Quality Data

Data is only useful if it results in actionable information. Test and design features of Reedholm WLR structures ensure high confidence decisions are made based on results.

Process Induced Damage (PID) Protection

Each structure using a gate oxide has protection against unwanted PID. A fusible link between the gate and substrate eliminates any potential difference between the two. The PID structure itself limits damage to only the process step(s) it is intended to monitor using these same fusible links. Structure testing begins with opening fuse links using methods designed to prevent structure damage while fusing.

Reference Structures

Where appropriate, a reference structure is provided with the monitor structure. This configuration allows normalization of test results, dramatically improving data quality.

The step coverage structure, for example, uses a featureless line for this task as shown in Figure 3.

When test results from reference and monitor structures shift in the same direction by a similar amount, normal process variation can be eliminated.

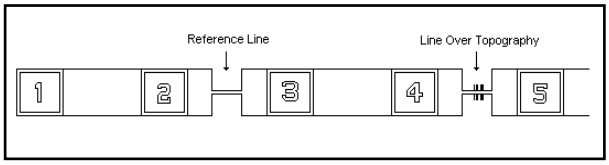


Figure 3 - Step Coverage Structure

Production Testing

With rapid testing, a quality WLR program provides statistically significant sample sizes in a short time span for quick analysis and process improvement feedback.

Self-Heating Elements

Elevated temperature is often used for acceleration. For electromigration, self heating from high current density stress accelerates failure with both current and temperature.

Others force current through either a highly doped layer of bulk silicon or a polysilicon heating element to generate heat. A metal line above the heater provides temperature feedback, where temperature is calculated from change in resistance.

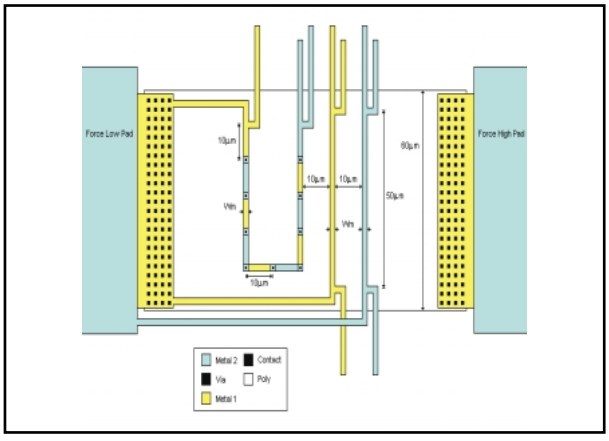


Figure 2 - Via Voiding Structure

The temperature coefficient of resistance (TCR) and ambient temperature resistance are needed for each layer of the structure. Figure 2 is a top view of the via and metal voiding structure showing the polysilicon heating element and thermometers.

High Acceleration Features

Some structure features are key to the high acceleration required for fast testing. For example, hot carrier testing utilizes an annular transistor where the area of most damage, the gate-drain interface, is maximized.

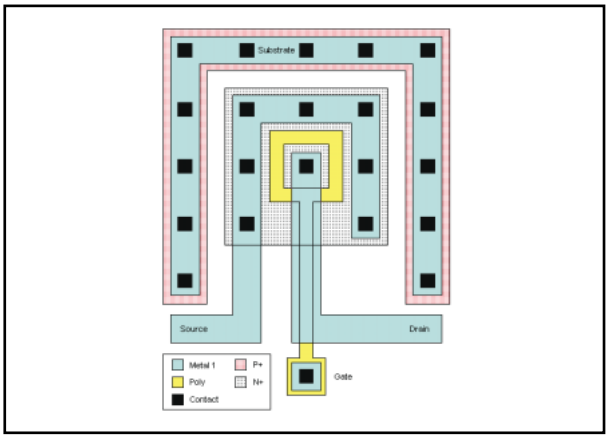


Figure 4 - Fast HC Structure

Documentation

Dimensions and Critical Parameters

Table 1 shows the design parameters required by the via and metal voiding structure.

