

Wafer Level Reliability for High-Performance VLSI Design

Bryan J. Root
and
James D. Seefeldt
Unisys Corporation
Semiconductor Operations
St. Paul, MN

Abstract:

As VLSI architecture requires higher package density, reliability of these devices is approaching a critical level. Previous processing techniques allowed a large window for varying reliability. However, as scaling and higher current densities push reliability to its limit, tighter control and instant feedback becomes critical. Previously, long-term package level testing that identified 100-year wearout mechanisms was adequate. Misprocessing resulting in slightly reduced wearout did not affect system performance. Due to scaling, however, normal lifetimes are approaching 20 years. Therefore, wafer level tests providing immediate feedback are essential to screen devices susceptible to any premature failure.

This paper describes several test structures developed to monitor reliability at the wafer level. For example, a test structure has been developed to monitor metal integrity in seconds as opposed to weeks or months for conventional testing. Another structure monitors mobile ion contamination at critical steps in the process.

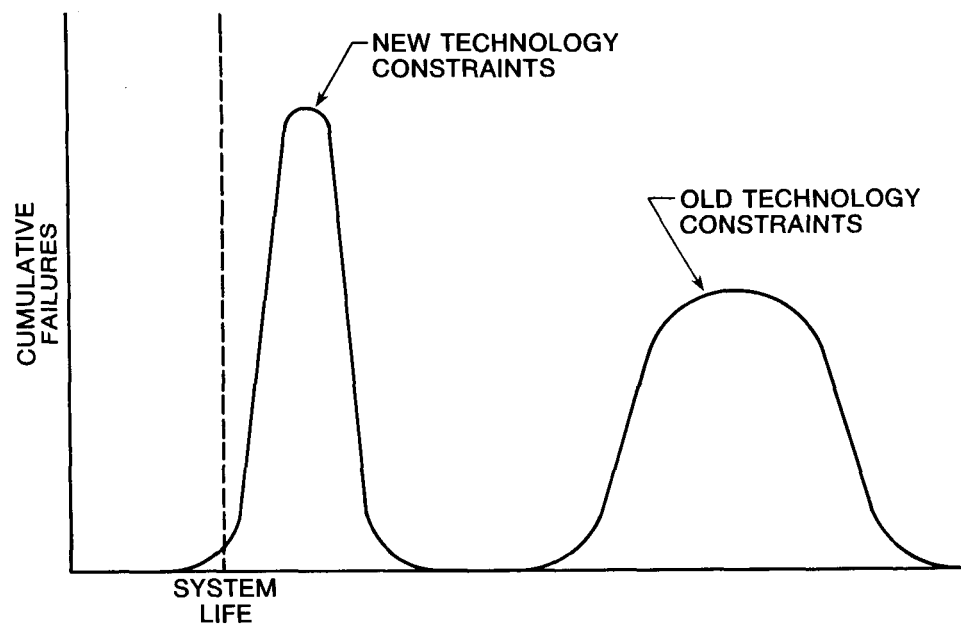
Thus the reliability jeopardy can be assessed during fabrication preventing defective devices from ever being placed in the field. Most importantly, the reliability can be assessed on each wafer as opposed to an occasional sample. Unisys Semiconductor is working on developing this technology.

PRECEDING PAGE BLANK NOT FILMED

1. INTRODUCTION

The electronics industry has seen explosive growth in the complexity of semiconductor devices. In only 20 years, computers previously built with discrete devices are now constructed with components containing hundreds of thousands of elements. As these devices become more complex, the reliability of these devices approaches a critical level.

Throughout the history of semiconductor devices, each new generation, even though more complex, exhibited a higher level of reliability. This is due to increased understanding of the physics of semiconductor devices as well as the advent of sophisticated instrumentation. As the industry moves from VLSI architecture to ULSI, the trend of increasing reliability must continue. The primary reason for this is because the lifetime of the devices has contracted to the point that any premature failure results in reduced system performance. Specifically, design rules in the $5\mu\text{m}$ range and higher normally show device lifetimes that might be in excess of 100 years. If the system life is expected to be a minimum of 20 years, and due to processing irregularities, the semiconductor device has a slightly reduced lifetime, system performance will not be compromised. However, as normal lifetimes of VLSI and ULSI devices approach 20 to 30 years, any premature device failure will degrade system performance.



Plot of Lifetimes vs. Cumulative Failures.
While System Life Has Remained the Same, Device Lifetimes Have Contracted.

The conventional methods of reliability screening, that of short-term burn-in to eliminate infant life failures, and long-term life tests at high temperature, will soon become inapplicable for many devices. The reason for this is increasing customization, cost, and shortened lifetimes. As an example, it has already been shown that standard burn-in of DRAMs,¹ while eliminating inferior devices, can substantially shorten the lifetimes of "good" devices to an unacceptable level.

Additionally, applying standard burn-in techniques to a small production run of semicustom devices could result in more devices being used in testing than are delivered to the customer. Thus, the trend is clearly forcing semiconductor manufacturers to adopt wafer level tests for process screening and reliability evaluation.

2. WAFER LEVEL TESTS

In the last four years, most US semiconductor manufacturers have organized research efforts in the area of wafer level testing for reliability, process screening and yield enhancement. This paper describes several wafer level tests under development at Unisys. The emphasis is directed toward ultra high performance bipolar process technology; however, many of these ideas can be applied to "standard" processes or modified to accommodate MOS reliability.

