DESIGNING MOS INPUTS AND OUTPUTS TO AVOID OXIDE FAILURE IN THE CHARGED DEVICE MODEL

Timothy J. Maloney
Intel Corporation
2250 Mission College Blvd.
Santa Clara, CA 95052
(408) 765-9389

ABSTRACT

MOS integrated circuits most commonly fail a mechanical handling, or charged device, ESD test due to an unanticipated oxide overvoltage during the stress. Reasons include on-chip ground loops, mismatches in the timing of voltage breakdown in critical devices, and direct flow of charge from a power bus to ground through an oxide. Weaknesses can be avoided through proper layout of protection devices and nearby circuitry. This paper develops design guidelines from charge flow analysis, device physics, and failure analysis of circuits under charged device stress.

I. INTRODUCTION

Design principles for MOS IC protection intended to survive a human body model (HBM) ESD test (Mil. Std. 883-C, Method 3015) are fairly well established now, and the same protection circuits (based on the npn bipolar snapback device) usually also survive a charged device model (CDM) test. But the CDM test, to a greater extent than the HBM test, can cause an unexpected failure in a product (usually due to gate oxide leakage) because of the circuit environment in which the standard protection device is placed. In this work we will develop the principles of good design and placement of input and output networks that include ESD protection against charged device failures.

In the charged device model, a component is charged up with respect to an external ground plane and the pin under test is suddenly grounded through a low inductance, dissipating the charge within nanoseconds[1]. The CDM test is a good simulation of ESD stress encountered in mechanical handling of ICs. An oxide overvoltage can arise from a variety of layout and circuit design weaknesses, and this paper will examine several of these:

1) Input buffers can suffer from a lack of proper on-chip ground bus wiring. If the input buffer device ground is returned to the Vss pad by a path different from that of its gate oxide's NMOS protection device, the IR (voltage) drop through the ground bus wiring for the main protection device can boost the input buffer's gate voltage to intolerable levels. Charge from Vss then flows to ground through the input buffer gate oxide, rupturing it.

2) A thin oxide can fail during a CDM event if it is part of a path to ground for a substantial quantity of charge, usually the charge from a power supply bus. This can happen when the charge comes through an avalanching NMOS device of an efficient forward bias path. Also, circuits near a pad often have shorter gate lengths than the output buffer or intended protection circuitry, leading to mismatches in time-to-snapback for various devices. This can cause oxide destruction due to rapid flow of charge from a power bus.

3) An excessive gate oxide voltage can occur due to negative charge accumulated on a node common to the gate of an avalanching NMOS device. The charge appears on the NMOS gate just as in EPROM programming but usually a device other than the avalanching NMOS device is threatened. An example is given in which a p-channel gate oxide is destroyed.

These oxide failures fall into two general categories. One is the case in which the gate node is directly connected to the pad, called a Type I failure. The failures under (1), above, and the breakdown timing failures under (2) are Type I. The other case, Type II, applies to the other cases in (2) and (3) and involves a failing gate oxide between an FET drain on the pad and its gate node. In all cases the failing gate oxide is not more than a diffusion and some series resistance from the pad.

After a brief review of the Intel approach to CDM testing, this paper will discuss various techniques for design and layout of ICs that avoid the above problems. As the IC designer may not always find it possible to use the most aggressive protection devices developed for the charged device model[2] because of constraints in process design rules, circuits, or layout geometries, the emphasis here will be on solutions that lie within a reasonable set of constraints, yet still eliminate mechanical handling failures on the test floor.

II. CHARGED DEVICE MODEL TESTING

An IC component's susceptibility to failure during mechanical handling is indicated by its performance in a CDM simulation test, similar to that discussed in Ref 1. The Intel experience with a simulation test, done with an inexpensive box constructed within the company, has been very positive (see Table I). For the tests in Table I, the inductance to ground through the switch is 50-100 nanohenries, considerably in excess of the package inductance (usually <10 nH[1,3]) in a "real world" mechanical handling event. Nonetheless, for products in Table I and others, the CDM simulation test has a remarkably good record of reproducing failure mechanisms, and of correlating failure voltage with the severity of mechanical handling problems. In every case of noticeable real-world mechanical handling problems, the affected pin has failed a charged device simulation test far short of a 1500V pass/fail voltage. Moreover, adoption (in 1985) of a 1500V CDM test tolerance goal for all new products (and related

10th annual
ELECTRICAL OVERSTRESS/
ELECTROSTATIC DISCHARGE SYMPOSIUM (1988)
SEPTEMBER 27—29, 1988 • ANAHEIM MARRIOTT • ANAHEIM, CALIFORNIA

pp220-227.

layout rules) has led to the virtual elimination of mechanical handling failures for products meeting these requirements. The CDM test has a much better record in this respect than an earlier "fast ESD" test (discontinued at Intel in 1985) that involved the discharge of 600V on 50 pf through 0 ohms to the pin under test. Mechanically sensitive parts did not usually fail this test, and never developed leakage in the proper location.

