Prognostic Techniques for Semiconductor Failure Modes

Douglas L. Goodman
Ridgetop Group, Inc.
7070 North Oracle Road, Suite 120
Tucson, Arizona 85704
(520) 742-3300

Abstract

Semiconductor reliability issues are beginning to emerge as a major impediment to long term reliability of critical systems such as Internet routers, ATM machines, and Automotive/Aerospace fly-by-wire systems. Semiconductors have certain defined failure modes that can contribute to end-of-life failures. These modes include time-dependent dielectric breakdown of the gate oxide (TDDB), hot carrier damage, and metal migration. All of these common failure modes are far worse at geometries below 0.25μ. Fortunately, there are methods proposed that counteract these common failure modes. This paper surveys the problems involved, and recommends a methodology for the inclusion of pre-calibrated prognostic cells that can be co-located with a host circuit to provide an "early-warning" of a system failure, so that appropriate corrective action can be taken.

Introduction

It is well understood that semiconductors have smaller and smaller process geometries and that ever-increasing levels of functionality are being integrated into new System on Chips (SoC's). Among the issues facing the designers of the SoC's are the proper scaling of the pre-designed Intellectual Property (IP) blocks as the die shrinks occur, plus assuring that the reliability of the SoC will be maintained[1].

There are age-related failure modes in semiconductors that limit the long term performance of the semiconductors, impacting the ability of designers to provide robust SoC's for demanding applications. These failure phenomena include oxide breakdown, hot carrier damage and metal migration. With the prognostic methodology proposed, the proper inclusion of pre-designed, pre-calibrated prognostic cells, fail-safe performance for demanding applications can be accommodated.

Further, due to the methodology in retrieving the information, and being at the IC level, the technique also provides the highest degree of diagnostic resolution for a system. This provides significant benefits to system-level fault diagnosis, as well as for provisioning of spares to maintain a high system uptime objective.

[1] - Seshan, K, et al, "The Quality and Reliability of Intel's Quarter Micron Process", Intel Technology Journal, Q3 1998

© 2000, Ridgetop Group, Inc. Tucson, Arizona

Failure Modes in Semiconductors

There are three common failure modes considered in this paper, gate oxide failure, hot carrier damage, and metal migration. All are exacerbated by shrinking process geometries and some are not fully understood. Each will be described in turn.

Gate Oxide Failure

This failure results from time-dependent dielectric breakdown (TDDB). In this case, the dielectric fails when a conductive path forms in the dielectric, shorting the anode and the cathode. There are some models for this phenomena[2] that have been released but there is still more work to be done to handle hi K oxides (e.g. nitrided oxides) and oxide layers thinner than 4 nm. At these extremely thin oxide layers, the dielectrics exhibit high tunneling currents, and the impact on device reliability is not well understood[3]. NIST is doing some work with Academic and Industrial Research groups to characterize the expected worsening reliability issues with this process-size related problem.

In the case of TDDB, adding protective circuits is not effective because the breakdown occurs over a period of time, even if the device is operated below the typical oxide breakdown voltage. Degradation is greatly accelerated by an increase in the electric field

Hot Carrier Damage

Hot carrier damage occurs as carriers move along the channel of a MOSFET and experience impact ionization near the drain end of the device. It is caused by gaining energy from the intense electric field of small process geometries. As the carriers gain enough energy, some surmount the Si-SiO2 energy barrier and are injected into the gate oxide film. The damage can occur at the interface, within the oxide or within the sidewall spacer. Charge trapping induced by this mechanism results in shifts in threshold voltage V_{th}, which can degrade high frequency switching and create mismatches in balanced signal path stages. Most semiconductor failure modes exhibit higher degradation rates at higher temperatures. The hot carrier effect, however, accelerates as the temperature decreases.[4]

Metal Migration

Another reliability concern is electromigration, which result from increased current densities. The current generation of highly integrated microprocessors, requiring dense

---

[2] - "Semiconductor Device Reliability Failure Models" International Semitech, Technology Transfer #00053955A-XFR

[3] Suehle, John, Dielectric and Interconnect Reliability Metrology Summary, National Institute of Standards and Technology (NIST), November, 2000 www.eeel.nist.gov

[4] "Failure Modes in Semiconductors", Toshiba Corporation Research Report, October 2000

interconnects and large amounts of current, has highlighted the concern for metal interconnect reliability.

