Product Reliability Trends, Derating Considerations and Failure Mechanisms with Scaled CMOS

Mark White[1,2], Duc Vu[1], Duc Nguyen[1], Ron Ruiz[1] , Yuan Chen[1], Joseph B. Bernstein[2]
[1] Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA 91109
[2] University of Maryland, College Park, MD 20742
Phone: 818-393-4173 Email: Mark.White@jpl.nasa.gov

ABSTRACT

As microelectronics are scaled into the deep sub-micron regime, space and aerospace users of advanced technology CMOS are reassessing how scaling effects impact long-term product reliability. The effects of Electromigration (EM), Time-Dependent-Dielectric-Breakdown (TDDB) and Hot Carrier Degradation (HCI and NBTI) wearout mechanisms on scaled technologies and product reliability are investigated, accelerated stress testing across several technology nodes is performed, and FA is conducted to confirm the failure mechanism(s).

INTRODUCTION

Product level lifetime prediction of scaled CMOS technologies for space and aerospace applications is a principle focus for NASA and the military given the steady infusion of advanced commercial-off-the-shelf (COTS) microelectronics. These devices are being considered for use across a wide range of missions and operating environments. The space/aerospace community and other high reliability users of scaled CMOS microelectronics rely on derating for stress management, the intentional reduction of stress drivers, e.g. temperature, voltage, etc. to improve device reliability and extend operating life. NASA relies on derating principles to provide adequate safety margins between the intended operating environment and the maximum stress limitations of the device. To this end, the desire to assess the reliability of emerging scaled technologies through faster reliability trials, more accurate acceleration and predictive models, and model validation is the precursor for further research and experimentation in this field.

MODELING

While junction temperature (Tj) reduction has traditionally been the primary derating focus, various SRAM field studies of commercial devices, and experimental research and modeling of the effects of duty cycle and Vdd stresses on the device, suggest that derating these elements with Tj can provide an order of magnitude or more improvement in reliability (FIT) [1][2][3]. The circuit design and application, however, must be robust enough to operate at the lower end of the device performance and specification limits. Recently, Srinivasan [4] conducted Processor RAMP modeling which provided FIT estimates across 180nm to 65nm technologies for a processor operating at worst case conditions. FIT estimates for TDDB, EM, Stress Migration (SM) and Thermal Cycling (TC) related failure mechanisms, and their relative contribution to total FIT are summarized in Figure 1. On average, the simulated failure rate (FR) of a scaled 65nm processor may be as high as 316% higher than a similarly pipelined 180nm device [4].

Generally accepted models for MTTF due to EM, SM, TDDB and TC used in Srinivasan's model have been published in JEDEC Publication JEP122-A [5] and are recapitulated here for completeness:

$$t_{fEM} \propto (J)^{-n} \exp^{\frac{E_{aEM}}{kT}} \qquad (1)$$

where J is the current density in the interconnect, E_{aEM} is the activation energy for electromigration, k is Boltzmann's constant, and T is absolute temperature in Kelvin. n and E_{aEM} are constants that depend on the interconnect metal used.

$$t_{fSM} \propto \left|T_o - T\right|^{-m} \exp^{\frac{E_{aSM}}{kT}} \qquad (2)$$

where T is the absolute temperature in Kelvin, T_o is the stress free temperature of the metal (the metal deposition temperature), and m and E_{aSM} are material dependent constants.

$$t_{fTDDB} \propto \left(\frac{1}{V}\right)^{a-bT} \exp^{\frac{\left(X+\frac{Y}{T}+ZT\right)}{kT}} \qquad (3)$$

where T is the absolute temperature in Kelvin, a,b,X,Y, and Z are fitting parameters, and V is the voltage.

$$t_{fTC} \propto \left(\frac{1}{T_{average} - T_{ambient}}\right)^{q} \qquad (4)$$

where $T_{ambient}$ is the ambient temperature in Kelvin, $T_{average} - T_{ambient}$ is the average large thermal cycle a structure on chip experiences, and q is the Coffin-Manson exponent, an empirically determined material-dependent constant.