3. WAFER LEVEL ELECTROMIGRATION TESTS

In April of 1985, two papers were presented at the International Reliability Physics Symposium on wafer level tests for electromigration susceptibility. The first, which is called the Standard Wafer-Level Electromigration Acceleration Test or SWEAT is described here.² The other test, Breakdown Energy of Metal (BEM), was developed at Intel.³

3.1 Important Considerations in the Development of Accelerated Wafer Level Tests

First, the test must be of extremely short duration to achieve adequate throughput in the production line. Second, the test cannot impact the reliability or yield of the normal functional part. Third, the test must correlate to established data for conventional testing.

The SWEAT test makes use of the well-known equation for electromigration:⁴

$$MTF = AJ^{-N} \exp(e_A/kT) \quad [1]$$

where: MTF is the median time to fail
A is determined by the metallization
J is the current density through the metal lines
N is the current density factor
 e_A is the activation energy of the electromigration
K is Boltzmann's constant
T is temperature in Kelvin

The important variables to note in equation 1 are the current density and temperature. In normal electromigration testing, the current density is often increased to 10X above operating current densities and the device is placed on a hot chuck or in an oven to achieve a failure in a timely manner. This type of testing has been found to be valid where significant Joule heating of the metal strip is not present. This is due to the inability to accurately monitor the temperature. Thus, this type of testing takes days and often weeks to obtain statistically significant data.

To obtain a highly accelerated test for electromigration (i.e., to obtain failures in seconds as opposed to hours), the temperature and the current density must be raised to extremely high levels. This presents a problem in control and monitoring. For precise acceleration factors, the current density and temperature must be known, controlled and updated continuously during the test. The SWEAT test controls the temperature by correlating it to the instantaneous power density, where the power density equation is given in equation 2.

$$\text{Power Density} \equiv P = (VI)/V_L \quad \text{Watts/cm}^3 \quad [2]$$

where: V is the voltage
I is the current
 V_L is the volume of the line under stress

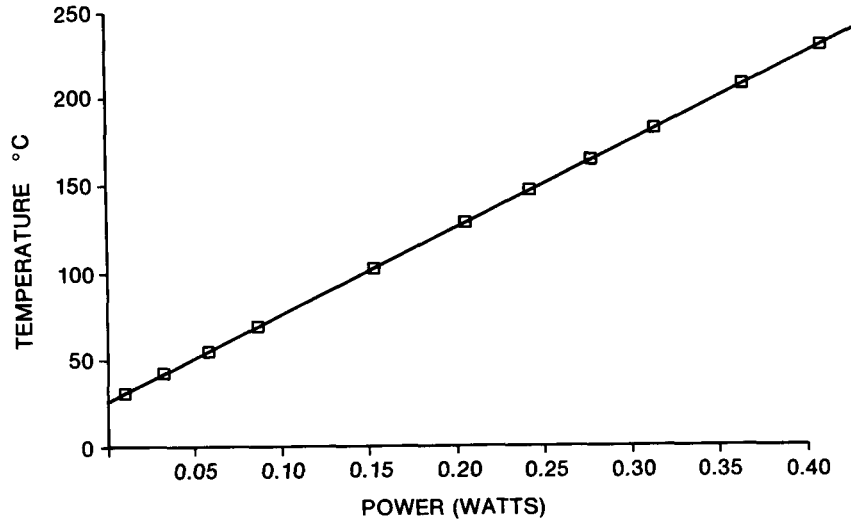
Thus, by monitoring the current through the test device, and controlling the temperature using power density, a constant acceleration factor can be applied. Equation 1 can be modified as:

$$TT_F = A(I/\text{cm}^2)^{-N} \exp [e_A/K(PQ - T_0)] \quad [3]$$

where the modified temperature expression is given by

$$T = PQ - T_0 \quad [4]$$

where P is the power density in equation 2 and Q is a constant slope of power density vs. temperature given in units $((\text{cm}^2 - \text{K})/\text{W})$, and T_0 is the ambient temperature. The power density vs. temperature curve is derived by first correlating resistance vs. temperature curves on a variable temperature wafer chuck. Power vs. temperature is then derived by raising the temperature of the line by increasing levels of power (I^2R); no external heat is used.



Plot of Temperature vs. Power.