Formation of metal voids induced by electromigration during normal SoC or IC operation will cause an interconnect open or high resistance resulting in malfunction or speed degradation. For its microprocessors, Intel has developed some mitigation techniques, involving the use of up to five metal layers, but the electromigration effects and resulting reliability concerns will still be present to varying degrees.

Prognostic Monitoring Concept

Prognostic monitors employ accelerated and calibrated stress conditions to increase their rate of degradation relative to the companion functional circuits in which they are incorporated, thereby assuring the monitor will fail before the circuit. Incorporating the monitors as part of sub-system circuits insures that they will see the same operational environment as the system components from fabrication, to test, to burn-in, to operation. This assures that any variation that would affect the system reliability (e.g. process-induced or installation-induced damage, voltage transients or spikes, temperature variations, etc.) will also affect the monitor causing its premature failure relative to that of the circuit. Thus, traditional methods of lifetime prediction based on conditions experienced by the integrated circuit up to system insertion and offline testing of other components is replaced by a technique that takes into consideration any impact the actual operating environment may have on system lifetime.

The prognostic monitor approach can be more readily appreciated by referring to the typical failure rate curve, shown in figure 1, that is typical of electronic components, circuits and systems (and for that matter most common items in general).

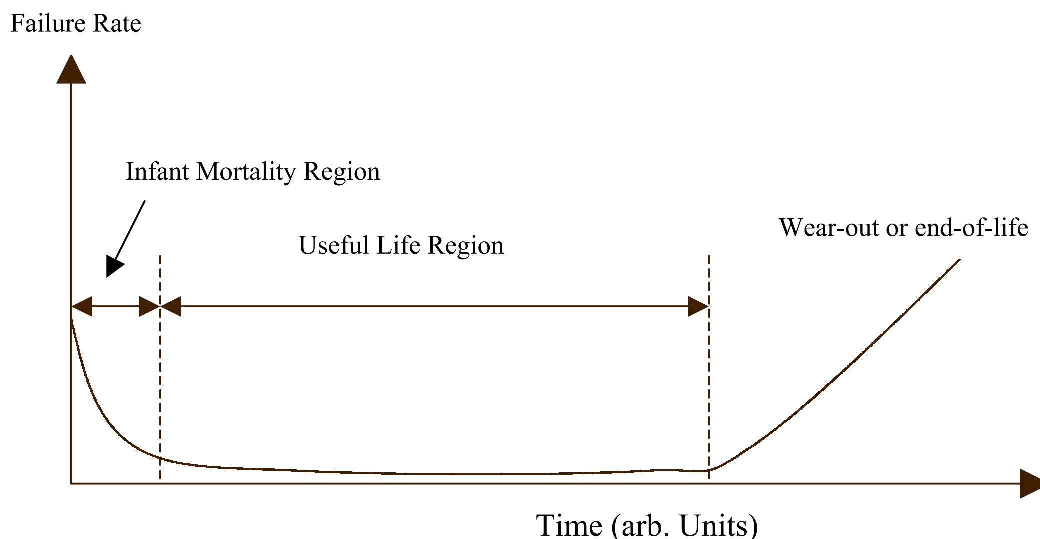

Figure 1 - Classical "Bathtub" Reliability Curve

© 2000, Ridgetop Group, Inc. Tucson, Arizona

Initially the failure rate is high but decreases markedly with time. This time period is known as early failure or infant mortality. Subsequently, the failure rate drops to a low and relatively constant value for a substantially long period of time known as the useful life. Ultimately, the components receive enough stress through continuous or intermittent operation that their failure rate increases with further operation time. This is referred to as end-of-life, or the wear out period.