Layout and Cross-Sections

Parameters		Symbols
Constant	Minimum via size	DR1
	Minimum metal overlap of via	DR2
	Minimum contact size	DR3
	Minimum contact space	DR4
	Minimum metal overlap of contact	DR5
	Minimum poly overlap of contact	DR6
	Metal 1 sheet resistance	Rm1
	Metal 2 sheet resistance	Rm2
	Poly sheet resistance	Rpoly
	Poly TCR	TCRp
	Field oxide thickness	h
	Field oxide thermal conductivity	k
Variable	Metal line width	Wm
	Metal line space	Sm
	Poly heater width	Wp
	Poly heater length	Lp
	Poly heater current	Ip
	Poly heater resistance	Rp

Table 1 - Via & Metal Voiding Structure

Top views and cross-sections clarify detailed structure features. Figure 5 is a top view of the long line electromigration structure. Figure 6 shows a cross-sectional view of the contact electromigration structure. These views show geometric relationships and how design rules are applied.

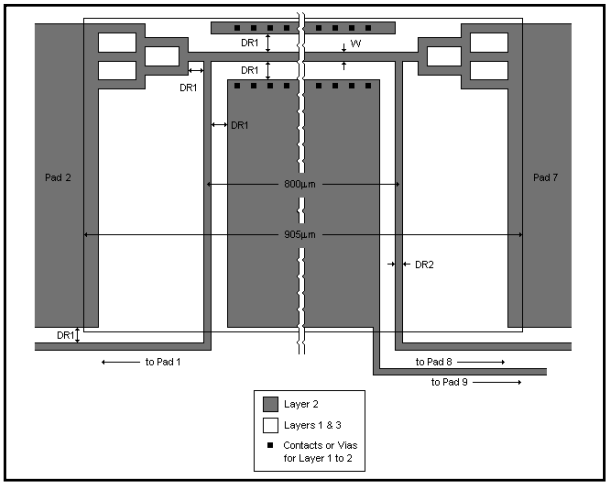


Figure 5 - Long Line EM Structure

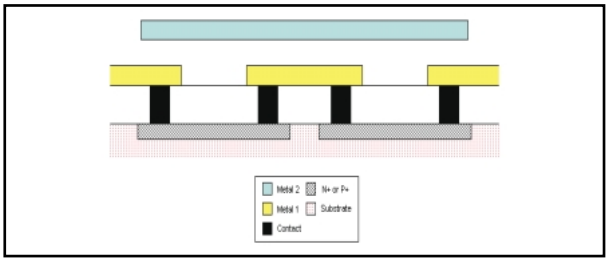


Figure 6 - Contact EM Structure

Calculations

Many structures require calculation of critical dimensions, based on process parameters shown in Table 1. These calculations ensure structure performance, accuracy, and that test system compliance limits are not exceeded. For example, the equations below calculate the power needed to increase a poly heater by ΔT°C.

$$Power = I_p^{\,2} R_p = k L_p W_p \Delta T / h$$

$$R_p = R_o (1 + TCR_p \Delta T)$$

$$R_o = R_{poly} L_p / W_p$$

All calculations are provided as standard documentation.

Other structure features are derived from characterization or modeling to increase user understanding of each structure. These factors contribute to data integrity and structure effectiveness. An example of the characterization driving test structure changes is shown in Figure 7, where the thermal profile of structure containing heat sinks was so bad that a change was made to an 800µm line without heat sinks.

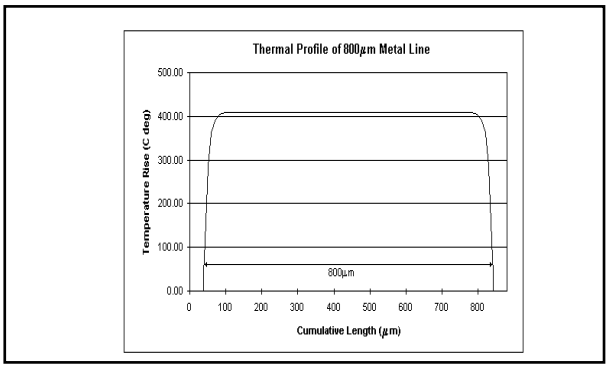


Figure 7 - Long Line EM Temperature Profile

References

Reedholm's documentation includes references to literature relevant to each structure or failure mechanism.