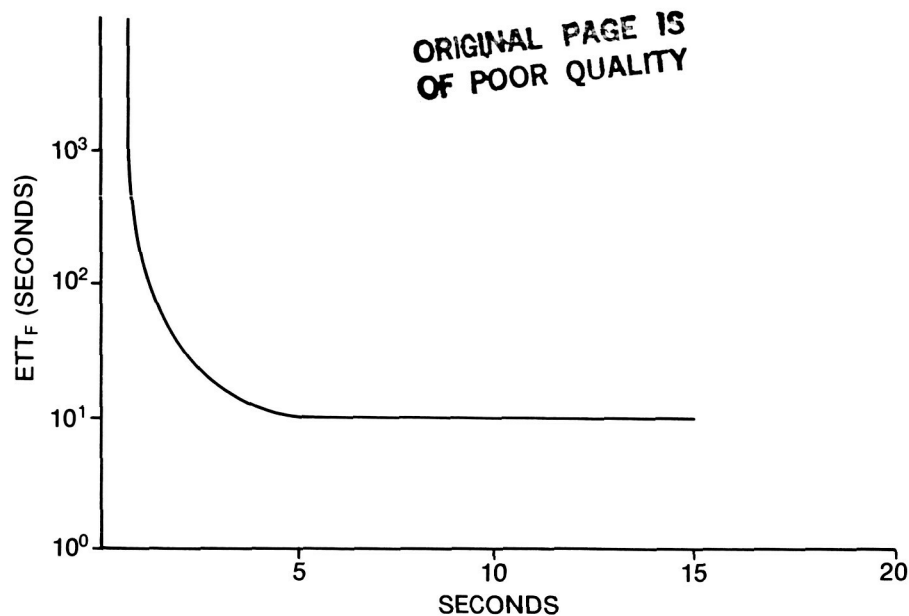
Note in All Cases This Has Been Found to Be a Linear Relationship.

3.2 Implementation of the Test

The implementation of the SWEAT test first requires characterizing what is considered to be minimally acceptable material. This is because the SWEAT test is designed as a simple pass-fail test. In a normal production mode, the benchmark established with the minimal acceptable material is compared with the production material. Thus, if the minimum specification material is set up to last an equivalent time to fail (ETT_F) of 15 seconds, then the current and power density in the production material is compared against this. Then equation 3 becomes:

$$ETT_F \equiv TT_F = A(I/\text{cm}^2)^{-N} \exp [e_A/K(PQ - T_0)] \quad [5]$$

where TT_F of the subject material is increased until it is equivalent to ETT_F .

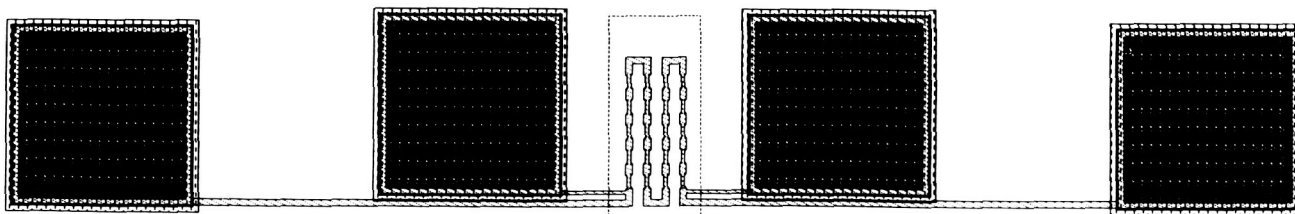


The Calculated TT_F Zeros in on the Target ETT_F During the Test.

3.3 Electromigration Test Structure

It has been known for some time that various physical parameters such as grain size, passivation, mechanical stress, metallization type and topography effect electromigration sensitivity. It is possible to concentrate several stresses in one area to enhance the electromigration characteristics.

High current gradients have been shown to enhance electromigration. Levine, et. al.,⁵ showed voids forming just inside the negative bond pad where stress from the electron wind changes rapidly. As discussed previously, temperature greatly enhances electromigration where voids tend to grow toward hot spots⁶ in metallization. Additionally, due to current sputtering techniques, metal thinning at steps causes current density gradients and heating, thus enhancing electromigration. A test structure was developed to concentrate these stresses improving the precision and resolution of the test. The structure essentially duplicates worst case conditions in a device. Also, real-time observations have shown voids traveling down substantial portions of the lines before enlarging to an open circuit. In the SWEAT structure voids are trapped in the narrow sections, thus enhancing the sensitivity.



The SWEAT Test Structure.

The test structure shown is a series of wide and narrow regions over topography. The wide areas are designed as thermal sinks to reduce overall temperature and emphasize the electron wind. The narrow regions route over steps which, depending on the structure and worst possible step coverage, can be up over poly or down over diffusion. In current designs there are 20 narrow regions to increase randomness and assure multiple grain boundaries.

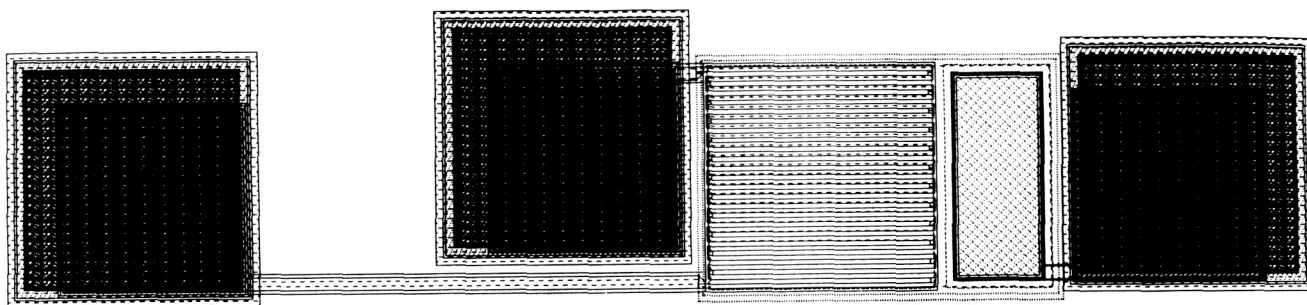
A modification of this structure can be used to monitor step coverage on a wafer to wafer basis. It involves using a SWEAT structure without topography adjacent to a SWEAT structure with topography. Comparing the two cancels most other mechanisms, but insulates step coverage as a potential failure.