During the infant mortality stage most failures are due to manufacturing defects. Thus, most circuits, especially where high reliability is a must, are subjected to stress conditions sufficient to simulate the early failure period in a test period known as burn-in. This insures that a circuit at system insertion has moved to the useful life period of the bathtub curve.

The useful life for the circuit is then estimated by off-line testing of other components/circuits that hopefully represent an average expected performance of all devices. But of course, these measurements have no means to include effects of the operational environment seen by circuits in use. For that matter, in general they do not see the precise fabrication environment of the circuits in use, due to the required duration of lifetime testing. This can be of major importance due to process variations that could prove to be catastrophic for a given circuit grouping, and not present in circuits under lifetime test.

Prognostic monitors circumvent the foregoing limitations in critical circuits by insuring that the test cell sees the exact environment that the actual circuit sees, but at an accelerated rate, thereby providing failure prediction. Thus, prognostic cells can detect the integrated effect of factors that have the potential to severely change the bathtub curve. The failure point of the prognostic cell is configured to lie just to the left of the wear-out region, as depicted in figure 2.

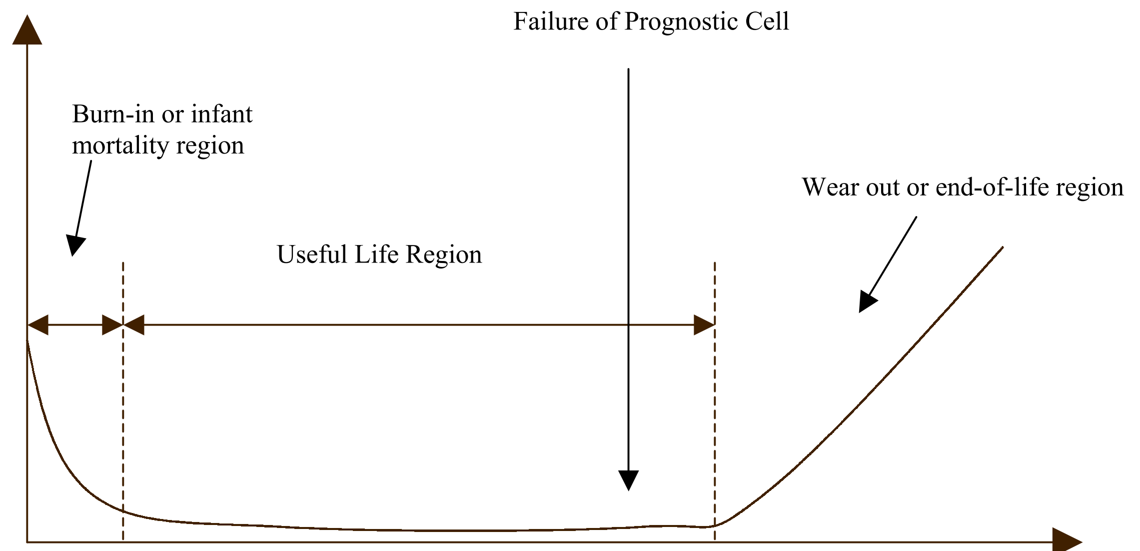

Figure 2 – Bathtub curve with indication of the Prognostic Cell failure specification

The purpose of the prognostic cell is to predict circuit failure. To achieve this, the prognostic cell must fail prior to the circuit on the same chip for all realistic operating conditions. The prognostic cell failure distribution of a particular chip must therefore be linked to the fabrication and operating conditions in the same way as the circuit failure distribution of that same chip. If the complete prognostic cell failure distribution lies before the onset of the circuit failure distribution, all prognostic cells will fail prior to circuit failure (Figure 3). The challenge to building a useful prognostic cell is to ensure this type of predictive capability without severely reducing useful circuit lifetime.

