KEITHLEY A GREATER MEASURE OF CONFIDENCE

Figure 1. Typical semiconductor reliability curve.

Wafer Level Reliability Testing – A Critical Device and Process Development Step

Dave Rubin and Yuegang Zhao, Keithley Instruments, Inc.

THE continuing push for more devices on a chip and faster clock speeds is driving the demand for shrinking geometries, new materials, and novel technologies. All these have a tremendous impact on the lifetime and reliability of individual devices due to increased fragility, higher power density, more complex devices, and new failure mechanisms. Processes that once produced devices with 100-year lifetimes may now yield only 10-year lifetimes – uncomfortably close to the expected operating life of systems using these devices. The smaller margin of error means that lifetime reliability must be designed in from the start and constantly monitored, from device development, through process integration, and into production; even small lifetime changes can be catastrophic to today's devices.

While reliability testing can be done at the packaged device level, many IC makers are migrating to wafer level testing for a number of reasons, including the need to test further upstream in the manufacturing process. Wafer level reliability (WLR) testing also eliminates much of the time, produc-

tion capacity, money, and material lost if the packaged device fails. Turn around time is much less, as a wafer can be pulled directly off the line and tested without the delay of sending the part away for packaging, which can be up to a two week process. Much of the testing is the same, allowing for relatively easy migration to wafer level testing.
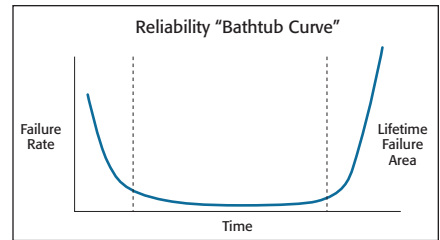
Stress-measure testing is a common technique used to evaluate operating lifetimes and wear-out failure mechanisms in semiconductor devices. This testing is focused on failures on the right side of the typical failure rate bathtub curve (*Figure 1*) – i.e., failures not associated with infant mortality or manufacturing failures.

Stress-measure tests can quickly generate curves that are extrapolated to predict the operating lifetime of devices. Such data are used to evaluate device designs and monitor manufacturing processes. Since typical device lifetimes are measured in years, techniques are needed to accelerate the testing. The most efficient method is to over-stress the device, measure degradation trends of key operating parameters, and extrapolate the data to the full lifetime. In *Figure 2*, for example, the lower right portion of the curve (collected data) was generated using high stress conditions. The data generates a line that can be used to predict device lifetime under normal operating conditions (upper left portion of the curve).

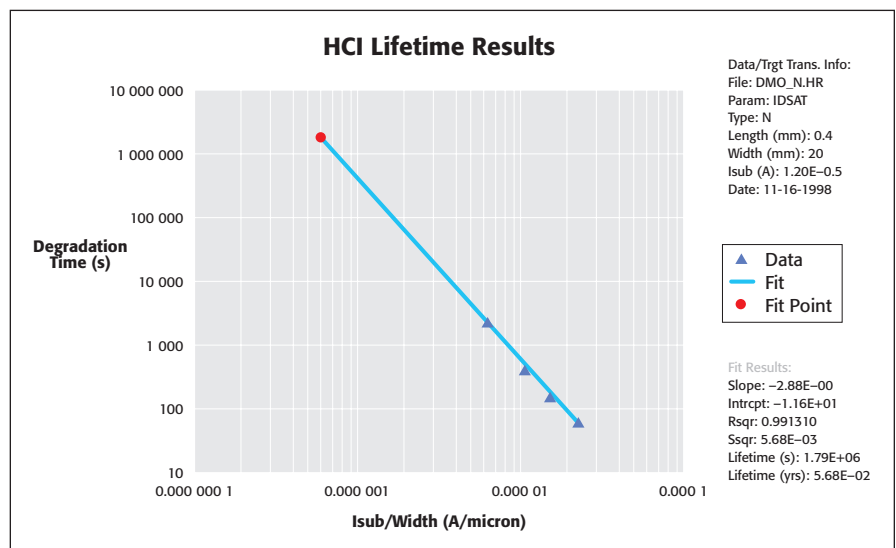Common WLR tests that use stress-meas-

Figure 2. Example of lifetime reliability extrapolation from HCI testing.

ure techniques have included Hot Carrier Injection (HCI) or Channel Hot Carrier, Negative Bias Temperature Instability (NBTI), Electromigration, and Time Dependent Dielectric Breakdown (TDDB) or Charge to Breakdown (Q_{BD}). These tests have become critical in mainstream CMOS device development and process control. (See Appendix for a description of traditional HCI and NBTI tests.)