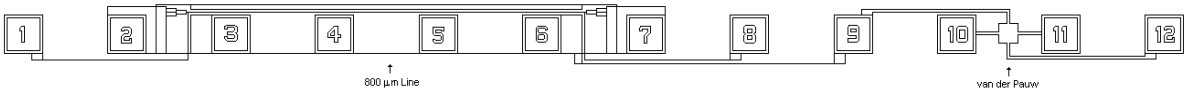
Test Structures

Structures are delivered in GDSII file format in 1x12 pad arrays that fit into 100μm scribe lanes or test element groups, using 75 x 75μm pads on 200μm centers. The structures listed in Table 2 are for a single poly, three metal process requiring 34 pad arrays. The quantity of structures is a function of process complexity.

Structure Type		Qty	Structure Type		Qty
Interconnects	Metal Electomigration	12	Gate Oxides	Gate Oxide Integrity (Area & Edge Intensive)	6
	Metal Step Coverage	15		Process Induced Damage	16
	Metal Stress Migration	6		Gate Oxide Charge Trapping (N & P Substrates)	2
	Via Electromigration	8	Interfaces	Fast Hot Carriers (N & P Substrates)	4
	Via Voiding	2	Insulators	Mobile Ion Contamination	4
	Contact Electromigration	6		Spin-On-Glass Charging	3
	Junction Spiking	2		Interlevel Dielectric Strength	8
				Top Passivation Strength	8

Table 2 - Structure Deliverables

Metal Electromigration

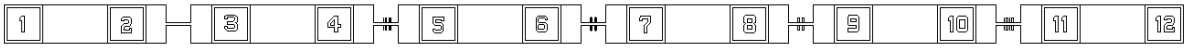


Physics: Mass flow in a metal line is accelerated by temperature and current. Acceleration factors based on Black's equation provide a sensitive indicator of process electromigration.

Structure: There are two pad sets for each metal layer: one minimum width and one 3μm wide. Each pad set has both a four terminal, 800μm line plus a four terminal van der Pauw structure.

Test Method: SWEAT or Isothermal EM.

Metal Step Coverage

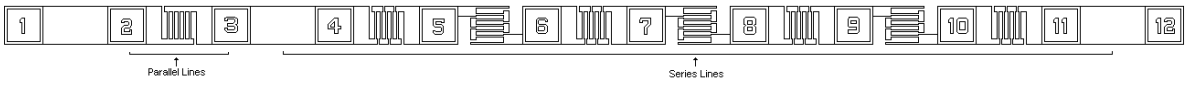


Physics: Lack of ideal step coverage over topography results in thinning and/or narrowing of a metal line. A reduction in cross sectional area results in higher current densities and higher temperatures during stress. Black's equation can be used to show that electromigration lifetime is a more sensitive indicator of thinning than the alternative of comparing fusing currents.

Structure: For each metal layer, a series of five 3μm x 50μm four terminal metal lines is used. One test line is flat, the other four are over topography selected for worst case, process specific step coverage.

Test Method: SWEAT or Isothermal EM.

Metal Stress Migration

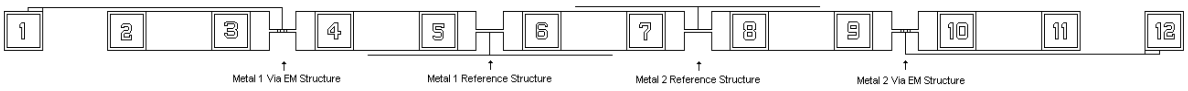


Physics: Because of the differences in thermal expansion coefficients between metal, substrate, and interlevel dielectric, voids can be generated in metal lines during cooling of the wafer.

Structure: Reference and void detection structures are provided for each metal layer. The reference consists of five parallel line segments, while the void detector has 35 series segments. By being in parallel, the reference structure effectively reduces the possibility that a void in one or more of the parallel segments affects the lifetime characteristics of the structure.

Test Method: SWEAT or Isothermal EM.