Methodology for Addressing Common Age-related Failure Modes

Building on earlier work conducted by the US Air Force[5], Ridgetop Group has developed and refined the InstaCell™ library of prognostic cells that have the characteristic of providing an early warning indication of a future end-of-life semiconductor failure. These cells are compact and self-contained, and can be co-located with a host circuit. These cells are configured with well-characterized acceleration mechanisms to elicit an early-warning of the host circuit failure. The cells are accessed via a JTAG bus (IEEE1149.1) to provide a means of accessing the information quickly and efficiently. This is depicted in figure 3:

Cell Libraries for each process (patent pending)

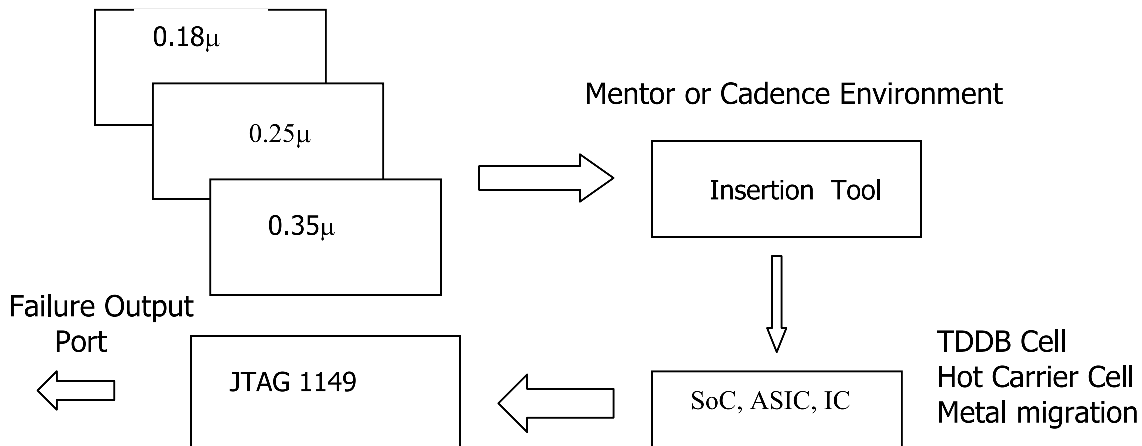

Figure 3 – InstaCell™ Architecture : Prognostics for semiconductor Failures

By carefully studying the statistics associated with the fault modes, Ridgetop has developed methods of calibrating the "early warning" period that cells can provide.

System Diagnostics

[5] Wilmar W. Sifre, Steven L. Drager and Martin J. Walter "A Prognostic Monitor for Microelectronics Systems Reliability", GOMAC Conference, 1996

By providing fault indications at the IC level, the highest degree of diagnostic resolution can be provided in a system. Ridgetop has designed the system so that the predicted failure indication is accessible via IEEE1149.1/JTAG. This is a well accepted process that employs a boundary scan technique. Using the IEEE1149.1, programming standards such as boundary scan description language (BSDL), hierarchical scan description language (HSDL), and serial vector format (SVF) can be accomodated.

The IC's are installed on PCB's, which are in-turn installed in modules which comprise the system. The InstaCell library can pinpoint the exact root cause of a failure, supporting improved designs, improved MTTR's, spares and provisioning considerations, and also corporate TQM initiatives.

Conclusion

This paper has shown that common semiconductor failure modes can be mitigated through a methodology developed by Ridgetop Group. By having an early detection system, the target system can be configured to take adaptive action to correct the predicted problem prior to a catastrophic failure condition. The InstaCell™ library of cells is available for semiconductor designers to include on their designs, and provides a measure of safety and assurance that the deployed systems will maintain robust performance while in service. Because this methodology is available, it is recommended that IC's and SoC's that are in safety-critical applications adopt a prognostic methodology as proposed in this paper.

Author

Doug Goodman is President of the Ridgetop Group. Prior to Ridgetop, he was a co-founder and President of Opmaxx, a semiconductor Design-for-Test firm acquired by Credence Systems in 1999. He has held various positions as VP of Engineering, Systems Engineer and Product Design Engineer at Tektronix, Analogy, Bunker Ramo and Delco. He holds a BS in Electronic Engineering from California Polytechnic State University and an MBA from the University of Portland. He is on the Program Committee of the IEEE Mixed Signal Test Workshop and has served on various industry boards.

© 2000, Ridgetop Group, Inc. Tucson, Arizona