New scale factors and materials now require modifications to these established tests and demand instrumentation features that can implement the new techniques. Two good examples are tests to overcome measurement challenges related to Negative Bias Temperature Instability (NBTI) in PMOS devices, and those associated with charge trapping phenomena in transistors with high κ gates.

Degradation Relaxation in NBTI Test

A unique characteristic of NBTI degradation is that it can relax when stress is off (see *Figure 3*) [1]. When gate voltage stress is turned off, the degradation of drain current and threshold voltage may recover and change back toward their original value [1]. The rate of recovery is strongly dependent on temperature. At room temperature, as much as 100% recovery has been reported. If stress is resumed on the gate after recovery, the degradation will follow the previous degradation curve. At higher temperatures, there will be a portion of the degradation that is irreversible. This is called degradation lock-in.

The recovery of I_d degradation when stress is not present presents a great challenge in parallel NBTI testing. The traditional approach, which for a long time has been used to test HCI degradation, is to stress devices in parallel, then disconnect from
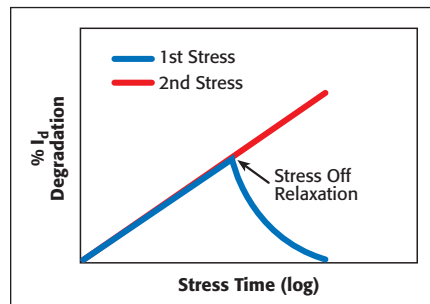
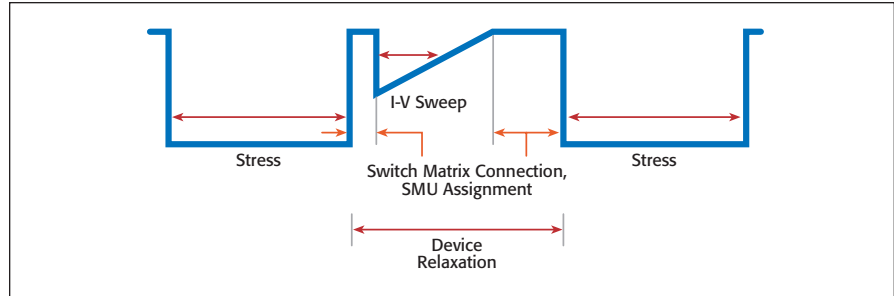*Figure 3. NBTI degradation and recovery.*

Figure 4. Possible device degradation relaxation during stress-measure transition.

the stress source and measure them in sequence (*Figure 4*). This technique presents two problems: (1) switching from stress to measure takes enough time that degradation recovers after the stress is switched off, and (2) devices in the queue during the measure phase will experience different amounts of relaxation. The degradation of the last device to be measured may be only a fraction of that for the first one. This measurement challenge requires switch-less stress and measure for multiple devices in parallel as well as a new technique to estimate V_t degradation, using only a few points instead of a full I_d–V_g curve to measure V_t degradation [2].

Another aspect of the NBTI recovery problem is associated with typical transistor operation, where it is turned on and off very often. When the transistor is off, NBTI degradation may recover. Therefore, if one uses the traditional DC stress and degradation technique, there will be no recovery effect and it may underestimate the lifetime of the transistor [3, 4].

One approach to handling this recovery dynamics problem is to use pulse stress instead of DC stress [5]. In this technique the transistor is pulsed alternately between on and off conditions. Then degradation of threshold voltage is measured as a function of pulse frequency. This test routine provides some very important information about the nature of recovery in different applications. For example, the switching frequency is not the same for transistors in different circuits with different functionality. The frequency dependency of NBTI degradation may reveal that some part of a circuit will fail before the rest.

Charge Trapping in High κ Gate Dielectrics

While high κ material can help solve ultra-thin gate leakage problems with leading edge processes, there is no free lunch. Associated with this advantage are several technical hurdles that must be overcome. One of them is the transient charge trapping in bulk traps when gate bias is turned on, resulting in a shift of the threshold voltage. This charge trapping problem introduces error in channel carrier mobility measurements due to a reduction in drain current [6]. In addition, it overshadows device parameter degradation in HCI, NBTI, and TDDB testing. This occurs because most of the observed degradation occurs as a result of charge trapping in the film [7], whereas device parameter degradation is actually the desired observation.

The issue in charge trapping is its transient behavior, i.e., its effect is strongly time dependent. The traditional DC method will not capture, or will only partially capture, its effect. A pulse method has now been introduced to study the transient charge trapping effect [6, 8].