TABLE I: CHARGED DEVICE TEST RESULTS
(selected 1980-85 products)

PRODUCT	MECHANICAL HANDLING PROBLEMS	CHARGED DEVICE TEST
PRODUCT A	SEVERE (fails at grounded gate-oxi edge)	-600V to -900V (fails at grounded gate-oxi edge)
PRODUCT B	BAD (fails at grounded gate)	-800V to -1300V (fails at grounded gate)
PRODUCT C	BAD	-1000V to -1200V
PRODUCT D	?	-1200V

The test box inductance results in a device waveform that is expected to be slower than in the real-world case, but also the 1500V is likely to be larger than real-world charging. In the simple series R-L-C model of Ref. 1, shown in Figure 1, for the usual case of $R << \sqrt{L/C}$, the current flowing after the switch is closed is given by

$$I(t) = V/\sqrt{L/C} \exp(-Rt/2L) \sin \omega t, \ \omega \approx \sqrt{1/LC}, \quad (1)$$

where V is the charging voltage, L is the ground path inductance (including the switch), R the device resistance and C the total package capacitance with respect to external ground. Thus the peak current is approximately

$$Imax = V/\sqrt{L/C}. \quad (2)$$

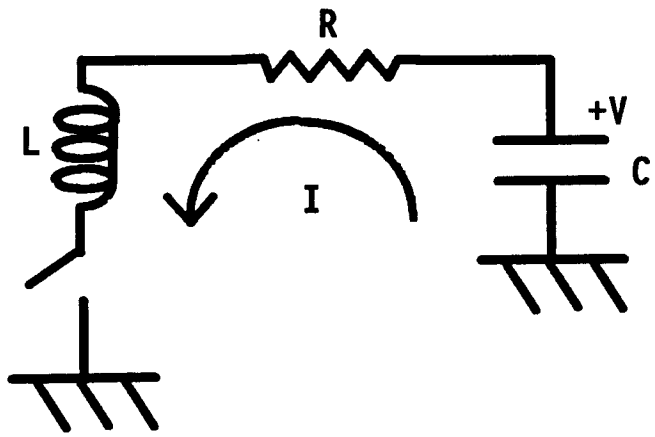

Fig. 1. Series R-L-C circuit model of charged device ESD event.

We have observed quite a few CDM failure voltages going as Imax. For various ground path inductances L, a $\sqrt{L}$ dependence on the failure voltage was observed for some of the designs to be discussed later.

The remainder of this article will be concerned with why oxide damage occurs in the charged device model, and how proper layout can avoid it. CDM simulation tests as discussed in this section were used in the evaluations, and leakage paths were identified by laser cutting or by hot spot detection methods like liquid crystal and emission microscope.

III. ON-CHIP GROUND LOOPS AND OXIDE OVERSTRESS

ESD protection of an input buffer transistor's thin gate oxide can usually be accomplished by clamping the voltage across the gate oxide with a grounded gate NMOS FET of the same kind, as shown in Fig. 2a. As gate oxides become thinner, NMOS drain breakdown voltages fall along with the gate oxide breakdown voltage[4]. Because thinner gate oxides are always planned for narrower gate lengths, devices break down faster, so that when the gate length of the thin gate (TG) protection device is close to minimum, NMOS snapback[5] is fast enough[6,7] to protect the input buffer gate oxides from damage. For example, while up to 4-5 micron gate lengths are adequate for protecting a 400Å gate oxide, 2-3 micron lengths must be used for equivalent protection of 250Å gate oxides. There is more about this subject in Section IV.

The scheme shown in Fig. 2 is for protection of an input buffer when the pin is intended for use as an input-only. The main protection device on the pad is a thick field oxide (TFO) device, which has been described in many references[8-10]. The TFO is most appropriate for NMOS or N-well CMOS, where it is in the p-substrate, not a well. It operates as a diode clamp in one direction (negative Mil Spec voltage, positive charged device voltage) and as an npn snapback device in the opposite direction. The metal gate connection on the TFO is not of great consequence[10]. The rest of the protection "pi network" is a series resistor (typically a few hundred ohms) and the aforementioned grounded gate NMOS device to clamp the gate oxide voltage on TNin and TPin. The series resistor limits the breakdown current through TG so that TG can be a fairly small device. Augmenting a TFO protection device in this way is not a new concept; Ref. 10 describes how adding the rest of the pi network to TFO was necessary for obtaining full ESD protection.

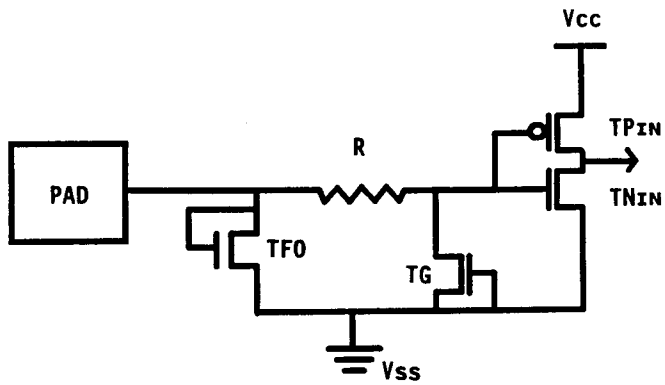

Fig. 2a. Protection scheme for input-only pin.

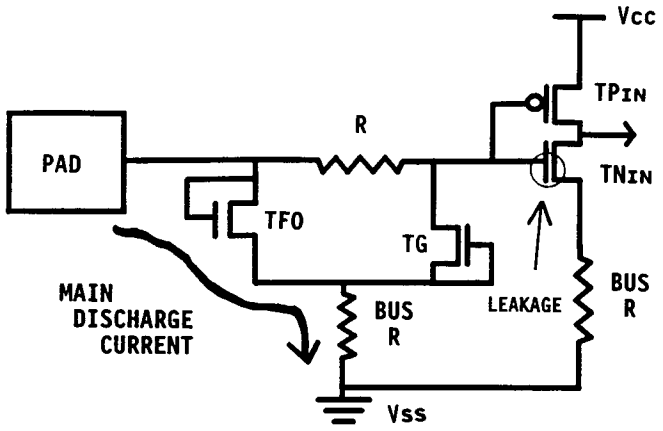

Fig. 2b. Deficient input protection scheme, with ground bus wiring resistances shown.