4. MOBILE ION CONTAMINATION

The phenomenon of mobile ionic charge is a well known yield reduction and reliability concern. Substantial attention has been applied to MOS device degradation due to unstable threshold voltages; however, in bipolar devices contamination by alkali ions such as sodium and potassium or negative ions and trace heavy metal contamination can also cause device failure.

Device degradation in bipolar devices due to mobile ion contamination manifests itself by creating inversion layers. These inversion layers occur in isolation areas creating leakage from device to device. Additionally, with the new ultrahigh performance double-poly emitter bipolar devices, mobile ion charge can cause junction leakages above acceptable levels. Older bipolar devices operate at such high currents that minor leakage does not affect performance. However, newer generation bipolar devices, due to scaling and lower operating currents, result in increased sensitivity to junction leakage.

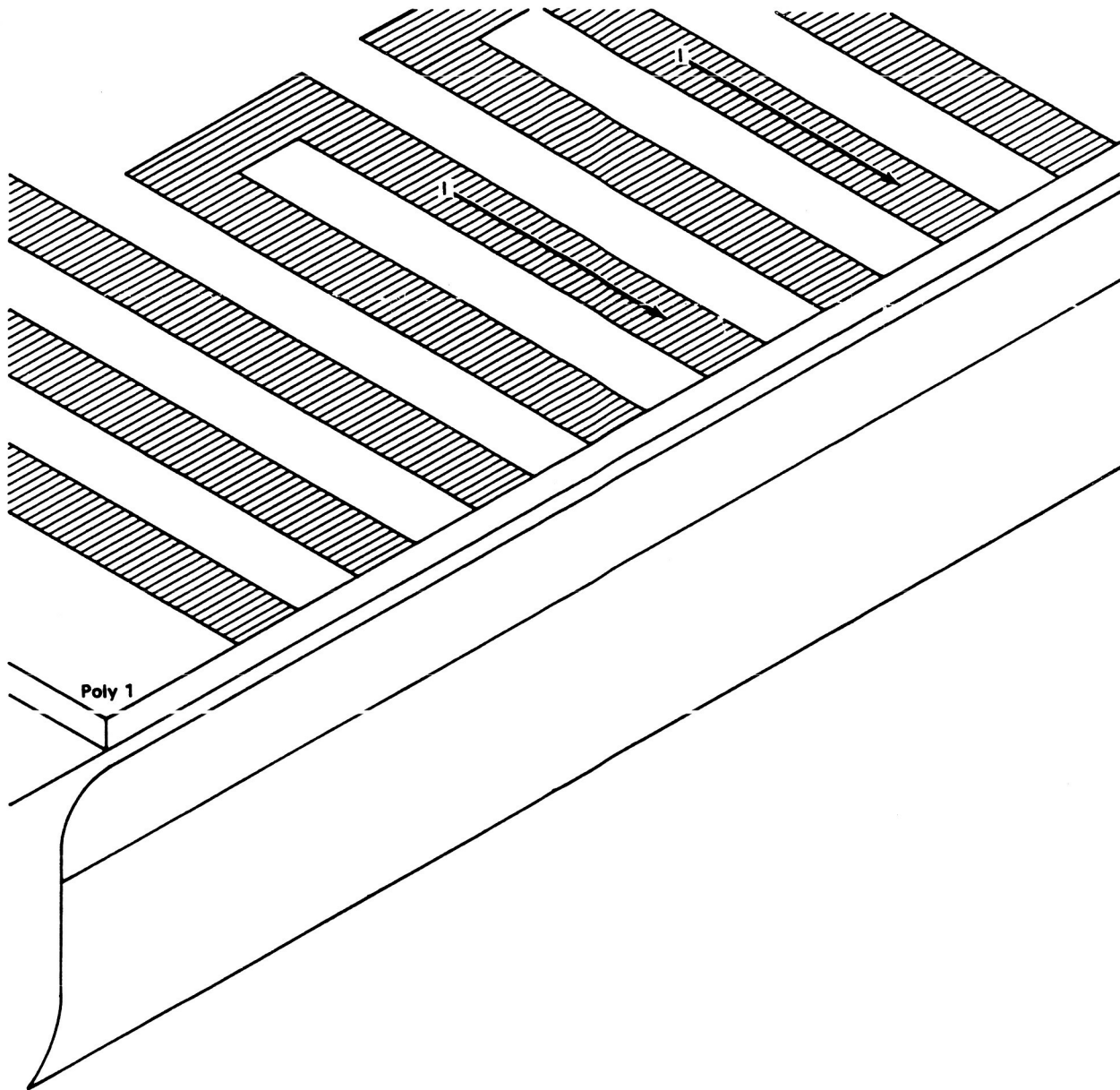
A rudimentary test structure for wafer level mobile ion contamination has been available for years in the form of CV dots. A new test structure has been designed, however, that avoids any external heating by employing self heating and can be placed into any scribe lane or dropout test die. This test structure is essentially a capacitor structure.



The Mobile Ion Test Structure.
Note the Serpentine Across the Top of the Capacitor Plate.

