An active MOSFET Rail Clamp Network for Component and System Level Protection

Michael Stockinger[1], Wenzhong Zhang[2], Kristen Mason[1] and James Feddeler[1]

[1)] Freescale Semiconductor, 6501 William Cannon Dr. West, Austin, TX 78735 USA

tel.: 512-895-8507, e-mail: m.stockinger@freescale.com

[2)] Freescale Semiconductor, High-tech Park F3, Ste. A1, No. 80 Fourth Ave., Teda, Tianjin 300457 P.R. China

Abstract **-** An on-chip protection network is introduced providing high immunity to component and system level stress. It uses active MOSFET rail clamps with novel RC triggering and clamp layout schemes.

I. Introduction

RC-triggered active MOSFET rail clamps offer effective solutions for component level (e.g. CDM and HBM) protection and are therefore widely used in advanced CMOS technologies. However, under system level stress, as described in the IEC standard 61000-4, they may suffer from triggering issues and poor control of clamp operation (turn on time and intensity) [1].

The protection design introduced in this work overcomes these issues making active MOSFET clamps an attractive alternative solution to SCRs for on-chip system level protection. Typical applications include harsh system environments like household appliances (washer and dryers), thermostats, power meters and motor controls where ESD and other transient stress events (e.g. due to power switching or motor noise) are likely to occur. These are typically low-cost applications with minimum board-level protection components.

II. The Protection Scheme

The protection network presented here was designed for CMOS microcontroller applications where the IO pads support an LCD-tolerant scheme, i.e. the IO pads can be driven to one of two independent supply levels, a main IO supply (VDD) or an LCD supply (VLL3). To support this, a dual-diode pad protection scheme with a dedicated ESD bus (a.k.a. “ESD ghost rail”) was used, as shown in Fig. 1. This protection network is similar to [2,3] using boosted and distributed NMOS rail clamps shunting current from the esd_bus to the vss bus during an overstress event.

A) IO Pads

Large dual diodes A1 and B provide the main protective current paths from the IO pad to the esd_bus and from vss to the IO pad, respectively. The IO cell contains an NMOS clamp M1. By making the A1 diode point to the esd_bus instead of the vdd rail current injection from the IO pad onto VDD is

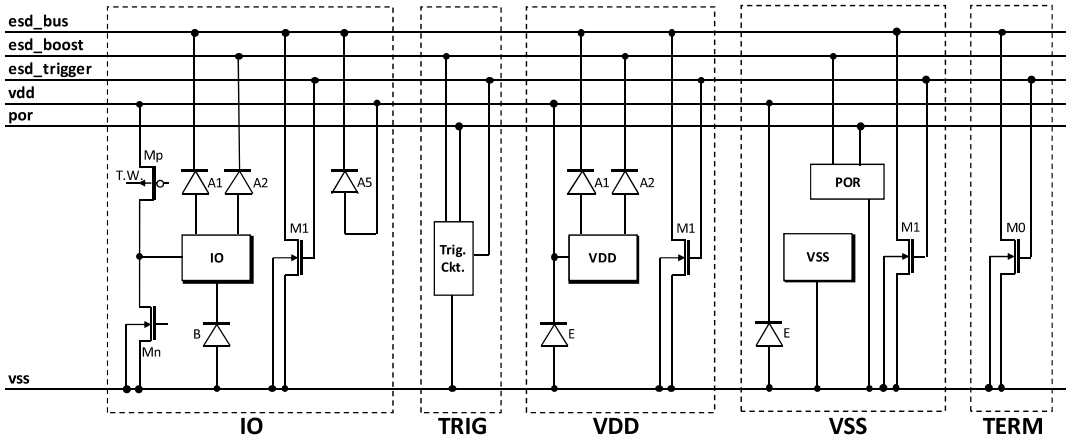

Fig. 1: Conceptual protection network.

avoided. This allows the IO pad voltage to rise above VDD, e.g. when driven to VLL3 by the LCD driver (not shown in Fig. 1). This enables the LCD-tolerant IO pad scheme mentioned earlier. A smaller diode A2 pointing from the pad to the esd_boost bus supplies power to the trigger circuit (TC) and enables the boosted rail clamp scheme [2]. The TC controls the Gate of the clamps via an esd_trigger bus.

The main output drivers Mn and Mp are also shown in Fig. 1. To avoid a parasitic diode (the Drain-Body junction of Mp) pointing from the IO pad to VDD, the PMOS driver Mp is implemented in a tracking Nwell biased at the higher of the IO pad or vdd bus voltages using an active switching circuit (not shown).

B) Supply Pads

The VDD pad protection shown in Fig. 1, with diodes A1, E and A2 and clamp M1, is similar to the IO pad protection. A small additional diode A5, distributed via the IO cells in the pad ring, works in parallel with the main diode A1 of the VDD cell. Without the diode A5, a positive zap on VDD may cause excessive voltage on the vdd bus at pad ring locations that are relatively far away from a VDD pad cell. Unlike traditional rail clamp networks where the vdd bus gets clamped directly to the vss bus, VDD is protected via diodes A1 and A5 to the esd_bus. These diodes must be distributed over the entire pad ring to limit the peak voltage occurring locally between the vdd and vss buses.

The diode E provides a discharge path for negative zaps on VDD. Similar to the distributed diode A5, a distributed parasitic Nwell-Psubstrate diode (not shown) pointing from the vss bus to the vdd bus is present in every IO cell. This diode is in parallel to diode E and avoids excessive negative voltage excursions on the vdd bus relative to the vss bus.