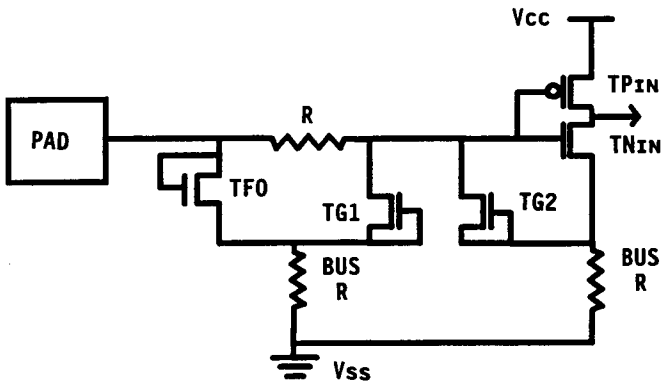

Fig. 2c. Improved input protection scheme, not affected by ground bus resistance.

The same sort of pi network scheme can be used for an input/output (I/O) pin. The TFO device is simply replaced by an output buffer, for which a self-protecting NMOS pulldown is the principal ESD protection device. Special design considerations apply for p-well CMOS, where NMOS devices are in a p-well[11].

The ESD protection scheme as shown in Fig. 2a is an idealized one that does not recognize ground bus wiring resistances as shown in Fig. 2b. Circuit designers will often separate power and ground buses on internal circuitry from "noisy" pad driver buses, so that ground paths are returned to a common point (the Vss pad in this case) as shown in Fig. 2b. Sometimes this common point is on a lead frame and bond wire inductance also appears on each separate path. Fig. 2b represents the most convenient (but faulty) input-only wiring scheme, where the protection elements are all in a cell near the pad and referred to pad driver ground for convenience, with no cross-coupling of the ground bus lines. Now imagine a CDM event with negative chargeup (-CDM). Once in snapback, most discharge current will flow through the TFO device, as it has the lowest resistance to ground. But as CDM events often have peak currents of over 10 amps[1,7], it is easy to see that metal resistance in the ground path (typically 25-40 milliohms/square) can easily cause an extra IR drop sufficient to blow out the oxide of TNin. Sometimes even a positive human body model (+HBM) zap with respect to Vss will have the same result, and of course in this case, the direction of charge flow is the same. The destruction of the TNin gate oxide is the most common kind of Type I failure.

To avoid these difficulties with ground bus wiring, the scheme shown in Fig. 2c is used. The pi network cell can remain undisturbed as long as TG2 is added at the input buffer, sharing a common ground connection with it.

The pi network cell can be arranged with a merged source (emitter), as shown in Fig. 3, so that it is not only compact, but also triggers at the NMOS snapback voltage. The NMOS device TG, current-limited by R, breaks down and forward biases the Vss junction common with TFO, thus triggering TFO immediately.

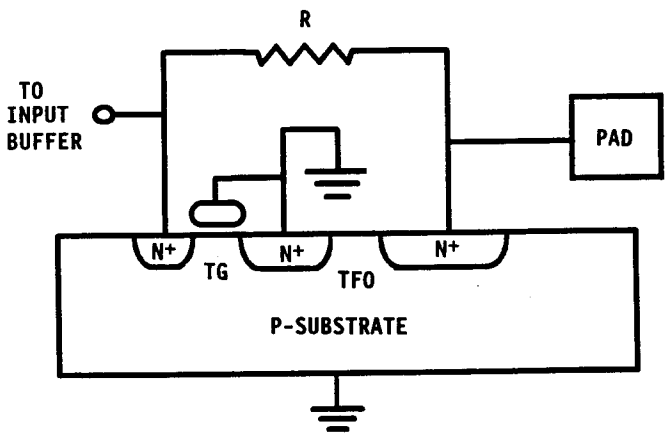

Fig. 3. Input protection network with a merged source (ground). TG's npn snapback triggers TFO snapback.

The above discussion takes a preference to the -CDM (or +HBM) test, and in the CDM case is not specific about where the actual "common" ground is located. There are reasons for this. As they throw the protection device into the weaker reverse breakdown mode[6], the -CDM and +HBM tests almost universally fail parts at lower voltages than their opposites, reported also by Ref. 7. (Notably, those test polarities, with the same direction of charge flow, have the greatest application to the real world, where for triboelectric reasons, components tend to acquire negative charge and people (clothing) tend to acquire positive charge.) We will drop further discussion of the HBM test here since the IR drop argument is very clear when the ground reference is the Vss pad.

The -CDM test causes snapback in the TFO device, which then passes Vss charge to the external earth ground. The connection with substrate charge is made through p+ substrate taps (in CMOS) and forward bias resistance, the latter depending on voltage polarity and time. This complicates the notion of a common Vss point for TG and TNin sources; while a common metal link may be through the Vss pad, there is also a parallel path through the substrate. Resistance to a common substrate point involves distributed substrate tap spreading resistance as well as time-dependent forward-bias resistance for the internal Vss line. Substrate connections may play an important role, but the common connection of TG2 and TNin sources should be accomplished with metal, not with substrate taps.