More recent work conducted at University of Maryland (UMD), Center for Microelectronics Reliability Engineering using SRAM SPICE modeling and analysis offers an improvement over earlier models. UMD simulated a 17-stage ring oscillator consisting of CMOS inverters and interconnecting capacitors to investigate the impact of voltage, frequency and temperature stress conditions on HCI, TDDB and NBTI on a scaled device. See Figure 2. The primary effects of HCI on device characteristics are threshold voltage drifting and transconductance degradation; increases in cell access time may result after long term operation. TDDB in Store-to-Vdd and Store-to-gnd result in increased leakage currents and a degradation of cell stability and static noise margin (SNM). NBTI leads to device mismatches in the SRAM cell and input offset voltages in the sense amplifier; SNM degrades as Vdd decreases. Taking a physics-of-failure approach with SRAM SPICE modeling to estimate MTTF of each failure

1-4244-0297-2/06/$20.00 ©2006 IEEE

Authorized licensed use limited to: Jet Propulsion Laboratory. Downloaded on February 10, 2009 at 16:55 from IEEE Xplore. Restrictions apply.

mechanism [2][3][6]. The SPICE simulation was based on the following device lifetime and derating models proposed by Walters [2][7]:

$$t_f = A_{HCI}(\frac{I_{sub}}{W})^{-n}\exp(\frac{E_{aHCI}}{kT}) \quad (5)$$

$$t_f = A_{TDDB}(\frac{1}{A})^{\frac{1}{\beta}}F^{\frac{1}{\beta}}V_{gs}^{a+bT}\exp(\frac{c}{T}+\frac{d}{T^2}) \quad (6)$$

$$t_f = A_{NBTI}V_{gs}^{-\frac{1}{\beta}}[\frac{1}{1+2\exp(-\frac{E_1}{kT})}+\frac{1}{1+2\exp(-\frac{E_2}{kT})}]^{-\frac{1}{\beta}} \quad (7)$$

Derating factor (D*f*) is defined as the ratio of the MTTF of a device operating at derated conditions (*MTTFd*) to its MTTF at rated operation condition ($MTTF_0$):

$$D_f = \frac{MTTF_d}{MTTF_o} \text{ and } MTTF_{derated} = Df(MTTF_{rated}) \quad (8)$$

Total derating factor is sought after to provide adequate safety margin against each of the common wearout failure mechanisms: EM, HCD and TDDB. Total D*f* thus becomes a function of the individual derating factors D_{fEM}, D_{fHCD} and D_{fTDDB}:

$$D_{fEM} = \left(\frac{V_{DD}^o}{V_{DD}}\right)^n \exp\left(\frac{E_{aEM}}{k}\left(\frac{1}{T_j}-\frac{1}{T_j^o}\right)\right) \quad (9)$$

$$D_{fHCD} = \exp\left(\theta\left(\frac{1}{V_{DD}}-\frac{1}{V_{DD}^O}\right)\right) \quad (10)$$

$$D_{fTDDB} = \exp\left(\gamma E_{ox}\left(1-\frac{V_{DD}}{V_{DD}^O}\right)\right)\exp\left(\frac{E_{aTDDB}}{k}\left(\frac{1}{T_j}-\frac{1}{T_j^O}\right)\right) \quad (11)$$

where Ea_{EM} and Ea_{TDDB} are the activation energy, Θ and γ are empirical constants, *Eox* is the oxide field, V^0_{DD} and T^0_j denote rated operating values for voltage and temperature, and V_{DD} and T_j represent derated values for voltage and temperature. The above three wearout mechanisms are related to the total derating factor *Df* with function *fd* :

$$D_f = f_d\left(D_{fEM}, D_{fHCD}, D_{fTDDB}\right) \quad (12)$$

If independency with respect to each is assumed for D_{fEM}, D_{fHCD} and D_{fTDDB} , Li [3] hypothesizes *fd* can be approximated with the linear relation:

$$f_d = C_{EM}D_{fEM} + C_{HCD}D_{fHCD} + C_{TDDB}D_{fTDDB} \quad (13)$$

where C_{EM}, C_{HCD} and C_{TDDB} are constants and their values to be determined from experiment or simulation. The total derating factor *Df* becomes unity when there is no derating for each failure mechanism.

Voltage and temperature are the two principle stress drivers for the failure mechanisms described earlier and the combined effect on circuit/product level reliability becomes very complex. Further experimentation and analysis with varying voltage and temperature stress combinations is necessary to stimulate and accelerate the wearout failure mechanisms; this will lead to a better understanding and validation of these models.

EXPERIMENTAL RESULTS

SRAM devices consisting of three mature technologies (0.13, 0.15 and 0.25um) were subjected to high temperature range of 125°C to 165°C and voltage ramp stress testing in 0.1Vdd and 10°C steps at constant frequency [8]. Objectives included validation of the manufacturer's FIT, identification of dominant failure mechanism(s) at the circuit/product level, and validation of the time-to-fail as a result of the dominant failure mechanism. First, temperature was fixed and voltage was stepped up 0.1Vdd, and secondly voltage was fixed and temperature was increased 10°C.