The VSS cell of Fig. 1 is equipped with a power-on reset (POR) circuit signaling the TC via the por bus whether the chip is in a steady powered-up state or not. The TC behavior will be influenced by the por signal. Both circuits will be described later.

C) Segment Termination

The TERM cell is needed for esd_bus segment termination. Its clamp M0 is typically many times larger than the clamp M1 of an IO cell. The clamp M0 can be designed using known termination techniques [2,4]. However, esd_bus segment termination may only be needed in special cases, e.g. for pad rings with significant gaps or between VDD domains with a strict noise isolation requirement. Multiple independent VDD domains may otherwise share one esd_bus - only the vdd bus needs to be interrupted. In such a case, the VDD supply with the highest voltage level determines the voltage level on the shared esd_bus and also carries the total leakage current of all the distributed rail clamps combined.

Fig. 1 also illustrates the main cell types needed in an IO pad library - IO, TRIG, VDD, VSS and TERM cells. TRIG cells must be placed frequently in a pad segment because each TC can only drive a certain maximum number of clamps. A single VDD/VSS cell pair may be sufficient for this protection scheme, but other considerations (e.g. simultaneous switching, bus voltage drop, radiated emissions) may require more supply cells. In such a case, multiple POR circuits can operate in parallel without interference.

III. The New Triggering Method

TCs for active MOSFET clamps are typically designed as slew rate detectors (e.g. an RC filter driving a chain of inverters) that engage on a fast voltage ramp starting at 0V [2]. While this works well for un-powered (component level) stress, e.g. HBM or CDM, it may cause issues for powered (system level, IEC 61000-4) stress, e.g. Powered ESD (PESD), Electrical Fast Transient (EFT) or power surge [1].

A conventional TC (e.g. an RC filter with two inverter stages, like the one shown in Fig. 2a but with constant resistance R) usually changes its detection behavior drastically when powered up. A significant supply voltage jump (e.g. 100%) may be needed to flip the first inverter stage of the TC due to the existing pre-charge on the capacitor of the RC filter. This high voltage may lead to damage during powered stress. Furthermore, a conventional TC may not turn on long enough to dissipate e.g. a power surge event, which can last much longer than e.g. an HBM event.

Some advanced TC designs act like a latch turning the clamps on as hard as possible for a specific duration, e.g. using an on-time control circuit [2]. Their on-time is usually designed to outlast the longest possible ESD event, e.g. HBM. Such a “latching” TC may upset microcontroller operation (causing resets or code failures) due to a collapsing VDD supply.

Using a voltage threshold instead of a slew rate

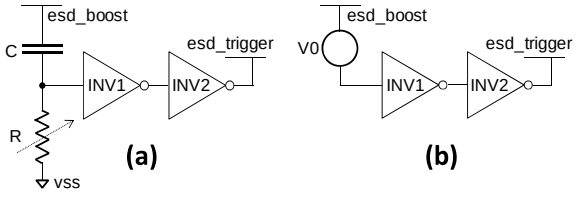

Fig. 2: Simplified trigger circuit schematics: (a) With variable RC stage, (b) illustrating regulation mode.

detection scheme would theoretically solve the first issue with the poor powered stress detection, but may still suffer from the second issue causing a collapsing supply when triggered. Such an alternate detection scheme would also require a stable voltage reference device, e.g. a Zener diode, which may not be readily available in a given CMOS process technology.

This work introduces a triggering scheme well suited for both powered and un-powered overstress. While maintaining the response of a conventional TC during un-powered stress, a new triggering mode is used for powered stress. This mode activates the clamps just enough to limit the voltage increase on the zapped pad, but does not cause a collapsing supply. In other words, the conductivity of the clamps is modulated in proportion to the magnitude of the applied stress. Such a dynamic overvoltage regulation scheme does not require any voltage reference (e.g. Zener diode) because the observed average supply voltage just before the stress event is used as a dynamic reference.

A) Powered versus Unpowered Stress Events

The new TC functionality is achieved by modifying the RC filter stage of a conventional TC, as illustrated in Fig. 2a. The R value is varied based upon two parameters: the power-up state of the chip (as indicated by the por bus) and the triggering state (as indicated by the esd_trigger bus). Table 1 shows an example of RC time values as implemented on a real chip (only typical values at 25^{O} C are shown).

Table 1: RC time values depending on por and esd_trigger signals.

	por=1, unpowered	**por=0, powered**
esd_trigger=0, event detection	~0.1 us	~10 us
esd_trigger=1, event response	~5 us	~20,000 us

For the detection of an un-powered stress event, e.g. CDM or HBM, a relatively short (~100 ns) RC time value is used to distinguish between a true stress event and a regular power-up ramp. False triggering during fast power-up must be avoided, but safe ESD triggering must also be guaranteed. Once the TC turns on, the RC time value is increased to ~5 us to ensure a sufficient on-time response of the clamps to fully dissipate any kind of unpowered ESD event.

For the detection of a powered stress event, e.g. PESD, EFT or power surge, a larger RC value (~10 us) is used because such events can cause pin stress waveforms with a slower effective rise time than e.g. an HBM event. It is assumed that this larger RC value will not cause issues with false triggering when the chip is powered up because board-level decoupling capacitance and power management typically prohibit such rapid supply voltage changes.