ORIGINAL PAGE IS
OF POOR QUALITY

The lower plate is formed by a silicon epitaxy mesa with an oxide layer cover. The upper plate is a P^+ polysilicon plate with a platinum silicide serpentine across the top. During the test, current is passed through the silicided serpentine increasing the temperature of the upper plate. The figure shows the characteristic cross section of the structure.



Cross Section of Mobile Ion Test Structure.

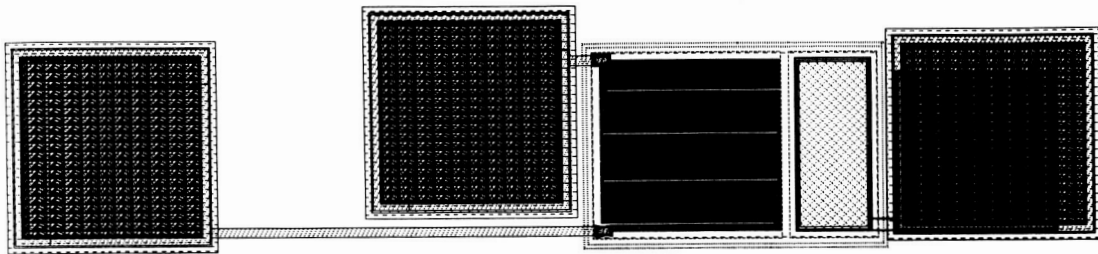
Since it has been shown that the ion mobility increases as temperature increases, the self-heated structure accelerates the associated mechanisms. The temperature and thus the required current is determined in the following manner. First, a linear relationship has been shown for resistivity vs. temperature for platinum silicide (PtSi), and a linear coefficient has been extracted. Using an external heat source temperature vs. resistance is determined. Second, using the resistant coefficient temperature vs. current can be accurate.

Thus, a minimum baseline is determined and C-V shifts of production material compared to it. Again, a simple pass-fail criterion is determined.

5. SCHOTTKY DIODE STRUCTURES

Schottky diodes⁷ are critical devices in VLSI, not only as circuit elements but also as process characterization structures. Schottky diodes are useful in measuring the density of dopants (N_D) trap centers (N_T) and minority carrier lifetime. Although, other tests are available for limited use, such as spreading resistance or DLTS, they are either destructive or require complex sample preparation and sophisticated instrumentation. Wafer level measurement allows rapid nondestructive analysis on the product wafer.

A Schottky diode test structure, using self-heating for temperature measurement, is shown below. The Schottky structure is formed by opening a serpentine structure to epi and then forming a PtSi layer in the exposed silicon area. The diode is heated by forcing a current laterally through the PtSi region. During heating the devices can be biased for testing at temperature.

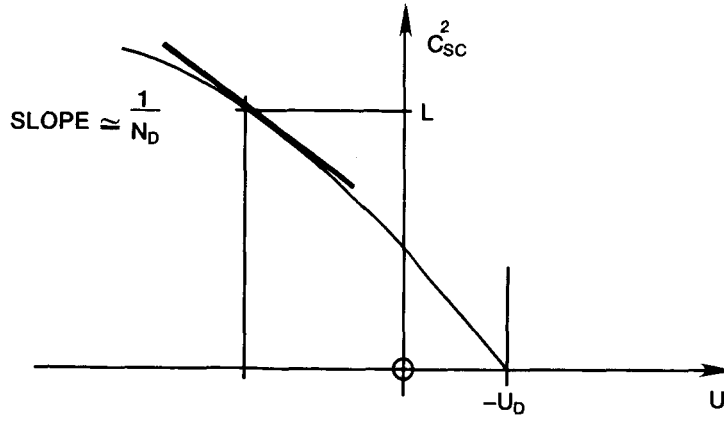


The Schottky Diode Test Structure.
Note the Serpentine PtSi Area.

Schottky diodes are very useful for measuring impurity concentration in doped silicon. When a Schottky diode is reversed biased, its depletion region spreads into the underlying lightly doped epi region. The capacitance of the reversed biased Schottky diode is a function of impurity concentration. The doping concentration can be determined from the relation in the following figure.

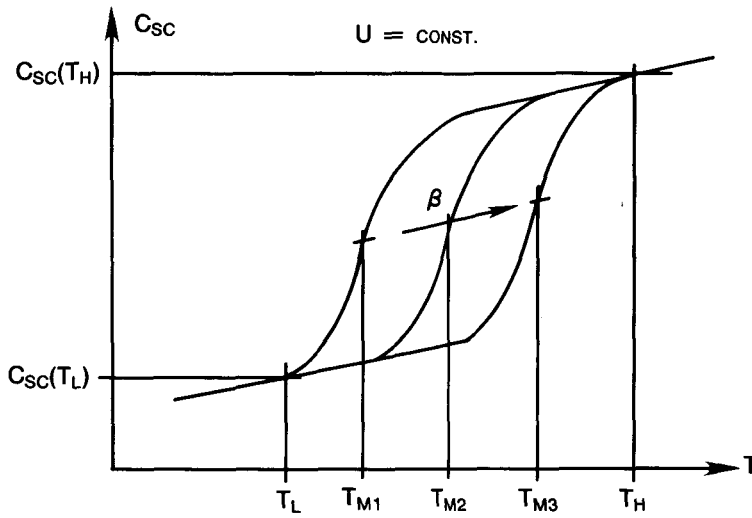
$$N_D = [2/(E_R E_0 q)] [d/dv(1/C_{SC}^2)]^{-1} \quad [6]$$

Where: C_{SC}^2 is the Schottky capacitance
 q is charge
 E_R is the dielectric constant
 E_0 is the permittivity of free space



Note That the Doping Concentration Can Be Obtained from the C_{SC}^2 vs. U Curve.