The experiment revealed that at specific times and voltage/temperature thresholds, large numbers of bit failures were recorded. The failures that were recorded at the same time represent a single failure event which was reflected on multiple addresses and therefore, counted as a single failure for reliability evaluation. Hard and soft failures were treated equally in the reliability evaluation because once a soft failure has occurred in a high-reliability, remote application, e.g. an un-repairable system, the address corresponding to the failure are circumvented and not used in future write cycles. Table 1 shows test conditions and time to 0.1% device-bit failures of the three technology nodes tested. The failures were exponentially distributed and the sample size included four devices/technology as a result of test capability constraints.

In this experiment, the manufacturer's FIT was validated with a weighted sum Exponential Model (14), normalized to 55°C and nominal Vdd. The weighted sum Exponential Model best correlated the manufacturer's published data (7-20 FIT) to the experimental data (19.482 FIT), normalized to 55°C and nominal Vdd operating conditions. Reference Table 2.

$$AF = \frac{\lambda(T_2,V_2)}{\lambda(T_1,V_1)} = (AF_T + AF_V)/2$$
$$= (\exp\left(\frac{E_a}{k}\left(\frac{1}{T_1}-\frac{1}{T_2}\right)\right) + \exp\left(\gamma_1(V_2-V_1)\right))/2 \quad (14)$$

FAILURE ANALYSIS

Upon functional failure, units were submitted for failure analysis. I-V curve measurements using a Digital Curve Tektronix 370 tracer revealed a 120 ohm resistive short in the input circuitry between Vcc and Vss. See Figure 3. Devices were then chemically decapsulated and subjected to internal optical examination. See Figure 4. Photon emission microscopy (EMMI) was implemented to pinpoint the failure site. Optical testing of advanced CMOS circuits exploits the near-infrared photon emission by hot-carriers in transistor channels. However, due to the continuous scaling of features size and supply voltage, spontaneous emission is becoming fainter and optical circuit diagnostics becomes ever more challenging [9]. EMMI revealed emissions in the area between the Vcc and Vss buses. The EMMI findings correspond with the I-V curve measurements. A device was subjected to Focused Ion Beam FIB/SEM inspection to determine root cause of the failure.

Authorized licensed use limited to: Jet Propulsion Laboratory. Downloaded on February 10, 2009 at 16:55 from IEEE Xplore. Restrictions apply.

See Figures 5 and 6. The differences in metal appearance in the upper and lower portions of the image reflect differences in stress conditions. Stress induced metal migration is evident in the lower region. See Figure 8.

DISCUSSION

Devices in this accelerated stress test experiment succumbed to thermal runaway upon reaching critical temperature and voltage thresholds. Ref. Table 1. The failures were caused by ESD and EOS stresses applied directly between Vcc and Vss pins (Input circuitry transistors). No damage was found in the memory cells as a result of EM, TDDB or HCD. FIB/SEM inspections show evidence of an ESD event and thermal/electrical stress induced metal migration damage. Buffer/voltage regulation circuitry protected the actual memory cells as damage was concentrated on the input transistors of the memories. The observed failure mode of a sudden large increase in memory cells (bit failures) was actually a result of the failure of the input circuitry (Low resistance 120-Ohm short between Vcc and Vss) of the device. These results demonstrate the necessity to conduct FA on accelerated stress test failures to confirm the actual failure mechanism(s) and as in this case study, avoid the false conclusion of memory bank failure.

SUMMARY

The stress test and failure analysis on 0.25, 0.15 and 0.13 SRAM technologies demonstrate that root cause of failure can be attributable to multiple, simultaneous failure mechanisms. Furthermore, it is not practical to assume no interdependency of the effect of voltage and temperature stresses on the wearout failure mechanisms. Different failure mechanisms will also be accelerated by certain voltage and temperature stress combinations. In conclusion, additional experiments are needed to refine and validate the models described earlier. Future work includes accelerated stress testing and modeling with bulk 110, 90 and 65nm CMOS technologies using frequency, temperature and voltage as the principal stress variables. The objective is to establish time-to-fail at the product level from either the dominant or multiple failure mechanism(s), and to establish appropriate product level derating criteria as a function of technology scaling.