When the TC turns on due to an overstress event during powered chip operation, a very large response time value (~20 ms) enables the new proportional regulation mode described in this work. In this mode a very large resistor R effectively makes capacitor C work like a voltage source V0 that stores the average voltage that has occurred on the esd_boost bus just before the event was detected (see Fig. 2b). Then, as esd_boost exceeds V0 by more than the threshold voltage of INV1, the two inverter stages effectively amplify any further voltage increase on esd_boost and provide the amplified voltage on the esd_trigger output. This makes the clamps conductive trying to counteract the voltage increase caused by the stress event and forming a regulative feedback loop. The voltage on an IO pad would be roughly limited to:

$$V_{IOmax} = V_{boost0} + V_{thINV1} + V_{A2} \qquad (1)$$

This describes the regulation margin. V_{boost0} is the average esd_boost voltage just before the stress event, V_{thINV1} is the threshold voltage ("voltage switch point") of INV1, and V_{A2} is the voltage drop across the diode A2 of the stressed IO pad. This formula neglects the small parasitic voltage drops along the esd_boost and vss buses, between the stressed pad and the TC.

B) Trigger Circuit Implementation

Fig. 3 shows the TC schematic. The capacitor of the RC stage is implemented using Mp5, and the variable resistive element is implemented using a chain of long-channel devices Mn4, Mn5 and Mn6. The RC filter node N2 feeds into a first inverter stage Mn9, R1 and then via node N3 into a second inverter stage Mp6, R2. Inverters with resistive ballast (poly resistors) are used to limit the open-loop gain, thus avoiding unwanted oscillations of this rail clamp network. SPICE simulations (in AC mode) were used as a design aid to ensure sufficient phase margin of the voltage regulation loop that this new triggering scheme represents. Process and temperature variations as well as worst case pad ring configurations (i.e. number/locations of TCs relative to IO pad cells) were taken into consideration.

Inside the TC, the signals Tb and T are derived from the esd_trigger signal using inverters Inv2 and Inv3, respectively. Inv2 has a low threshold voltage to detect even the weakest clamp triggering. The por signal is first inverted by Inv1 and then coupled to a simple latch circuit (cross-coupled inverters Inv4 and

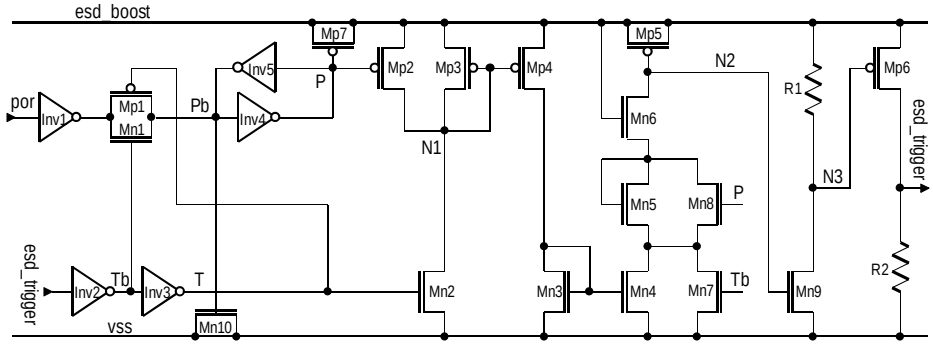

Fig. 3: Trigger circuit schematic.

Inv5) via a transmission gate Mn1, Mp1. This transmission gate, controlled by the Tb and T signals, and the latch provide a stable internal signal P that stores and preserves the correct state of the por signal from the beginning to the end of an overstress event. During a stress event the por signal integrity may suffer, which could negatively impact TC operation if the por signal were used directly. The two additional helper capacitors Mp7 and Mn10, which are attached to the P and Pb nodes, respectively, ensure that during an unpowered stress event (when esd_boost ramps up from 0 V) the internal P node always comes up high.

The T, Tb and P signals are also used to control the variable resistance of the RC stage. During the detection phase of an unpowered stress event, when signals P and Tb are high, only Mn6 is effective in the resistive chain since Mn5 is shorted out by switch Mn8, and Mn4 is shorted out by switch Mn7. Once the TC turns on (response phase), signal Tb goes low and Mn4 becomes effective, too, in addition to Mn6. The effective on-resistance of Mn4 is controlled by the resistance of Mn2 via two current mirrors Mp3, Mp4 and Mn3, Mn4. The current mirror ratios are skewed to amplify the effective resistance of Mn2, which allows for the use of smaller Mn2 and Mn4 devices to preserve layout area.

During the detection phase of a powered stress event, both Mn5 and Mn6 are effective in the resistive chain. Once the TC turns on (response phase), Mn4 becomes effective, too. However, the effective resistance of Mn4 is now much larger than during an unpowered stress event because Mp2 is turned on as well, absorbing much of the Mn2 current and significantly weakening the current through the current mirror.

C) POR Circuit Implementation

Fig. 4 shows the POR schematic. It sets the por output signal to high whenever the esd_boost voltage ramps up from 0 V, i.e. during regular power-up of the chip or during an unpowered stress event. At the beginning of such a ramp all the capacitive elements Mp1, Mn1, Mn2 and Mn3 are discharged. The internal porh signal comes up in a high state due to its capacitive coupling to esd_boost and the porl and porclr signals come up in low states due to their capacitive coupling to vss. The cross-coupled inverter Inv1 and the NOR gate (Nor1) form a latch circuit, which is set via the capacitors during an esd_boost ramp and which will later be reset by the porclr signal.

When the latch is set, the internal POR state signal porint is driven high via the two inverters Inv2 and Inv3. The Inv3 is implemented as a Schmitt trigger to obtain clean signal transitions. The por output signal is created from the porint signal via another two inverters Inv4 and Inv5.