If the Schottky diode is kept at a constant reverse bias and heated, the capacitance will change. If the experiment is performed at several heating rates, mid-band gap impurity characteristics can be measured.



C_{SC} vs. Temperature Can Be Used to Obtain Mid Band Gap Impurity Characteristics.

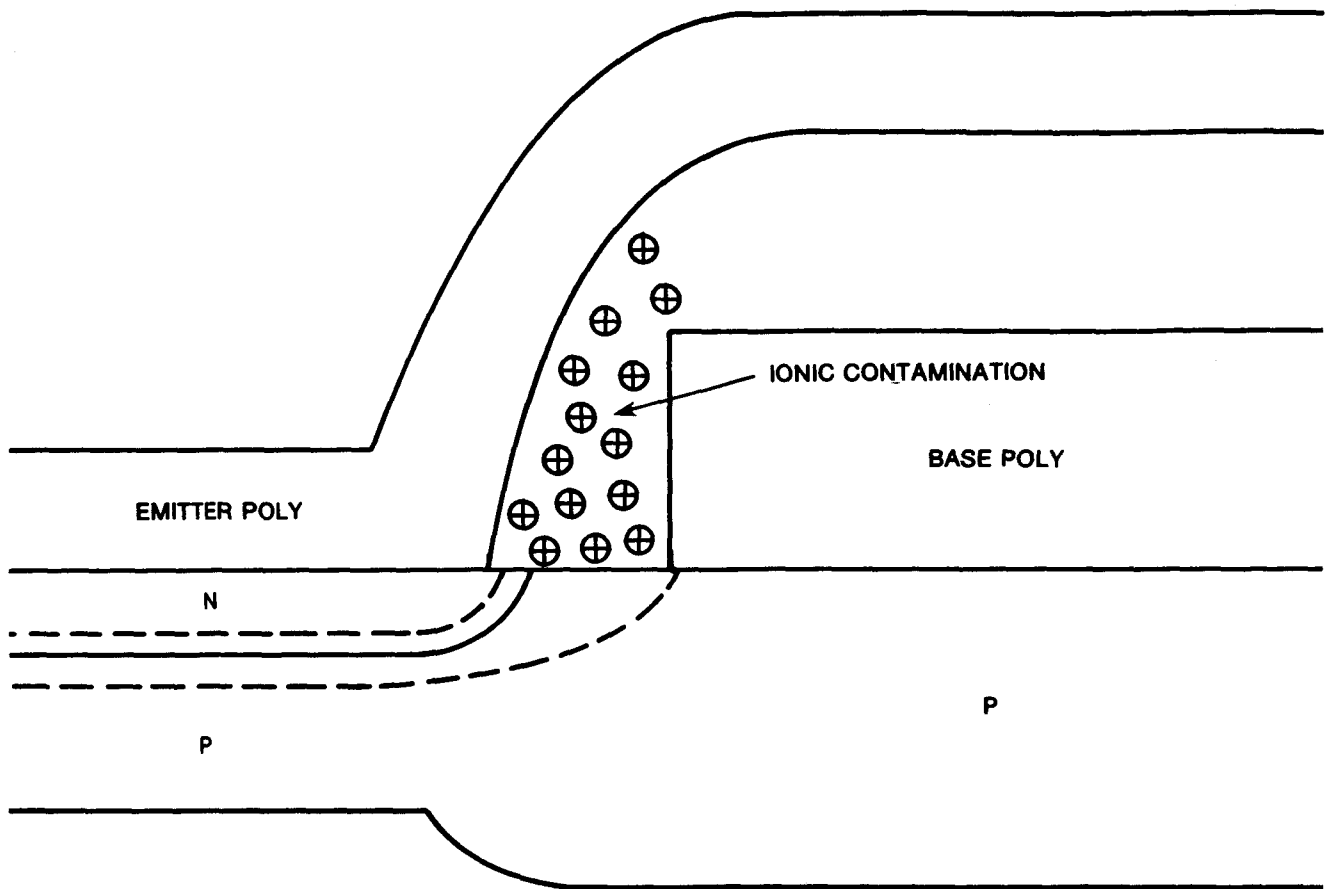
$$N_T = [2(-V)/(qE_R E_0)][V_D(T_L) - V_D(T_H)][C_{SC}^2(T_H) - C_{SC}^2(T_L)] \quad [7]$$

V_D is the voltage drop across the depletion region. From this an activation energy (E_T) can be determined:

$$E_T/KT_M = 1N[E_T/KB(1/t_2Kt_M/E_T)] \quad [8]$$

where B is the rate of increase in temperature.

An important concern is the integrity of the sidewall in a double polysilicon bipolar device. It can be seen that mobile ion contamination in the oxides over the space charge region can cause junction degradation. The figure below shows the sidewall region under normal conditions. However, if ions are present, a leakage path can be generated in the junction as shown in the second figure on the next page.