REFERENCES

[1] Bornstein, W, et al, “Field Degradation of Memory Components from Hot Carriers,” IEEE IRPS (2006)

[2] University of Maryland - Center for Reliability Engineering, AVSI 17 Project (2002-2006)

[3] Li, X, et al, “Simulating and Improving Microelectronic Device Reliability by Scaling Voltage and Temperature,” IEEE ISQED (2005)

[4] Srinivasan, et al,"The Impact of Technology Scaling on Lifetime Reliability." International Conference on Dependable Systems and Networks. June 2004.

[5] Li, X, et al, “SRAM Circuit Failure Modeling and Reliability Simulation with SPICE,” IEEE Transactions on Device and Materials Reliability (2006)

[6] Failure Mechanisms and Models for Semiconductor Devices. *JEDEC Publication JEP122-A* (2002)

[7] Walters, J.D., “Methods to Account for Accelerated Semiconductor Device Wearout in Long life Aerospace Applications,” PhD Dissertation, University of Maryland (2003).

[8] White, M, et al., “Impact of Device Scaling on Deep Sub-Micron Transistor Reliability – A Study of Reliability Trends Using SRAM,” IEEE IRW (2005).

[9] Tossi, A, et al., “Hot-Carrier Photo emission in Scaled CMOS Technologies: A Challenge for Emission Based Testing and Diagnostics,” IEEE IRPS (2006).

Acknowledgements

The research described in this paper was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.

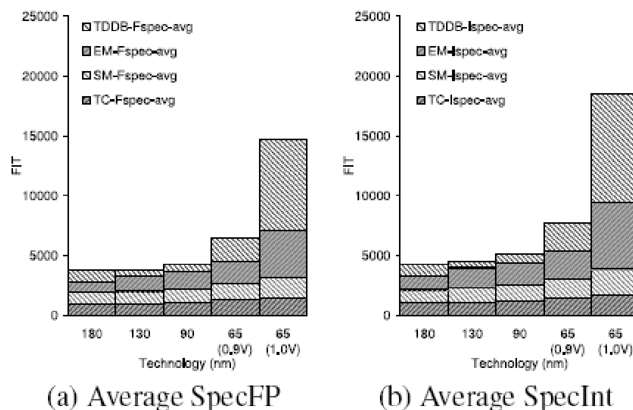

Figure 1. FIT values for processor averaged for w/c conditions application for model (a) and model (b) with relative contribution of each mechanism [4].

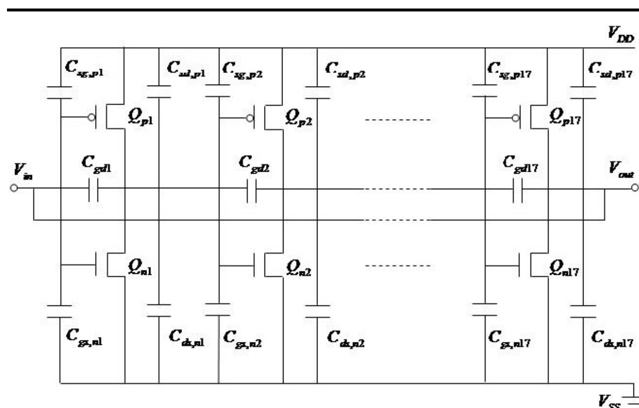

Figure 2. Schematic diagram of the ring oscillator used as a test circuit for monitoring process variations. Its oscillation frequency is sensitive to SPICE model parameters [3].

Table 1. Technology node and stress conditions vs. time to failure of 0.1% of the bits in a device.

Tech. Node	Vratio (Vapp/Vnom)	Temp C	Time to 0.1% Device-Bit Failures (Hrs)
0.13	1.4	165/155	588
	1.5		
0.15	1.6	165	528
0.25	1.7	165	768

Authorized licensed use limited to: Jet Propulsion Laboratory. Downloaded on February 10, 2009 at 16:55 from IEEE Xplore. Restrictions apply.

Table 2. Step-stress accelerated test results compared to manufacturer's data.