Long-channel device Mp3 and capacitive element Mn3 form an RC delay stage whose time constant is used to define the delay of the porclr signal relative to the start of the esd_boost ramp. During the ramp, when porint is high, Mp2 is turned off and Mp3 is turned on. The resistive element Mp3 charges capacitive element Mn3 until the node rc reaches the switch point of inverter Inv6, which is implemented as a Schmitt trigger to obtain clean signal transitions. When the output of Inv6 goes low, the Nand1 gate switches the porclr signal to high and the latch resets causing the porint and por signals to go low. Then the device Mp2 turns on and quickly pulls the rc node

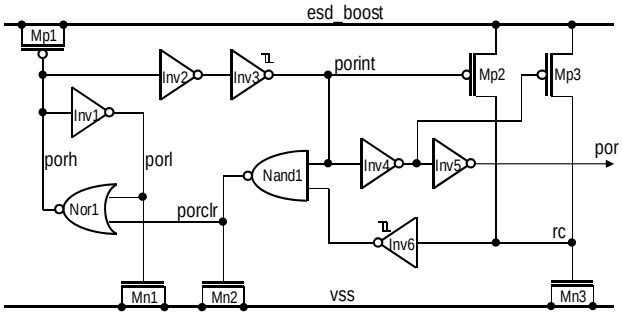

Fig. 4: Power-on reset (POR) schematic.

high. The time constant of this RC filter is set large enough to keep the por signal high during regular power-up ramps that might otherwise be falsely detected as powered stress events.

IV. The IO protection design

The overstress protection method presented here was implemented in an IO pad library for microcontroller applications using a triple-well technology with embedded flash memory built in a 90nm technology node [5]. The MOSFETs are 3.3 V devices intended for IO interfaces, the allowed supply voltage range of this pad library is 1.7 V to 3.6 V.

A) Protection Diodes and IO Drivers

All protection diodes were implemented as perimeter intensive, poly-bounded, multi-finger diodes. A detailed description of these diodes with TLP data and compact model information can be found in [3,6].

Referring back to Fig. 1, the diodes A1, A2 and A5 are P+/Nwell diodes whereas diodes B and E are N+/Pwell diodes in an isolated Pwell (IPW). The Nwell isolator tubs needed to form these IPWs are tied to the vdd bus. Diode B and driver Mn were laid out as a combination device with their fingers in an alternating sequence as shown in Fig. 5. This dissipates the heat generated during a stress event better and increases the effective It2 failure currents of both devices since each individual device is spread out over a larger area and only one of the two devices can conduct at a time. The diode B conducts during a negative stress on an IO pad, the driver Mn may go into snapback conduction during a positive event. A combination device for diode A1 and driver Mp was not feasible because their Nwells are not at the same net. The Nwell of diode A1 is connected to the esd_bus while the Nwell of driver Mp is implemented as a tracking well, as mentioned earlier.

Driver Mn has silicide protection (SiProt) [7] on the Drain (Fig. 5) to further increase its failure current (It2) and failure voltage (Vt2) [8]. Driver Mp and clamp M1 are laid out as fully silicided devices. The Mn/B combo device and the clamp M1 reside in IPWs with their Nwell isolator tubs tied to the vdd bus.

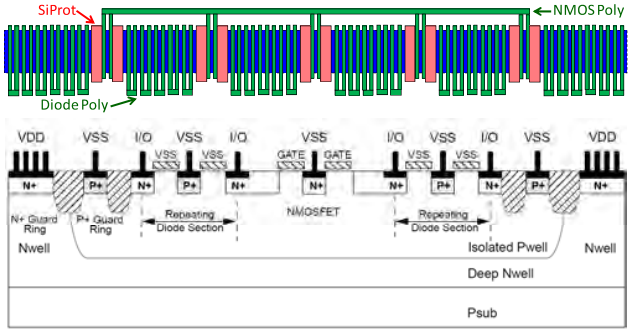

Fig. 5: N+/Pwell diode/NMOS driver combination device. Top: Layout view. Bottom: Cross section.

Fig. 6 contains TLP IV data of an NMOS driver test structure with a layout style that resembles that of the actual diode/MOSFET combination device used in the IO cell. The It2 and Vt2 values of this SiProt NMOS are much larger compared to a fully silicided device (e.g. the NMOS clamp discussed later). This is a general advantage of SiProt devices and is due to the local ballast resistance added by the unsilicided regions causing more uniform bipolar conduction and better heat dissipation [8]. As is typical of thermally induced damage, the device failure point has a strong dependence on pulse length (Fig. 6, top). At very short, CDM-like pulse length, this device is almost indestructible (Vt2 above 15V).

The TLP gate bias dependence of the NMOS driver structure can be seen in the bottom plot of Fig. 6. With a 50 ns pulse length, the Vt2 is ~7.2 V and the It2 per unit gate width is ~8 mA/um, which is where leakage starts to increase noticeably. This Vt2 value will serve as an important optimization constraint later. The data also show that Vt2 does not change over a relatively wide range of gate bias Vgs.

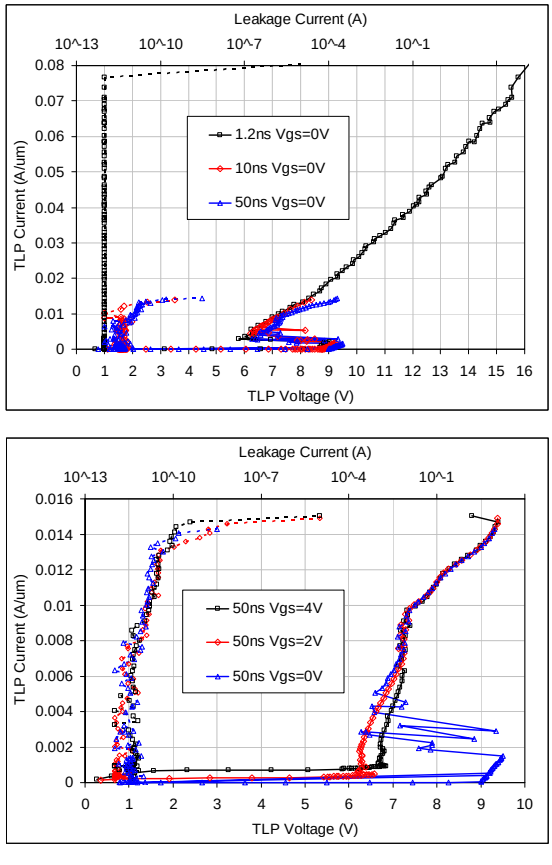

Fig. 6: TLP IV data of an NMOS driver test structure.