It is interesting to compare CDM results in the two flavors of CMOS (p-well and n-well) as well as NMOS. Table II summarizes the general trends of this experience. The test to watch is -CDM for NMOS (see Ref. 7) and N-well CMOS. The +CDM test in these cases has immediate forward-bias access to substrate charge[7], as can be seen in Fig. 4a, and few problems with this polarity are observed. Our experience with p-well CMOS CDM testing has been limited to mostly outputs and I/Os, but not many problems have been observed. The good -CDM performance in output networks for p-well CMOS is again explained by forward-bias access to substrate charge, this time through the p-channel output device. +CDM in p-well CMOS can trigger the vertical npn device through the well to gain access to substrate charge. The vertical npn is expected to be quite rugged when the collector is the well-substrate junction[11].

TABLE II: GENERAL TRENDS OF CHARGED DEVICE MODEL TESTING

	-CDM	+CDM	SUBSTRATE CHARGE
P-WELL CMOS			
INPUT	-----	---	
OUTPUT	OK	OK	Vcc
N-WELL CMOS	sensitive	OK	Vss
NMOS	sensitive	OK	Vss (through bond wires)

NMOS and n-well CMOS equivalent circuits for the CDM are compared in Figs. 4a and 4b. An input-only pin is shown for NMOS in Fig. 4a. The connection between Vss and substrate is made only through bond wire inductance, as shown, and is even more tenuous if there is a substrate bias generator. For n-well CMOS, an I/O pin is shown because external Vss may sometimes have no substrate taps and be separately bonded out, as it is also used to carry noise spikes from output inverters. An input-only in N-well CMOS will typically have the TFO device on substrate tap ground, to give it direct access to substrate charge. Some diodes resulting from n+ to substrate junctions are also pictured in Fig. 4. But even these circuit models are not complete; for example, power bus resistance, the distributed R-C character of power buses, and the distributed L-C character of the capacitance and inductance at the pin, is not shown. Also, there is unsatisfactory isolation of the substrate charge in Fig. 4b, which is remedied by an extra TG device, to be discussed in the next section. Circuit simulations of the charged device event, using expanded versions of the kind of circuit models shown in Fig. 4, are a topic of current research.

This section has presented the first example of what could be called charge flow analysis of CDM events. One must visualize charge being stored on all circuit nodes, charge which then flows to external earth ground by the shortest path as the switch is closed. A knowledge of device physics and of electrical properties of materials used in silicon ICs is then used to locate possible weak points. The next section will deal with this further.

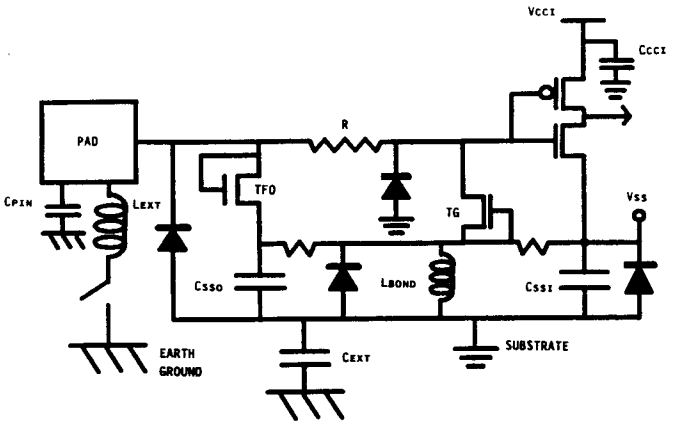

Fig. 4a. Charged device model equivalent circuit for NMOS input-only pin.

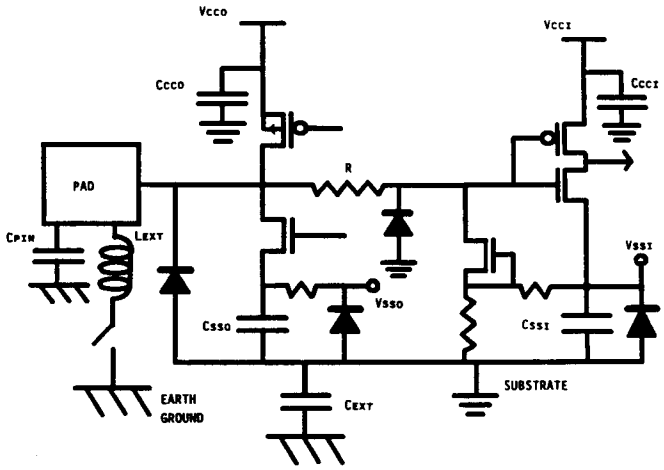

Fig. 4b. Charged device model equivalent circuit for n-well CMOS input/output pin.

IV. POWER BUS ISOLATION AND BREAKDOWN TIMING CONSIDERATIONS

An analysis of any charged device model failure should begin with a consideration of where charge is initially stored and how it flows to the grounded pin. Capacitances are of crucial importance. The capacitances shown in Fig. 4 are device to external ground capacitance, Cext, pin capacitance Cpin, and Vss and Vcc power bus capacitances Cssi, Csso, Ccci and Ccco. Cext depends on external geometric conditions, with 10-15 picofarads (pf) a typical value. While power bus capacitances with respect to substrate can range from tens to hundreds of pf and are almost always greater than Cext, note that they are in series with Cext in our simplified model. Clearly there are also capacitances from external ground to each node but these tend to be small compared to package and substrate capacitances, so we approximate by lumping them into Cext. In the discussion that follows, the pin capacitance Cpin will also be considered to be lumped into Cext. Cext supports the charging voltage and determines the energy to be dissipated when the switch is closed, but the charge flow pattern through the circuit determines whether or not the charge will be exhausted safely.