Via Electromigration

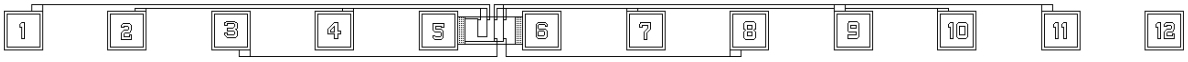


Physics: For plug and barrier technologies, electromigration occurs in the metal leading from the via. Otherwise, electromigration is monitored in the via itself.

Structure: For each pair of metal layers, four structures are provided. Two accelerate electromigration in metal layers around the vias, while the other two are reference structures (one for each of the two metal layers).

Test Method: SWEAT or Isothermal EM.

Via and Metal Voiding

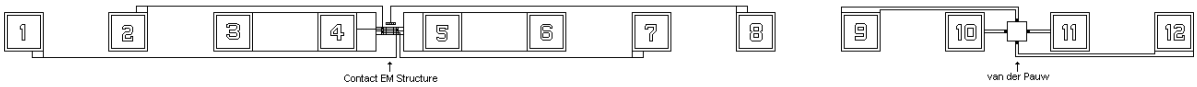


Physics: Problems can arise in and around vias due to contamination, reduction in contact area, etc., that are exacerbated by thermal cycling. Likewise, void growth in metal lines due to thermal cycling is also a concern.

Structure: For each pair of metal layers, one structure is provided for accelerating voiding within vias between the two layers and within the layers themselves. The structure consists of a short via chain, a reference structure for each of the metal layers, and a heating element.

Test Method: Resistance of the structures before and after each thermal cycle is monitored and compared to the reference structures.

Contact Electromigration



Physics: Unbound silicon atoms, whether alloyed with the metal or scavenged from beneath the contacts, are transported through the metal by electron flow. The silicon precipitates in the metal or at the metal-silicon contact.

Structure: For each type of metal-silicon contact, (N+, P+, and polysilicon) one pad array is provided for accelerating contact degradation. Each pad set has a contact structure and a van der Pauw structure for sheet resistance measurement. The structure is self heated by current through the contacts using a thermometer from the second metal layer for temperature control. A substrate tap is provided for N+ and P+ structures so leakage can be monitored during stress.

Test Method: Structure resistance, as well as junction leakage for the N+ and P+ structures, is monitored during high current stress.

Junction Spiking

Physics: Silicon lattice atoms beneath a metal contact can be replaced under thermal stress by metal atoms. If enough metal atoms displace silicon, the junction formed by the contact region and oppositely doped silicon beneath the region can be compromised.

Structure: For each type metal-silicon contact (N+ or P+) one structure is provided for accelerating junction spiking of metal into silicon beneath contacts. Each structure has a polysilicon heating element and substrate tap for measuring junction degradation. The contact structure is the thermometer element, providing thermal feedback to the stress algorithm.

Test Method: Temperature is the accelerating agent. No current is forced through metal-silicon contacts.

Gate Oxide Integrity

Physics: Gate oxide reliability is sensitive to processing effects at poly and field oxide edges as well as within the gate capacitor area. These effects can be differentiated using multiple gate capacitor designs.

Structure: Three types of polysilicon gate oxide capacitors are provided for each channel type. One is a polysilicon plate over an equal size active area. Second is polysilicon bars over an encompassing active region. Last is a polysilicon plate over bars of active region.

Test Method: While either the Jramp or Vramp tests produce approximately the same results, the Jramp test is less dependent on oxide thickness, permits more control of trap generation stress level, and provides more precise indication of trap generation.

Process Induced Damage

Physics: Various types of processing, and especially those which create or use plasmas, generate large surface to substrate potentials that cause damage to thin oxides.

Structure: Multiple pad sets with minimum size annular transistors are connected to antennas about the size of a probe pad. There are three types of antennas: plates for area dependence, meshes for edge dependence, and vias/contacts. A reference transistor with no antenna is included in each pad set. Each transistor is isolated from subsequent process damage with a fuse to the substrate.

Test Method: Gate-substrate fuses are blown and transistor parameters are measured before and after tunneling current stress.

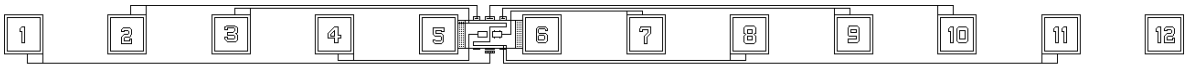
Fast Hot Carrier and Gate Trapping

Physics: Under normal operating conditions, hot carriers are created near the transistor drain causing localized damage within the gate oxide and at the Si-SiO_2 interface.