TLP data of a fully silicided PMOS driver test structure are shown in Fig. 7. This PMOS structure also has the expected strong It2 dependence on pulse length (Fig. 7, top). With a 50 ns pulse length (Fig. 7, bottom) and over a relatively wide gate bias range, the worst case Vt2 is about -7.2 V and the It2 per unit gate width is about -13 mA/um. There is a noticeable dependence of Vt2 and It2 on the gate bias Vgs.

B) IO Cell Floor Plan

Fig. 8 shows the device and bus floor plan of an IO cell corresponding to the schematic of Fig. 1. Special effort was made to place protection devices directly under buses and pad landings that they connect to. This minimizes parasitic voltage drops and the risk of damaging interconnects lines. Buses are implemented in the top two metal layers of a four metal backend process, with the top layer being a thick metal. Lower metal layers are used for device and guard ring strapping and for local interconnections. The widths of the main buses (vdd, vss, esd_bus) were derived from a tradeoff between supply bus voltage drop constraints requiring wide vdd and vss supply rails and maximizing overstress protection performance requiring wide esd_bus and vss rails.

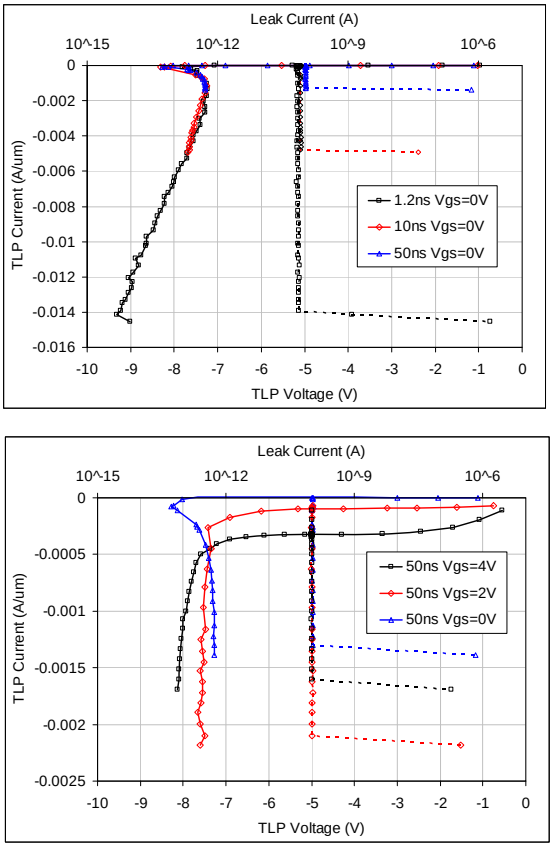

Fig. 7: TLP IV data of a PMOS driver test structure.

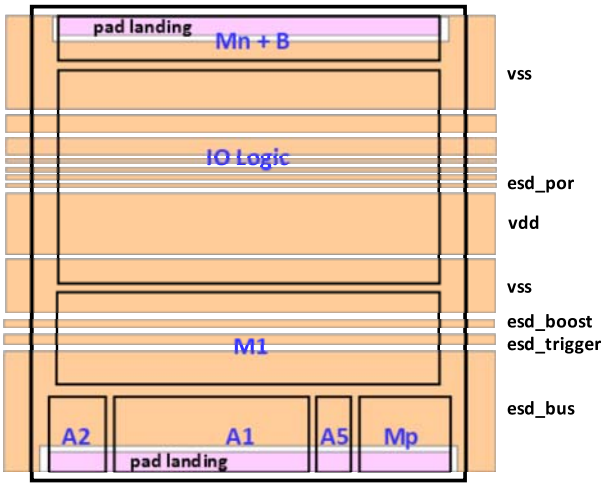

Fig. 8: IO pad cell floor plan.

C) Rail Clamp Device

The protection circuit of this work was not designed to function only in active MOSFET clamp conduction. Similar to [9] we use a "hybrid" mode where the clamps are allowed to enter bipolar conduction ("snapback") to extend their robustness into the upper range of common overstress events, e.g. CDM or PESD (with its initial high-current spike). This "hybrid" clamp conduction mode offers superior layout area efficiency, but also requires additional design techniques as outlined in the following.

The new TC scheme for powered operation is one key enabler for reliable snapback operation of the clamp. Due to its immediate, "proportional" Vgs increase in response to a powered stress event, the bipolar turn-on voltage Vt1 of the clamp is much reduced from the Vgs=0 V case [10]. A low Vt1 enables more uniform bipolar turn-on of a large multi-finger clamp device and it may even promote simultaneous turn-on of individual clamp devices in parallel, e.g. in adjacent pads. As mentioned before, conventional TCs may experience triggering issues in powered operation and may therefore not be able to raise Vgs significantly at the beginning of a powered stress event.