IV.A Separated Power Bus Lines

Now we turn to an example of Type II oxide failures, where the drains of PU and PD are on the pad and their gate oxides are failing. Fig. 5 is a simplified schematic of an N-well CMOS circuit failing the charge device test due to power bus isolation. The pullup (PU) and pulldown (PD) transistors are referred to Vcco and Vsso (the external power buses), while they are driven by gates referred to Vcci and Vssi, the internal power buses. P-substrate taps are on Vssix and are common with Vssi through on-chip metal, but are in a separate loop. Vssi and Vsso are bonded to separate pins while Vcci and Vcco are connected through bond wires to a single pin. The 100/2.2 FET called TG, grounded to Vssix through polysilicide, breaks down and provides a slightly resistive path to the pad for charge stored on the substrate and on Vssi.

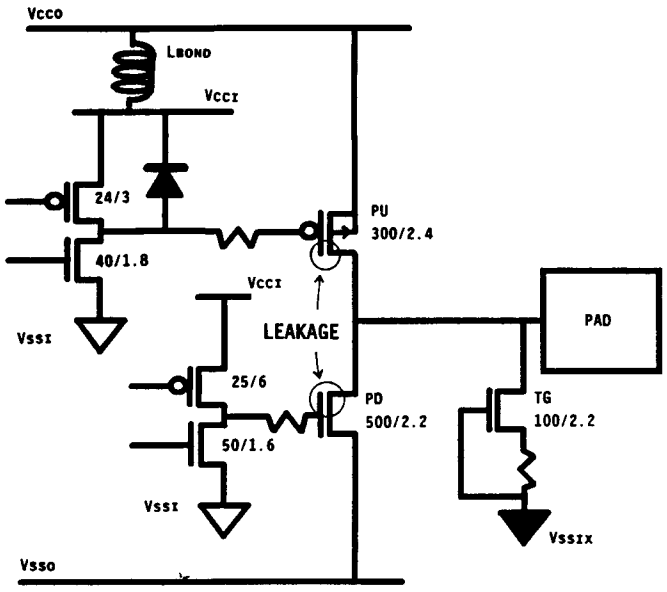

Fig. 5. N-well CMOS output circuit. PU and PD gates suffer in -CDM test. Vssix is substrate-tapped and is common with Vssi through a long metal connection.

The output pin fails the charged device test for -1400V, and leakage is observed (with liquid crystal hot spot detection) from gate to drain of both the PU and PD devices. The equivalent diode of one p-channel gate driver device is drawn in Fig. 5, to express how negative charge would flow from Vcc to pad, passing through the PU gate oxide. Both PU and PD gate oxides have ready access to Vssi and substrate charge through the NMOS pulldown devices on Vssi. As the devices on the pad must have longer gate lengths (here, 2.2μm for NMOS) to keep down subthreshold currents, the internal inverters provide faster access to the Vssi charge than TG because of the $Le^2/2D$ time dependence for breakdown[6,7], Le the gate length and D the carrier diffusion coefficient. The resistance to the substrate of TG is also important. Vcci negative charge can come through the PU diode, with resistance and time dependence based on how close the p+ drain contact is to a well tap, but also can flow (again with resistance) from well to substrate through the np diode there. Vcci and Vcco are connected by bond wires, as we said before, and there is another np diode to the pad through PU. Stored charge thus finds its way to external ground by a variety of paths, some destructive (as through PU and PD gate oxides) and some harmless (as through TG or PU and PD source-drain.

The most common failure location in output networks like Fig. 5 is the gate oxide of PD, the n-channel device, although failures of both PU and PD gate oxides have been occasionally observed in tandem. For a network like Fig. 5, charge coming through an efficient np diode from Vcci is of some importance but charge on Vssi, coming through very fast grounded gate NMOS devices, is of greater importance here, and primarily determines the failure voltage. There have been a few cases in which charge from Vcci, coming through an efficient np diode, has been primarily responsible for gate oxide failure of a small p-channel pullup device on the pad.

The stored charge on Cext amounts to 15 nanocoulombs (nC) in the case of 1500V charging voltage and Cext=10 pf. This is quite large compared to any charge stored locally near the pin under test, such as on nodes driving the output inverters, because Cext represents all the charge on the substrate, die cavity, package pins and other package parts. Moreover, the breakdown charge for gate oxides is typically in the range of 0.1 to 1.0 coulomb/cm^2, meaning that because gates are always more than a few square microns in size, something on the order of the Cext charge is required to destroy the gate. However, as suggested by some time-dependent oxide breakdown data[12], the breakdown charge required for the nanosecond time scale of the charged device event may be less than the above quantities, usually measured on the scale of milliseconds to seconds. This explains why gates with hundreds of square microns can be destroyed by -1500V. The principal argument of this section is that if the external power bus Vsso is sufficiently isolated from the internal power buses and substrate, large chunks of the Cext charge can flow through an internal power bus and then through a gate oxide on an output buffer, destroying it. The right design strategy can help shunt the Cext charge harmlessly.

Devices like TG in Fig. 5 help to provide a direct harmless path to ground for power bus and substrate charge, and raise the failure voltage of the pin under test. Its effectiveness was shown when TG was removed by laser cutting the drain and the failure voltage dropped from -2000V to -1400V.