Figure 5 shows two different test setups for a Single Pulse Charge Trapping (SPCT) measurement. In each measurement setup, a pulse is applied to the gate of the transistor, while its drain is biased at a certain voltage. The change in drain current, resulting from the gate pulse, appears on the scope. The difference between these two setups is that the one in *Figure 5b* has much higher bandwidth and, therefore, can capture much faster pulse responses (down to tens of ns). At such high speed, traps are unlikely to respond. Therefore a "clean" transistor response with minimum charge trapping effect can be measured [9]. *Figure 6* shows two SPCT measurements with long pulse width ("square") and short pulse width ("triangle"); the hysteresis in the long pulse width I_d–V_g curves is the result of charge trapping. This hysteresis is absent in the shorter pulse SPCT measurement [8] as a result of less time for charges to be trapped in the film.
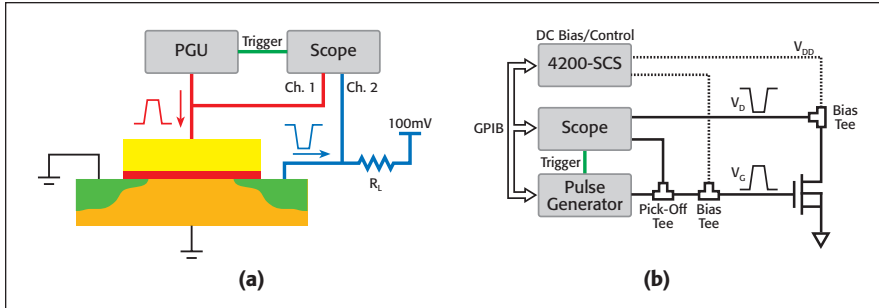
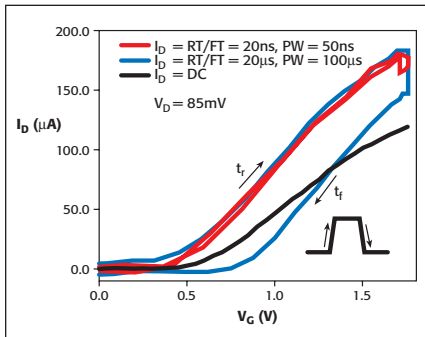Figure 5. Two different Pulsed I-V test setups to study transient charge trapping inside high κ gate stacks.

Figure 6. Single Pulse Charge Trapping measurements: long pulse width with hysterisis using the setup in Figure 5a, and short pulse width with little hysterisis using the setup in 5b. DC result is shown as reference.

Since there is much less charge trapping effect with very short pulse widths, the drain current measured is higher than under DC conditions (*Figure 7*). This results in a higher predicted channel carrier mobility when pulse I-V data are used to generate a model, which is more representative of transistors that are switching very fast (i.e., those that will not experience full charge trapping effects).
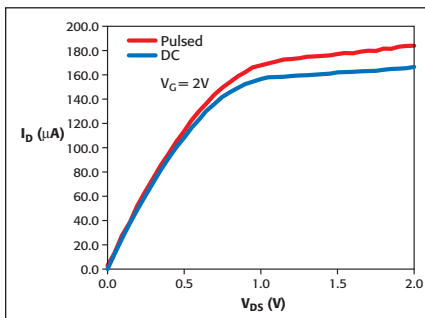

Figure 7. Comparison of V_d–I_d curves between pulsed and DC measurements.

Because of the complications in trying to characterize charge trapping effects for transistors with different functionalities, it would be nice if modeling engineers did not have to worry too much about it. They

can come closest to this ideal by using instrumentation and a test set-up that avoids the artifacts associated with DC or slower pulse measurements. The resulting models will help optimize designs for the best operating conditions. Additionally, in continuing efforts to improve film quality and remove as much charge trapping degradation as possible, process engineers need these measurement techniques to characterize and track improvements.

Besides looking at charge trapping during normal transistor operation, one can also stress the gate so that charges are intentionally injected into it. This is called charge pumping. The purpose of doing this is two-fold: (1) you can control the amount of injected charge; and (2) you can see if there is any interface damage due to stress and how the damage affects charge trapping behavior. The damage to the interface can be seen when measuring charge pumping current after each stress.

Reliability Test Instrumentation Trends

As the previous text indicates, reliability tests have evolved to match the needs of new device designs and materials. While HCI is still an important reliability concern, engineers must now worry about NBTI for PMOS, charge trapping for high κ gate transistors, and cross effect between NBTI, TDDB and HCI, such as NBTI enhanced hot carrier, and TDDB enhanced NBTI. To deal with these new phenomena, measurement methodology has evolved from DC stress and measurement to a point that both DC and pulse stress are used to study degradation relaxation effect. Furthermore, instrumentation now includes more comprehensive device characterization suites, which include DC I-V, C-V, charge pumping, and

charge trapping. *Table 1* summarizes some WLR test trends.