Furthermore, a Pwell pump was implemented using a new layout method (see Fig. 9). Pump fingers were inserted periodically into a multi-finger clamp device. They consist of abutted N+/P+ regions shorted by silicide, but not contacted. The N+ regions of pump fingers are the Source of an NMOS transistor Mpmp whose Drain and Gate are shared with the clamp M1, as shown in the conceptual schematic of Fig. 9. The P+ regions of pump fingers act as additional, actively driven Pwell ties. When the Gate is turned on, Mpmp conducts and injects current into the IPW underneath. The injected current flows to the Pwell perimeter tie

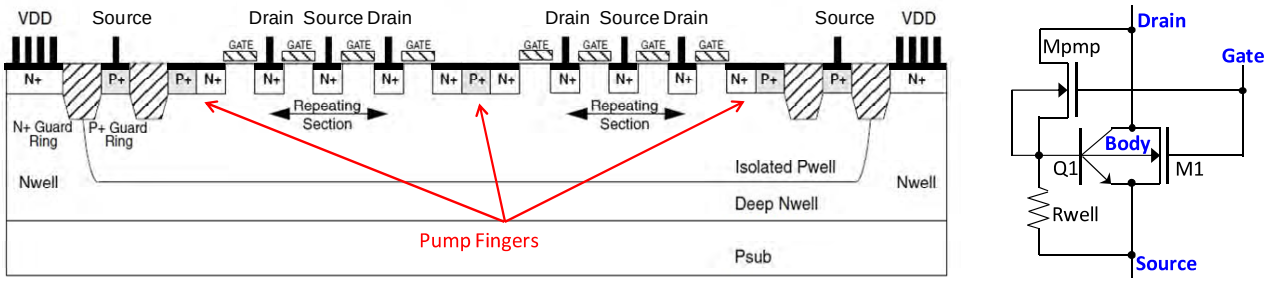

Fig. 9: NMOS clamp structure with active Pwell pump. Left: Cross section. Right: Conceptual schematic.

ring, which is shorted to the Source of the clamp. This current flow produces a voltage drop across the well resistance (Rwell) between the pump fingers and the Pwell tie ring. Therefore, the local Body potential under the clamp fingers gets elevated facilitating uniform bipolar turn-on of the lateral NPN (Q1) due to a resulting Vt1 reduction of the clamp [11]. During normal, powered-up chip operation, the device Mpmp stays turned off and the Pwell pump is inactive. In that case, the Body potential is solely controlled by the perimeter Pwell tie ring and the Vt1 of the clamp is not reduced.

Note that these clamp device improvements and the new proportional triggering scheme complement each other, but may also be used independently in some designs. Though a dedicated esd_bus ("ghost" rail) was used here, conventional protection circuits with the vdd bus as the clamp Anode rail could equally benefit from this method.

Fig. 10 shows TLP IV data of a large Pwell-pumped clamp test structure like the one used in the IO cell. The expected TLP pulse length dependence of the failure current It2 can again be observed (Fig. 10, top). The relatively low on-state voltage of this fully silicided clamp in bipolar conduction (between about 4.7 V and 5.5 V) helps to keep the IO pad voltage low during a positive stress event on an IO pad, as will be shown later. The relatively high It2 at very short pulse length (greater than 10 mA/um with 1.2 ns pulses) can be used as an effective protection strategy against CDM and the high-current spike of a PESD event.

V. Circuit Optimization

The protection devices were sized based on SPICE network simulations. A circuit optimization technique similar to [3] was applied where the ESD device sizes were defined as optimization parameters and a closed-loop iterative solver strategy was used to minimize the total layout area of the IO cell.

A) Optimization Targets and Constraints

The It2 failure levels of the protection devices (diodes and clamps) were one set of optimization constraints. The Vt2 failure levels of the output drivers Mn and Mp defined another set of constraints. The IO pad voltage relative to the local vss and vdd buses had to be kept below these Vt2 values during a stress event.

The unpowered performance targets for this IO library were 4 kV HBM and 750 V CDM. The powered performance targets were 8 kV PESD (ESD gun stress), 36 V Langer-IC EFT (P201 probe) and 480 V Langer-IC EFT (P301 probe) [12]. The PESD and Langer-IC EFT test methods apply a contact-discharge directly onto a chip pin. Special test boards were used to evaluate transient immunity while the chip was powered up to its nominal supply levels and was running a primitive code loop that invokes most of the functional modules of the chip and which also controls a blinking LED sequence. Decoupling

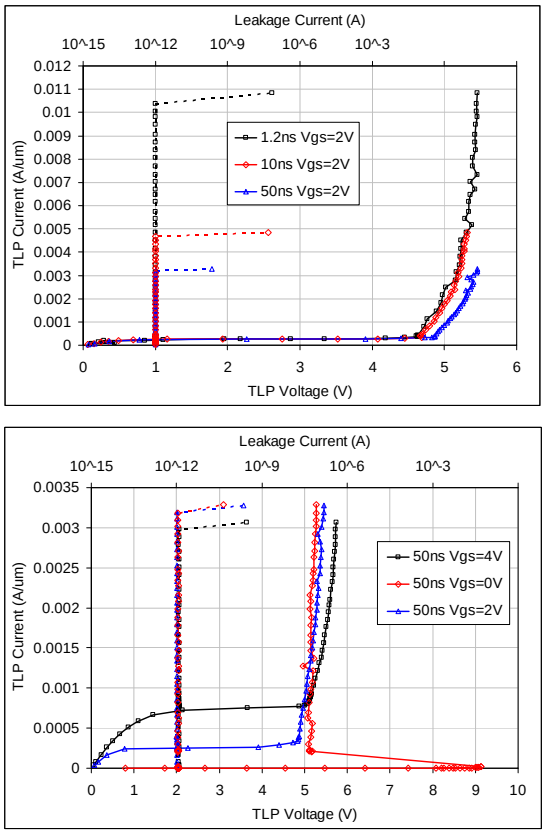

Fig. 10: TLP IV data of a clamp device test structure.

capacitors were attached to each supply pin on the board. IO pins were either left floating, driving out high or low, or configured as inputs with external pull-up or pull-down resistors. This test board configuration allowed damage screening by detecting a supply current increase, non-recoverable errors in the code execution, or an increase in pin leakage.