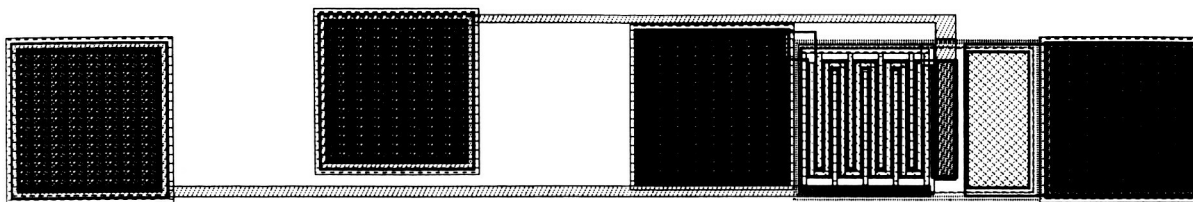
Ionic Contamination in an Unbiased Device.

6. WAFER LEVEL DEVICE RELIABILITY

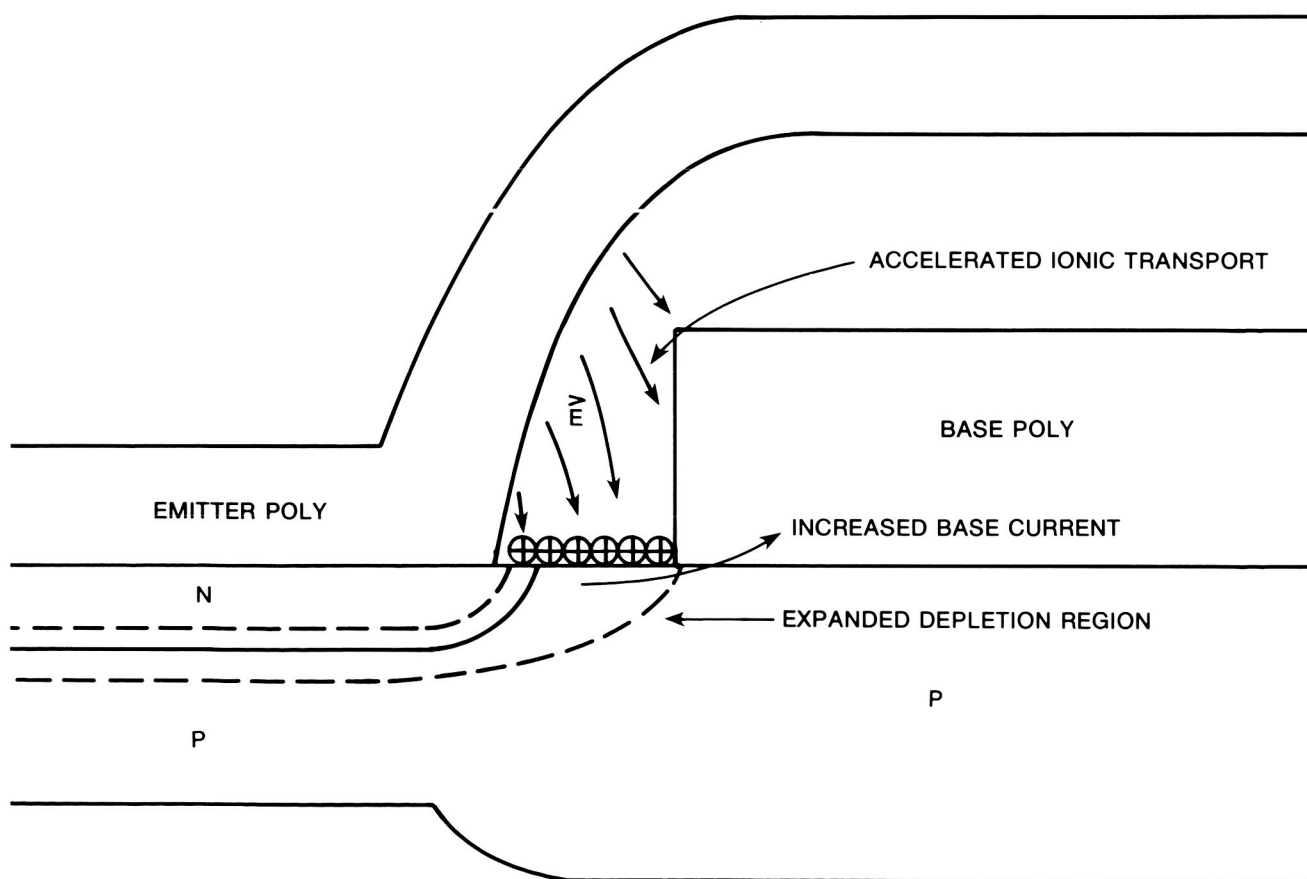
Historically, bipolar device reliability has been of minimal concern; however, as scaling approaches submicron sizes and current densities increase, possible failure mechanisms must be investigated. The use of deposited oxides for isolation on the device level is a potential reliability concern.

As device spacings are reduced, ionic contamination can increase parasitic leakage current. Additionally, as device currents are reduced, these parasitic currents could significantly degrade device performance.