Another obvious strategy for preventing gate oxide damage due to power bus charge is to prevent large voltage swings on the internal nodes by clamping the power buses with respect to one another, using NMOS FET triggered protection devices as in Fig. 3. Several such sets of clamps (Vcc to Vss for both sets of power buses) were applied to a developmental product with outputs designed as in Fig. 5, with the effect of raising the failure voltage from less than -1500V to about -2000V. Performance varies with position along the bus lines, as would be expected, and considerations of power bus and well tap resistance do affect placement of the clamps.

A final strategy for preventing gate oxide damage for circuits such as in Fig. 5, is to attach small grounded gate NMOS devices to the PU and PD gates, referring the ground to Vsso. This limits the gate voltage that can build up due to the isolation of Vsso from the substrate. Sometimes there are layout problems with such a scheme and other strategies are preferred.

IV.B Breakdown Timing As A Cause of Oxide Destruction

The discussion in the previous section pointed out how the gate length of a grounded gate protection device governs how fast snapback and voltage clamping will be achieved[6,7]. In some cases, this protection device gate length is crucial, although oxide thickness and the gate lengths of other nearby devices are usually part of the analysis. An example of such a case is shown in Fig. 6, where a Fig. 3-like input protection device protects the gates of a stacked-gate input buffer in another developmental product. Vss metal and substrate taps tie all the circuits to a local ground reference, so there are no grounding problems as discussed above. But well short of -1500V in the charged device test, the 42/2.4 stacked NMOS device develops a gate-source short as shown, which is traceable to the 5μm gate length of the 40/5 grounded gate device that triggers TFO (see Fig. 3). The 40/5 device breaks down more slowly than the 40/2 NMOS device at the bottom of the input buffer stack; consequently substrate charge has a chance to pass through the 40/2 device and the vulnerable gate oxide (250Å thick) before the protection device diverts it through a harmless path. The stacked gate is not essential to this failure; it is a Type I failure occurring for different reasons than the ground loops in Section III. In this case, emission microscopy was able to pinpoint the oxide leakage location as gate-to-source on the 40/2 NMOS device.

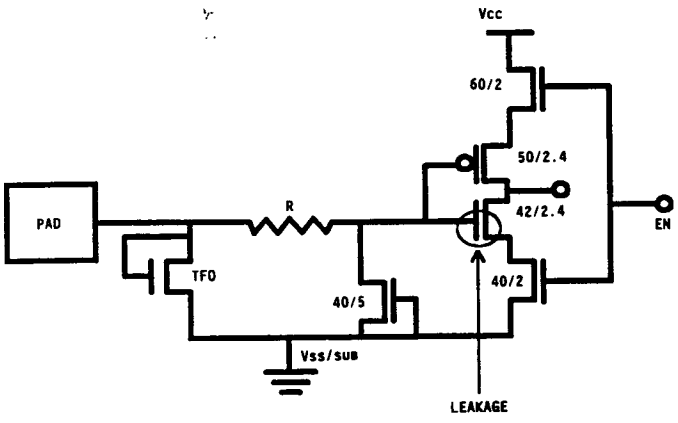

Fig. 6a. Stacked gate n-well CMOS input-only with failure in -CDM test caused by mismatch in NMOS breakdown timing. Leakage location is noted.

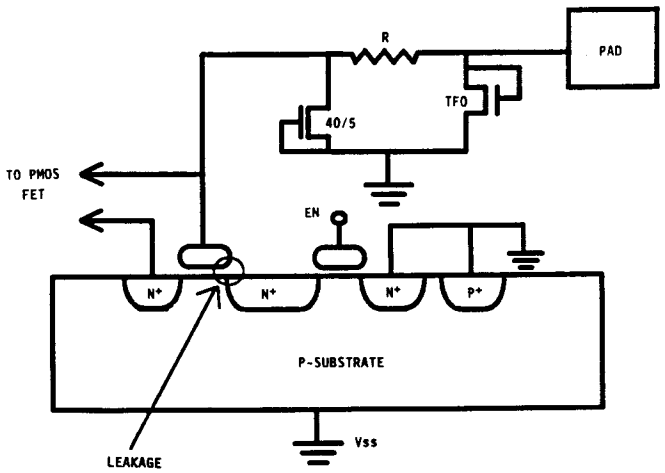

Fig. 6b. Partial cross-sectional view of stacked gate input-only; leakage location is again noted.

In the next stage of development of this product, the grounded gate length was shortened to 2μm and no further problem with the charged device model was seen. The 5μm length was a holdover from a 400Å gate technology, where no problem with the charged device ESD test had been seen. These 250Å gates were, as expected, more susceptible to the overvoltage spike that takes place during the protection device turn-on time, but a 2μm gate length on the protection device was sufficient to remove charged device test sensitivities.

V. OXIDE DESTRUCTION BY EPROM-LIKE GATE CHARGING

Our final example qualifies for the charged device ESD category because there were some slight problems observed, for negative chargeup, but in this case there also was a HBM ESD failure below +2000V, with leakage in the same location. Consequently the testing and failure analysis concentrated on the more widely accepted HBM test.

The circuit in Fig. 7 includes a weak 17/5 micron PMOS pullup device, tied to the same inverter that drives the large NMOS PD gate. A positive Mil Spec or negative charged device zap will cause gate-drain leakage (Type II) on the weak p-pullup, but not if the link from the PD gate to the weak pullup is broken at points A or B. Note that a break at point A still allows the inverter to be connected to the weak PMOS gate, yet it doesn't fail. The weak PMOS gate fails because of a direct link to the NMOS PD gate, and the reason is that in breakdown, the NMOS gate acquires negative charge, much as an EPROM floating gate would. This depresses the voltage of the weak PMOS gate so that gate-drain oxide breakdown takes place. While the parasitic diodes of the inverter will have some limiting effect on the gate voltage, they may not be fast-acting due to layout.