A study of the nature of these various unpowered and powered stress events has shown that the 8 kV PESD gun stress poses the highest risk for damage with this protection scheme. The PESD stress was therefore used as a benchmark for circuit optimization. Typical PESD current waveforms were measured using an inductive current probe placed around the tip of the ESD gun while stressing an IO pin on the test board. Fig. 11 shows a measured 8 kV PESD current waveform scaled to the discharge voltage. Also shown is the simulation model used for circuit optimization. It is an empirical current waveform fitted to measured PESD data. A very brief initial current spike of about 3.2 A/kV is followed by a longer current pulse with a peak current of about 1.5 A/kV. The used ESD gun had good discharge voltage scaling behavior as confirmed by additional 4 kV and 6 kV tests, which delivered almost identical waveforms (not shown).

Comparisons between PESD and TLP failure levels of protection structures have shown that the main PESD current pulse is roughly equivalent to a 50 ns TLP pulse with a slightly lower zap current. Therefore, we generally use 50 ns TLP data to estimate device failure levels for PESD. This is also the reason for the relatively large amount of 50 ns TLP data provided in this paper. However, we must also realize the potential danger of the initial PESD spike causing damage due to overvoltage and briefly pushing MOSFETs into bipolar conduction. The 1.2 ns TLP data can be used to estimate device damage levels in this case.

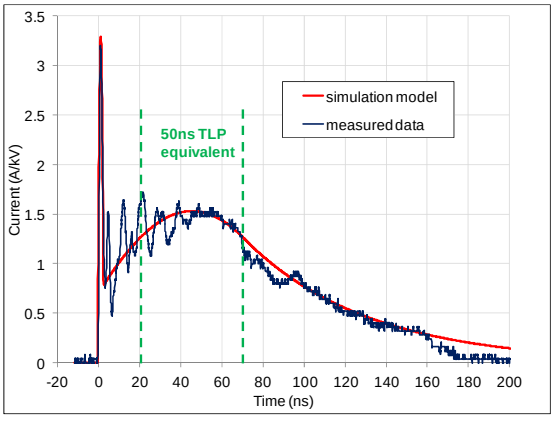

Fig. 11: PESD waveform measured vs. simulation.

B) Simulation Setup

Special high-current models implemented in VerilogA were used to accurately describe the strong forward bias regime of protection diodes [3,6]. Furthermore, an empirical VerilogA model was created for the bipolar mode of the clamps, based on the TLP data shown in Fig. 10 and additional TLP data not shown in this paper. This model takes into account the dependence of the Vt1 on gate bias Vgs and on body bias Vbs (Pwell pumping). In the "off" state, the clamp bipolar model gives zero current. The snapback regime is described by a simple linear IV curve with a given holding voltage and on-resistance. A smooth transition between the "off" and "on" states is provided with a programmable turn-on delay. The regular active MOSFET model is working in parallel.

A large, un-terminated IO pad ring with a single VDD/VSS pad pair was used as a worst-case simulation benchmark for circuit optimization [3]. The VDD was powered up to the typical supply voltage (3.3 V). Positive PESD stress was applied to an IO pad while an adjacent VSS pad was grounded.

Table 2 lists the extracted bus resistance values and optimized ESD device sizes. The diode sizes are measured in total junction perimeter and clamp MOSFET sizes in total gate width. The total IO cell area occupied by protection devices including all related guard ring structures is ~3330 um^2.

Table 2: Bus resistance and device size information.

Diode A1	550 um
Diode B	535 um
Diode A2	128 um
Diode A5	37 um
Diode E	671 um
Clamp M1	1122 um
Clamp M0	4480 um

R_esd_bus	0.153 Ω
R_vss	0.122 Ω
R_vdd	0.142 Ω
R_esd_boost	0.43 Ω
R_esd_trigger	3.9 Ω

C) Simulation Results

Fig. 12 shows the 8 kV PESD simulation response of the optimized IO protection circuit. During the main PESD pulse (~50 ns long), the input voltage on the IO pad is limited to ~7 V, which is below the Vt2 of the NMOS driver Mn (Fig. 6). The IO pad voltage due to the brief initial PESD spike reaches ~8.2 V which is well below the NMOS driver Vt2 for very short pulses (Fig. 6). The bottom plot of Fig. 12 zooms into the initial PESD spike, the duration of clamp bipolar conduction is highlighted in the plot. The spike current density of all clamp devices stayed well below the measured It2 of Fig. 10 (with 1.2 ns TLP pulses).

