Lumpur, Malaysia, in 1963. He received the B.S. degree from California Institute of Technology, Pasadena, in 1985. He is currently working toward the Ph.D. degree in electrical engineering at the University of California, Berkeley. His doctoral research involves modeling and testing of electromigration and computer-aided reliability.

Peng Fang received the M.S.E.E. and Ph.D degrees in electrical engineering from the University of Minnesota, Minneapolis, in 1987 and 1988, respectively. He also received the B.S. degree from Sichuan University and the University of Electronics Science and Technology of China in 1982 and 1985, respectively.

He joined the Semiconductor Department of UESTC in 1982 and engaged in researches of MOS integrated circuits. From 1988 to 1989 he was a visiting Research Associate in the Micro-
electronics Center, the University of South Florida worked on WSI and VLSI interconnects. He is currently a postdoctoral researcher in Electronics Research Laboratory, University of California, Berkeley. His research area include noise in solid state devices, hot electron effects in MOSFET’s, electromigration, and stress-induced migration in VLSI interconnects.

Nathan W. Cheung (M'76–M'82) is currently Associate Professor in the Department of Electrical Engineering and Computer Sciences at the University of California-Berkeley and Associate Faculty of Lawrence Berkeley Laboratory. His active research area include integrated circuit processing, integrated circuit reliability, and thin-film processing technology. He had held research positions at Exxon Research Laboratory at Linden, NJ, Bell Telephone Laboratories at Murray Hill, NJ, and the National Research and Resource Facility for Submicron Structures at Cornell University, Ithaca, NY.


Chenming Hu (S'71–M'76–SM'83–F'89) received the B.S. degree from the National Taiwan University and the M.S. and Ph.D. degrees in electrical engineering from University of California, Berkeley, in 1970 and 1973, respectively.

From 1973 to 1976 he was an Assistant Professor at Massachusetts Institute of Technology, Cambridge. In 1976 he joined the University of California, Berkeley, as Professor of Electrical Engineering and Computer Sciences. While on industrial leave from the University in 1980-1981 he was manager of nonvolatile memory development at National Semiconductor. Since 1973 he has served as a consultant to the electronics industry. He has also been an advisor to many government and educational institutions. He conducted research on semiconductor devices and materials, electrooptics, solar cells, and power electronics. His present research areas include VLSI devices, hot-electron effects, thin dielectrics, CMOS latchup, nonvolatile semiconductor memories, power semiconductor devices, and GaAs devices. He has authored or coauthored one book and over 200 research papers. He is an Honorary Professor of Peking University, China. He was a National Science Council Invited Chair Lecturer, Republic of China, in 1987. He has been awarded several patents on semiconductor devices and technology.

Dr. Hu was Associate Editor of IEEE TRANSACTIONS ON ELECTRON DEVICES from 1986 to 1988 and Vice Chairman of IEEE Electron Devices Society, Santa Clara Valley Chapter, 1980–1982. He was a recipient of General Telephone and Electronics Fellowship and the University of California Science Fellowship. He has been listed in American Men & Women of Science, California Men & Women of Science and Technology, Who is Who in Technology.