Table 1. Recent wafer level reliability test trends.

	Traditional	New
Main Device Degradation Mechanism	HCI	HCI
		NBTI/PBTI
		Charge Trapping
		NBTI-HCI
		TDDB-NBTI
Methodology	DC Stress	Pulse/DC Stress
	DC Measure	Comprehensive Characterization (I-V, C-V, CP, SPCT)

These evolving test requirements are challenging engineers to find the right instrumentation for efficient device and process development. The tool selected should be sensitive enough to capture all the pertinent details of parameter degradation due to stress and flexible enough to adapt to non-traditional WLR tests, such as stress C-V, charge pumping, etc. This tool should also be extendable so that one does not need to buy a completely new system every time a new test issue comes up. Finally, the tool should be easy to use so that one can focus valuable time on interpreting data, not learning to use the test system.

In terms of features, a modern reliability test stand must provide the following:

- Hardware and software that accelerates testing without compromising accuracy and extrapolated lifetimes.
- Semi-auto or auto-prober with a thermal chuck.
- Manipulators or a parallel probe card with low leakage.
- Drivers to control instruments, probers, chucks, create tests, execute tests, and manage data.
- Flexibility to accommodate user-changeable tests and stress sequences for new materials and failure mechanisms.
- Analysis software that provides easy extraction of final lifetime predictions from accelerated short-term tests.

Conclusion

Evolving design scales and new materials are making wafer level reliability testing more critical than ever. This is also driving the demand for reliability testing and modeling much further upstream – especially into

New Instrument Solutions

One example of new instrument solutions for advanced materials and processes is the Keithley Model 4200-SCS Semiconductor Characterization System with wafer level reliability test enhancements. This system allows engineers to easily put different measurement techniques together for timely data collection and is configurable from two to eight SMUs. An optional preamp has 0.1fA resolution. The Model 4200-SCS can also control other instruments, such as a switch matrix, C-V meter, and pulse generator without user programming. This can be done using GPIB, Ethernet, or RS-232. The interactive software has a test plan manager, interactive test setup interface, Excel-like data sheet, easy graphing capability, and more. The Model 4200-SCS also has the flexibility to be used in an interactive manual mode (for single test operation during development) or in more automated production use cases.

The KTEI 5.0 software that comes with the system makes reliability testing quick and easy. Its intuitive point-and-click interface, combined with built-in test routines, makes setting up and running reliability tests as easy as getting I-V curves. The latest version of the system comes with a variety of standard stress-measure tests ready to run on boot-up, which are easily modifiable for customized testing. Its superior sensitivity and low-level measurement capabilities make it ideal for accurately tracking the smallest change in a degrading parameter.

One feature that turns the Model 4200-SCS into an ideal reliability development station is the enhanced wafer level stress-measure loop built into the test plan manager. It includes a stress-measure loop with exit logics and a site loop for stepping through sites on a wafer. By taking advantage of these two loops, one can easily set up a customized wafer level reliability test without any programming. *Figure 8* shows the bundled HCI sample project. The graph shows a particular parameter being tracked over time, with each dot representing a different measure cycle after a stress cycle. On the left is the sequencer showing the order of measurement tests and the overall structure of the project.
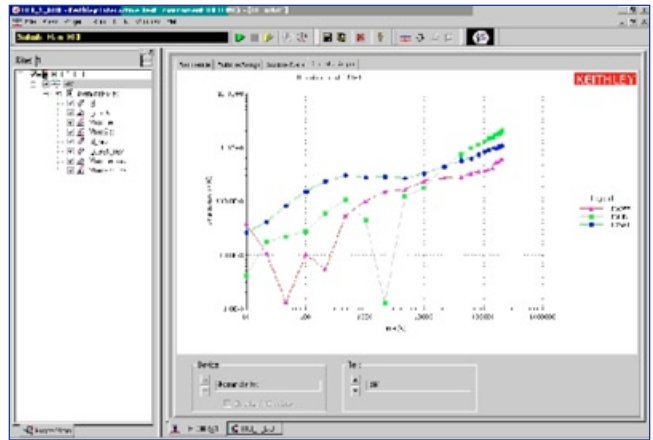

Figure 8. Model 4200-SCS HCI test screen with real-time data plots.