Countermeasures against such a failure of the weak PMOS device all seek to reduce the gate-drain voltage during the zap. A separate inverter for the weak pullup is a simple solution, but may cost too much real estate. A more efficient n+p+ diode in parallel with the inverter's diode to the substrate can speed up gate voltage clamping, while a

resistor in series with the weak PMOS drain (not a problem because the weak pullup is intended as a gated resistor) and closer spacing of the well tap to the PMOS drain will reduce the PMOS drain voltage.

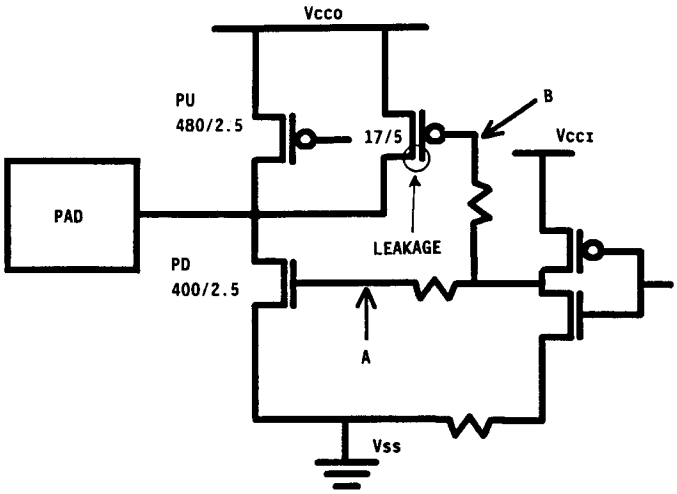

Fig. 7. N-well CMOS output circuit with +HBM and -CDM failure in weak PMOS pullup gate-drain due to charge accumulated on the gate common to the NMOS pulldown.

VI. GENERAL DISCUSSION: DEVICE CONFIGURATIONS AND OXIDE SENSITIVITY

The reader may ask why the gate oxide of the NMOS PD device in Fig. 7 is not also threatened in this event, since its gate-drain voltage is the same as that of the weak PMOS pullup, except for the IR drop in the drain-gate diffusion region. The same question can be asked about the 40/5 grounded gate device in the example of Section IV.B (Fig. 6). The voltage across that device's grounded gate to drain should actually be larger than the source-gate voltage of the 42/2.4 device that develops an oxide short. What's going on? Basically, high voltages to forward-biased junctions are more dangerous than to reversed-biased junctions. Both of the weak oxides mentioned in Sections IV.B and V are cases of a large voltage from the date to a forward-biased FET junction, leaving an abundance of carriers to be attracted to the gate. Notice that the "threatened" NMOS devices in each case had similar voltages from gate to a carrier-poor reversed-biased drain. Figure 8 illustrates the device configuration for each of these cases.

Notice also that all the Type I failures in Section III, due to on-chip ground loops, fit this model as well by failing from a gate to a forward-biased source junction! We should mention that the lowest failure voltages ever observed have been Type I failures as in Section III. Here the charge on internal Vss and substrate can flow directly to the TNin source and through the gate oxide. The other cases discussed here require substrate charge to flow through an npn device, pn diode or other obstacle before reaching the gate oxide.

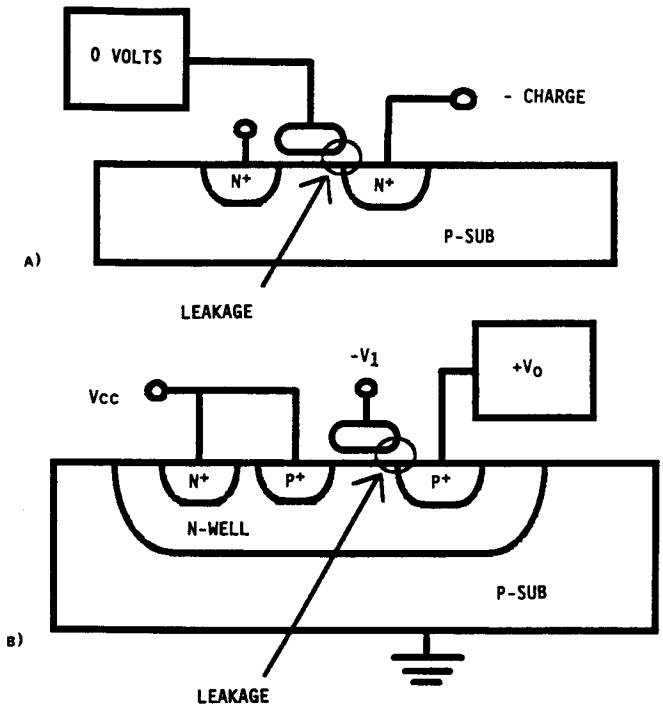

Fig. 8. Cross-sectional illustration of failing oxides from a) Fig. 6 and b) Fig. 7. In each case the gate breaks down to a forward-biased FET junction.

As a practical matter, the gate oxides of NMOS devices that go into snapback do not generally give us trouble in the charged device ESD model unless their gate voltage drops far below the source voltage, as was the case for the PD failures discussed in Section IV.A. (Most tests have been for oxides 250Å thick or more.) It is useful to look at the pre- and post-snapback simulations of field pattern and current density in Ref. 5 to get a feel for how the NMOS snapback event forces the current subsurface and relieves the intense concentration of carriers at the silicon-oxide interface. In this way the NMOS snapback event appears to relieve some of the stress applied to its own thin oxide, especially on the forward-biased side. The two unfortunate gates in Figs. 6-8 could not benefit because the 42/2.4 gate was not able to go into snapback, and the weak pullup was PMOS (with no snapback mode[5]), not NMOS. The lesson for failing NMOS devices that do avalanche during the zap is that the gate voltage is at fault, as in Section IV.A, and that limiting the gate voltage should be sufficient to cure the problem.