Structure: Annular transistors with drains enclosed by gates are used for these failure mechanisms. Three gate areas are provided for each transistor polarity, starting with design rule minimums. To minimize process induced damage, transistor gates are shorted to respective substrate connections with metal 1.

Test Method: Gate-substrate fuses are blown and transistor parameters are measured before and after a stress cycle. Hot carrier stress is performed by forcing a relatively large current from drain to source with the gate tied to the drain. Gate trap stress is performed using the constant current mode of the Jramp test.

Mobile Ion and SOG Charging



Physics: Ion mobility increases with temperature. A large positive bias and sufficient heat can push all unbound positive ions to the Si-SiO_2 interface. Conversely, a large negative gate bias and heat can pull positive ions to the metal-oxide interface.

Structure: Each metal and polysilicon layer has two field transistors (one for SOG and one for mobile ions). A bulk silicon or polysilicon heating element is also provided, with a metal line thermometer providing the temperature feedback.

Test Method: Threshold voltage of the field transistors is measured before and after thermal cycles. During the thermal stresses, metal gate bias is applied to attract and repel mobile charge.

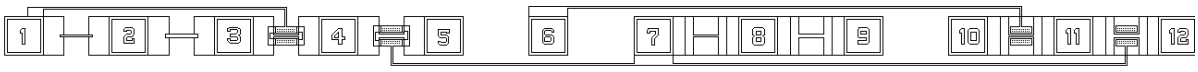
Interlevel Dielectric Strength

Physics: Dielectric processing problems (thinning, reduced densification, voids, cracks, and filaments) can lead to reduced reliability or breakdown voltage.

Structure: Multiple types are provided, including polysilicon-silicon, poly-first metal, and between subsequent metal layers. For each, a flat structure is included with one over worst case topography.

Test Method: Because of the high breakdown voltages involved (>5MV/cm), Reedholm's 1500V module is required. A ramped breakdown measurement is made to detect the structure breakdown voltage (V_{BD}).

Top Passivation Strength



Physics: Thermal expansion coefficient of metal is almost two orders of magnitude greater than passivation. Thus, passivation strength would affect the fusing point of a metal line.

Structure: The structure is based on the theory that fusing of a top metal line is a relative indicator of top passivation strength. Minimum width and 5μm wide lines are over various stacks of underlying metal features.

Test Method: A ramped current is used to fuse each line.

Implementation

Program Scope

Implementing WLR requires an approach that builds confidence. Extensive work with the structures is needed to establish their ability to properly identify potential reliability problems for a given process. Designing experiments that check the process corners is advised prior to putting them into production use. Establishing their sensitivity to previous process related reliability problems is key to getting buy-in from the process engineering team that they provide actionable information.

Start-Up Sequence

The following sequence has been used in successful implementations world-wide:

- Non-disclosure agreements are exchanged.
- Process design rules are supplied.
- Structures are customized and delivered.
- Customer design rule checks are performed.
- Structure design review meeting is held.
- Structures to be used are selected.
- Masks are generated and wafers fabricated.
- Training on the structures and testing is provided.

Role of the Test System

Good structures and algorithms are rendered useless without the right test equipment. The structure and physics dictate the stress and measurement limitations of an effective WLR program. Reedholm testers are tailored for the task of collecting meaningful WLR data to provide better performance, speed, and accuracy.

Power is Critical

With systems capable of forcing 600V through standard matrix relays and options enabling up to 1500V, any thin or thick oxide layer can be properly evaluated. Thick films like interlayer dielectrics (ILD) may still be up to a micron thick even in the most advanced processes. Breakdown voltage for such an oxide would be as high as 1000V, with no measurable increase in tunneling current at the 100V to 200V available from standard dc systems.