In addition, when performing traditional WLR tests such as HCI, electromigration, and TDDB/QDB, the biggest advantage of the Model 4200-SCS is that it can easily include customized WLR test routines. For example, the voltage waveform shown in *Figure 9* illustrates different test modules, such as C-V, I-V, and charge pumping that can be simply combined in the test plan manager for a looping sequence without programming.
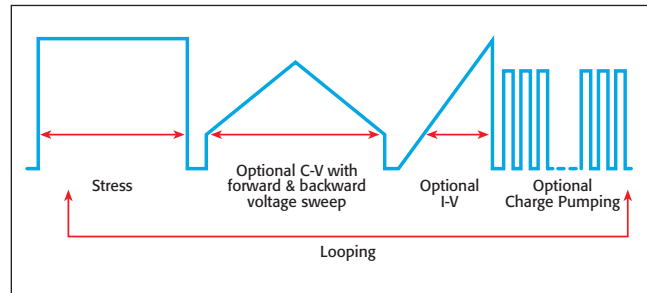

Figure 9. Customized reliability test routine incorporating different stress-measure protocols.

APPENDIX – TRADITIONAL HCI AND NBTI TESTS

Hot Carrier Injection

HCI has been one of the key reliability tests in the last couple of CMOS generations. This is a process where high lateral electrical fields in a MOSFET generate hot carriers (high energy electrons or holes) that can damage the MOS gate oxide interface and degrade the device's I-V characteristics. This phenomenon gets worse as channel length decreases, because the lateral electric field in the channel is a function of gate voltage divided by channel length. As channel lengths have been decreasing proportionally faster than gate voltage, the increases in lateral electrical fields are causing higher energy carriers and more potential damage to the gate oxide. This damage is due to the high kinetic energy of accelerated carriers that produce electron/hole pairs through impact ionization.

Degradation will be seen in the device's I_{DS} (*Figure 10*), transconductance, and threshold voltage. Degradation first slows down the operation of the device and will eventually cause it to stop working all together. The HCI test measures how fast a MOSFET transistor degrades when voltage stress is applied and uses stress conditions to accelerate the degradation for quicker results that can be extrapolated to lifetime predictions under normal operating conditions (see *Figure 2* of the text).
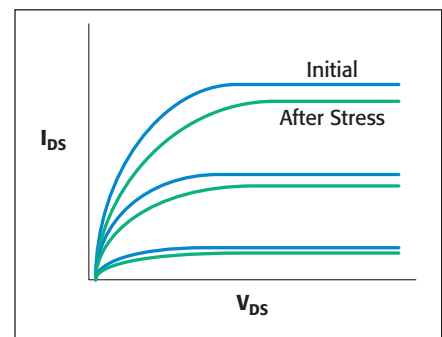

Figure 10. I-V curve showing HCI induced I_{DS} degradation after voltage stress.

the R&D process. Instrument manufacturers are responding with new reliability test tools that are faster, more sensitive, and highly flexible to help drive down the cost of testing and shorten the time to market. KEITHLEY

Negative Temperature Bias Instability

NBTI is a failure mode that is problematic in PMOS transistors and getting worse as threshold voltage continues to drop. NBTI degradation is measured by time dependent shifts in threshold voltages and is associated with slower operation, more leakage, and lower drive current under negative bias stress at high temperature.

The NBTI test is typically a stress-measure sequence loop. During the stress, negative gate bias voltage is applied with the rest of transistor terminals grounded. Between two consecutive stresses, drain current is measured at normal operating condition. Degradation of drain current or threshold voltage is plotted as a function of stress time. All the stress voltages and subsequent measurements are done at high temperature (for example, 135°C).

References

[1]   S. Rangan et al, IEDM 2003
[2]   M. Denais et al, IEDM 2004
[3]   M. Ershov et al, APL 2003
[4]   M. A. Alam et al, IEDM 2003
[5]   G. Chen et al, IRPS 2003
[6]   A. Kerber et al, VLSI 2003
[7]   B Lee et al, IEDM 2004
[8]   Y. Zhao et al, ICSICT 2004
[9]   C. Young et al. SSDM 2004

About the Authors

Dave Rubin joined Keithley Instruments as a Senior Industry Marketing Manager for the Semiconductor Business Group. He was previously a VLSI chip designer at AMD and a business consultant at McKinsey and Co. He has a B.S.E.E. from Arizona State University, and an M.B.A. and MSM-IS from Case Western Reserve University.

Yuegang Zhao is a Lead Applications Engineer with the Semiconductor Business Group of Keithley Instruments, Inc. in Cleveland, Ohio. He received his M.S. in Physics from the University of Wisconsin, Madison, and his B.S. in Physics from Peking University, Beijing, China.

KEITHLEY