It is well known that most failure mechanisms are accelerated by temperature, voltage and current. Thus, a bipolar transistor with self-heating capability has been designed. The test structure, shown below, is a double polysilicon device. The emitter poly formed in an extremely long serpentine pattern with pad connections on each end. In this manner, current can be passed through the emitter, not through the device, directly heating the device junctions.



Bipolar Device Reliability Test Structure.
Note the Serpentine Emitter Can be Used to Heat the Device.



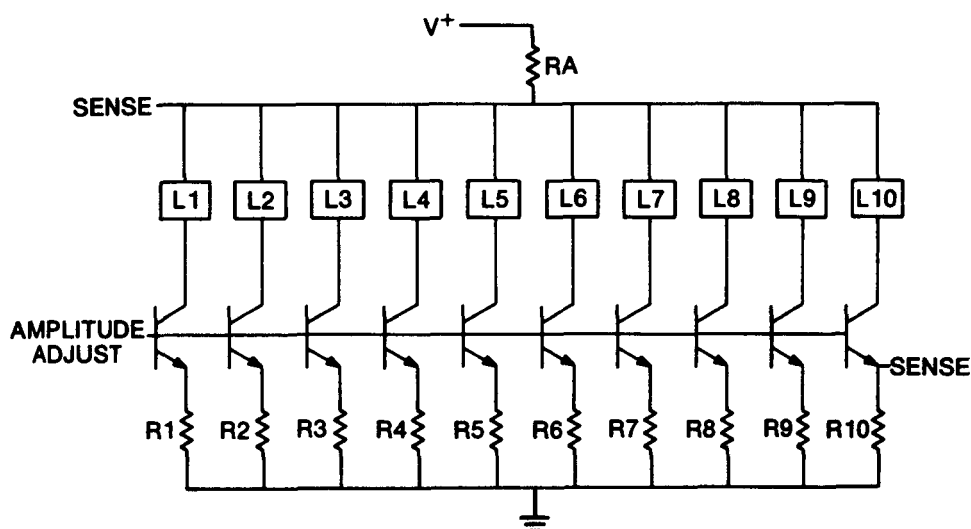
Biasing Causes Concentration of the Mobile Ions Creating a Leakage Path.

The structure shown in the figure* is also useful for basic device characterization. It is important to know device parameters over a wide temperature range. Typically wafers are heated on a hot chuck to make these measurements; with the heater structure the device can be heated to a known temperature and biased for characteristics, thus eliminating the need for an expensive hot chuck.

7. FUTURE WAFER LEVEL TESTING

It is clear that as the sophistication of semiconductor devices increases, wafer level testing must match that level. The future will show the level of integration of test structures increases. This increased sophistication will require drivers, sensors, and data collection to be implemented on the chip itself reducing the need for specialized off chip equipment. An added benefit to this is increased speed and ease of data collection. Thus, an entire series of tests could be multiplexed onto a small set of bond pads, increasing the usable scribe lane area.

An example of reliability integration is shown below where standard electromigration structures have been implemented.



Wafer Level Test Structure of the Future.
Integration of Drivers, Sensing and Calibration Entirely Into the Device.

Ten lines can be stressed using five bond pads where 40 pads would have been necessary in earlier structures. Pad 1 is the supply voltage. Pad 2 senses voltage changes across R1 indicating individual line failures. Pad 3 is the current adjust input. Pad 4 is the ground. Pad 5 is for calibration.

Another circuit that evaluates contact electromigration implements a multiplexed output to sense the status of each contact (the contacts are in parallel and of varying sizes). High temperature AC measurement can also be implemented due to the stability and AC characteristics of advanced bipolar devices.

8. CONCLUSIONS

The pace at which reliability research advances must match the increasing sophistication of semiconductor devices. The transfer of emphasis from expensive and time-consuming package-level testing to wafer level testing will accelerate. The absolute necessity of this can be seen in the shorter product cycles, increased customization, and increasing performance vs. price characteristics. The future will see increased reliability integration, computer aided reliability and more accurate models of die level failure mechanisms.

* C_{SC} vs. Temperature Can Be Used to Obtain Mid Band Gap Impurity Characteristics.

REFERENCES

- 1 W. Meyer, D. Crook, (Intel), "A Non-Aging Screen to Prevent Wearout of Ultra-Thin Dielectrics", I.R.P.S. (1985).
- 2 B. Root, T. Turner, (Mostek), "Wafer Level Electromigration Tests for Productional Monitoring," I.R.P.S. (1985).
- 3 C. Hong, D. Crook, (Intel), "Breakdown Energy of Metal (BEM) – A New Technique for Monitoring Metallization Reliability at Wafer Level," I.R.P.S. (1985).
- 4 J. Black, (Motorola), Proceedings of the Third International Congress on Microelectronics, p. 141, Munich, (Nov. 1968).
- 5 E. Levine, J. Kitcher, (IBM), "Electromigration Induced Damage and Structure Change in Cr-Al/Cu and Al/Cu Interconnection Lines," I.R.P.S. (1984).
- 6 R. Thomas, D. Calabrese, (RADC), "Phenomenological Observations on Electromigration," I.R.P.S. (1983).
- 7 J. Zemel, Nondestructive Evaluation of Semiconductor Materials and Devices, Plenum Press, (1978).