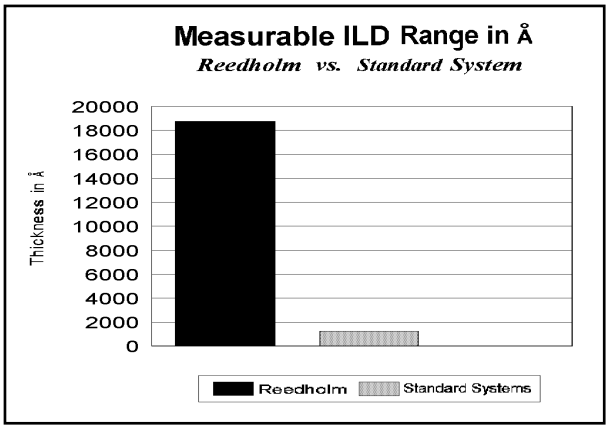


Figure 8 - Measurable ILD Thickness at V_{BD} = 8MV

Many WLR tests use heat as an acceleration factor. As a result, there is a need for high current since doubling the current quadruples power and hence heating. Reedholm's programmable power supplies provide 400mA for test structures using on-chip heaters and 550mA for self heated structures with options enabling up to 2A. High power is critical to such WLR tests as electromigration, via voiding, junction spiking, and mobile ion contamination.

Speed Impacts Accuracy

Typical WLR tests for gate oxide integrity involve ramping voltage or current until failure is observed in the oxide. With ramp times as short as a few seconds, any uncertainty in time to failure will translate directly into an uncertainty in the cumulative charge sustained by the structure. Unless timing enhancement techniques are used, the maximum uncertainty will be the measurement loop time. Since so much of the cumulative charge occurs in the last few ramp steps, this uncertainty can lead to large errors. For some systems, a 100ms force/measurement cycle is all that is attainable.

Reedholm software and system architecture provides ramps as fast as 3.2ms with timing uncertainty of less than 50ns. The time at each ramp step is known with that certainty, inclusive of any meter range changes. This resolution yields the best possible calculation of cumulative charge to breakdown (Q_{BD}).

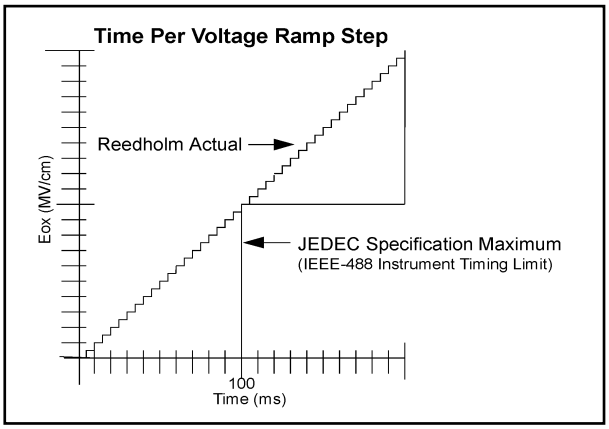


Figure 9 - Timing Resolution Impact on Vramp Testing

Sensitivity Increases Knowledge

Measurement sensitivity on the tester backplane is irrelevant if it does not translate into sensitivity at the probe tips. A typical tester and probe card combination may provide only 0.1nA sensitivity at the probe tips.

Reedholm can measure with 0.05pA/V resolution; even lower using digital averaging techniques at the probe tips. Design, material selection, and extensive characterization allow those current levels to be realizable using modern automated probing systems. Orders of magnitude more sensitivity translates into the ability to detect thinner oxide and smaller transistor problems earlier and more accurately.

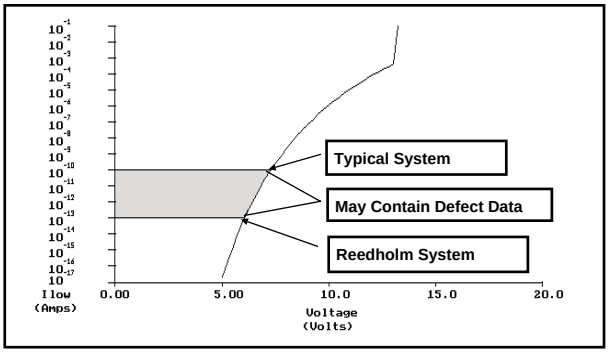


Figure 10 - Fowler Nordheim Tunneling Current (Typical)