VII. CONCLUSIONS

Mechanical handling failures on the manufacturing test floor can be anticipated with a simulation of the charged device ESD test that reproduces the "real world" failure mechanism and location with remarkable accuracy, even when the simulator's ground inductance is 50-100 nH and considerably larger than package inductance. Failure thresholds often scale with maximum current, allowing simple R-L-C theory to be applied.

Gate oxides that fail in the charged device test are always near a pad, and either the gate (Type I) or drain (Type II) connection may lead to the pad. Analysis of charged device test failures has led to some general design principles for avoiding oxide damage. These failure mechanisms and design principles are summarized in the Introduction. Some design strategies are aimed directly at limiting the voltage across vulnerable gates. Most others indirectly limit the gate voltage through layout and voltage clamping of power and ground buses. As most of the package charge flows through these buses, there is the distinct possibility of excessive voltage excursions, which of course threaten thin gate oxides. Proper layout will not happen by accident, because many things that circuit designers like for other reasons (such as separated internal and external power buses) must be handled very carefully to avoid charged device ESD damage.

A novel failure mechanism reported here was related to EPROM-like charging of the gates of avalanching NMOS devices, with damage occurring only to a PMOS device with its gate connected to the NMOS gate. This led to a discussion of oxide sensitivities as they relate to all the voltages applied to a device, and it is clear that there is much room for further work. Most of the oxide failures discussed in this paper have a large voltage between the gate and a forward-biased FET junction. An NMOS device in avalanche will not generally give problems in the charged device model unless the gate voltage is excessive, as has happened with output devices on a separate power bus.

As stated above, further work is needed on the issue of oxide sensitivity as it relates to devices and configurations. NMOS vs. PMOS, gate to forward-biased vs. reverse-biased FET junctions, and gate length dependence all need to be studied more. Some published ESD-related device and oxide studies[4,12,13], provide a good starting point. Further work on circuit simulation of CDM events in VLSI products is desirable, and should expand on the circuit models illustrated herein to include distributed R-C and L-C effects in the device and package, as well as charge flow through the oxide. Finally, the phenomena discussed herein should be studied for the latest technologies, including thin gate oxides below 250Å and salicided diffusions, as there were too few CDM data on these at present to determine their possible influence.

ACKNOWLEDGEMENTS

The author thanks Bruce Euzent for managerial support and Edith Coon for manuscript preparation. Many of the concepts expressed in this paper could not have been formulated without a great deal of test data and failure analysis from many Intel people. Major contributors of these results include Ted Krings, Paul Churilla, Hamid Rangchi, Rick Shoemaker, Ron Hansen, Quan Bui, Tina Cantarero, and Kutlu Aricanli.

REFERENCES

1. P. Bossard, R. Chemelli, and B. Unger, "ESD Damage From Triboelectrically Charged Pins", EOS/ESD Symposium Proceedings, 1980, pp. 17-22.

2. L. Avery, "ESD Protection Structures to Survive the Charged Device Model (CDM)", EOS/ESD Symposium Proceedings, 1987, pp. 186-191.

3. L. Avery, "Charged Device Model Testing: Trying to Duplicate Reality", EOS/ESD Symposium Proceedings, 1987, pp. 88-92.

4. W-S. Feng, T. Y. Chan, and C. Hu, "MOSFET Drain Breakdown Voltage", IEEE Elec. Dev. Lett. EDL-7, pp. 449-450, (1986).

5. T. Toyabe, et. al., "A Numerical Model of Avalanche Breakdown in MOSFETs", IEEE Trans. Elec. Dev. ED-25, 825-832 (1978).

6. N. Khurana, T. Maloney, W. Yeh, "ESD on CHMOS Devices: Equivalent Circuits, Physical Models and Failure Mechanisms", Proceedings of IRPS, 1985, pp. 212-223.

7. L. DeChiaro, S. Vaidya, and R. Chemelli, "Input ESD Protection Networks for Fineline NMOS—Effects of Stressing Waveform and Circuit Layout", Proceedings of IRPS, 1986, pp. 206-214.

8. T. Hulett, "On Chip Protection of High Density NMOS Devices", EOS/ESD Symposium Proceedings, 1981, pp. 90-96.

9. R. Rountree and C. Hutchins, "NMOS Protection Circuitry", IEEE Trans. Elec. Dev. ED-32, 910-917 (1985).

10. R. Wilcox and R. Doucette, "The Elimination of ESD Failures from Silicon Gate Logic Technologies", EOS/ESD Symposium Proceedings, 1985, pp. 55-58.

11. T. Maloney, "Contact Injection: A Major Cause of ESD Failure in Integrated Circuits", EOS/ESD Symposium Proceedings, 1986, pp. 166-172.

12. E. Harari, "Dielectric Breakdown in Electrically Stressed Thin Films of Thermal Si02", J. Appl. Phys. 49, 2478-2489 (1978).

13. Y. Fong and C. Hu, "The Effects of High Electric Field Transients on Thin Gate Oxide MOSFETS", EOS/ESD Symposium Proceedings, 1987, pp. 252-257.