Test level	Cumulated test time	Equivalent op. time @55deg&nominal voltage			
		Case1 (Multiplication)		Case2 (Weighted Sum)	
		AFv Exp. Model (1)	AFv Power law (2)	AFv Exp. Model (1)	AFv Power law (2)
stress level 1	576	32464923.04	237589693.1	310353.6276	2170970.594
stress level 2	384	43090951.76	315354698.1	217390.3382	1457801.649
stress level 3	384	434116546.9	3918127282	1998870.897	17871738.22
stress level 4	384	824942335.4	7445532987	2017841.11	17890708.43
stress level 5	384	8310819403	77740152267	19965232.78	186422071.3
stress level 6	384	12452806266	1.16485E+11	19985188.96	186442027.5
stress level 7	335.8	1.09721E+11	9.14211E+11	175611815.3	1462841979
stress level 8	133.6	4.39858E+11	2.85782E+12	703819229.5	4572690225
Total equiv. time:		5.71677E+11	3.97817E+12	923925922.4	6447787521
Failure rate @55C &Vnom (FIT)		**0.031**	**0.004**	**19.482**	**2.792**
Failure rate reported by Manuf: 7 – 20 FIT					

Case 1 – refers to assumption a.
Case 2 – refers to assumption b.
(1) - Voltage Acceleration Factor according to Exponential. Model ($\gamma = 7$)
(2) - Voltage Acceleration Factor according to Power Law Model (k=34)
(3) – Mfr's FIT reported at 60% CL. ALT comparison also at 60% CL.

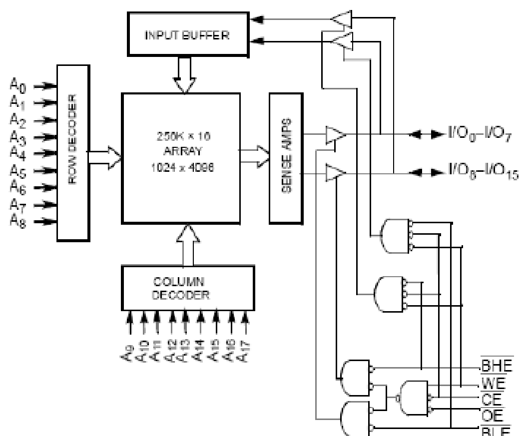

Figure 3. 256K X 16 Static RAM functional diagram

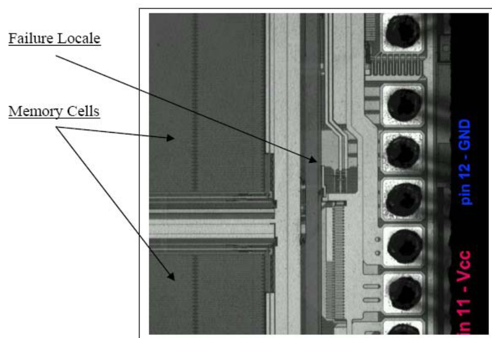

Figure 4. Decapsulated Optical Overview

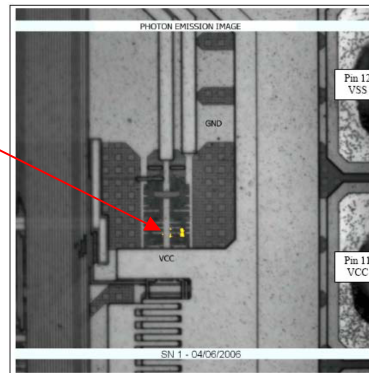

Figure 5. Photon emission image shows emissions between pin 11 (Vcc) and pin 12 (Vss)

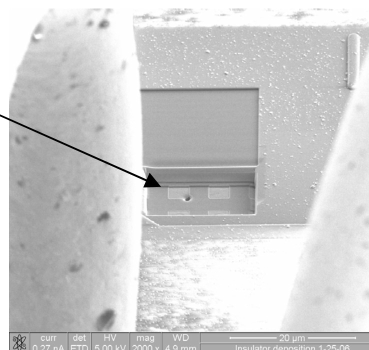

Figure 6. Close-up of the defective region milled with the FIB instrument directly over the area that produced photons in the emission microscope.

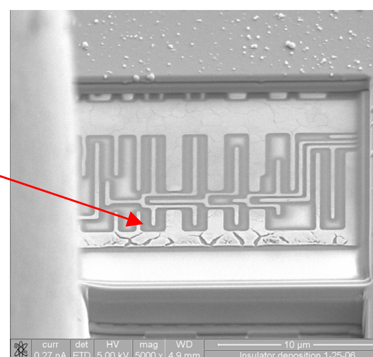

Figure 7. SEM of the defective Region milled with the FIB instrument.

Authorized licensed use limited to: Jet Propulsion Laboratory. Downloaded on February 10, 2009 at 16:55 from IEEE Xplore. Restrictions apply.