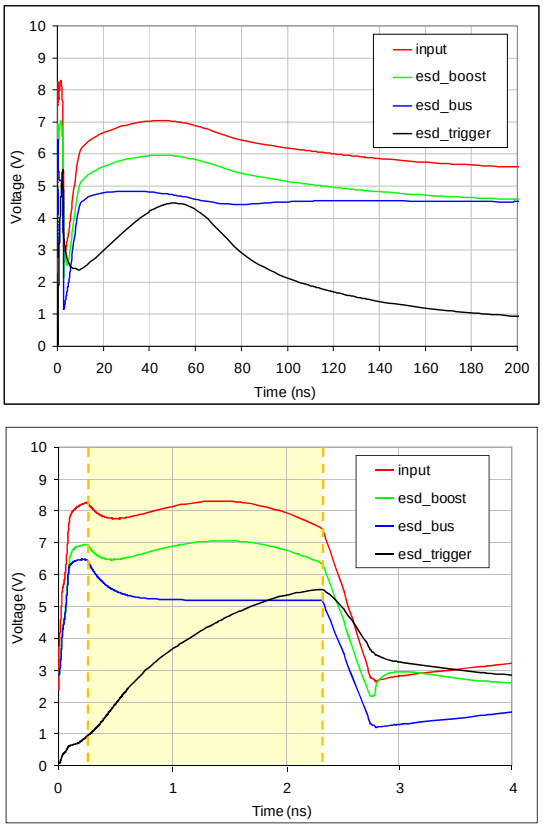

Fig. 12: Simulated PESD response.

Fig. 12 also gives a good insight into the new proportional clamp triggering scheme described earlier. The esd_trigger voltage waveform closely follows the PESD current peak of Fig. 11, with the maximum esd_trigger voltage occurring at ~45 ns. This means that the conductivity of the clamp network is modulated in proportion to the applied stress current and the IO pad voltage is clamped within a certain regulation margin.

Additional transient simulations were performed to confirm that the new proportional triggering scheme does not negatively impact normal chip operation, e.g. due to noise on the VDD supply causing unwanted clamp conduction. A single sinusoidal voltage peak with varying amplitude and a given frequency was applied to the VDD pad of an IO segment and the transient current response was monitored. Fig. 13 shows the simulated response at room temperature and with typical process conditions, just to illustrate the general behavior. A variety of pulse frequencies were simulated to gauge the impact of pulse rise-time. It can be seen that the main frequency window for clamp conduction is somewhere between 100 kHz and 1 GHz. Outside this window, the TC either does not detect the pulse (too slow to be recognized as a stress event) or the TC cannot react fast enough to the pulse.

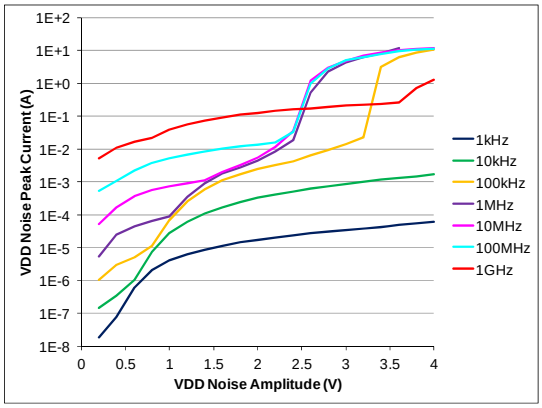

Fig. 13: VDD noise immunity simulation results.

Within this frequency window, the amplitude of the pulse must furthermore be larger than ~2.5 V to cause significant clamp turn-on. Given that there are typically large decoupling capacitors attached to VDD on the board, such large voltage noise is not expected as a result of normal chip operation.

VI. Test Results

This protection circuit was first implemented on an IO test chip and then on real microcontroller products. Table 3 shows the measured product performance levels on IO, VDD and VSS pads. LU tests (at 125C) with 100 mA injection and 50% supply overvoltage testing also passed.

Table 3: Measured Performance Levels (“Damage Free”).

HBM	CDM	PESD	P201	P301 (IOs only)
4 kV *	1000 V *	8 kV	36 V *	480 V *

**) Not tested at higher zap levels.*

Additional pulse characterization was done on a product wafer. The correct functionalities of the POR and TC, according to the event detection/response design targets of Table 1, were confirmed using supply voltage ramps and transitions with various rise times and voltages. TLP measurements (50ns pulse length) of an IO pad vs. an adjacent VSS pad were performed and the results are shown in Fig. 14. These I-V curves illustrate the new active MOSFET mode regulation scheme. The unpowered IV curve shows the case where the clamps get turned on to their highest conductivity level according to this triggering scheme. The powered IV curves (Vdd=1.5 V, 2.5 V and 3.6 V) illustrate the proportional triggering mode. The V_{IOmax}, which is the IO pad voltage at which the clamps start to turn on according to Formula (1), is about 2.5 V higher than the applied Vdd. For example, V_{IOmax} is ~5 V for the case Vdd=2.5 V. For

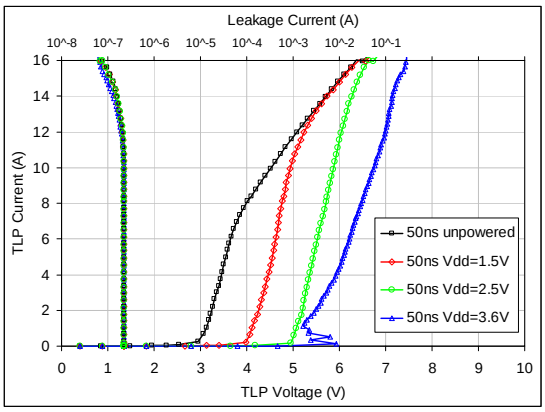

Fig. 14: TLP IV data of an IO pad vs. VSS.

the case Vdd=3.6V, the clamps appear to turn-on a bit harder than expected at low to medium current levels. A preliminary analysis of this phenomenon has indicated that this may be due to substrate interaction between the P+/Nwell diode A1 and the NMOS clamp M1, which are placed in close proximity (Fig. 8). However, although not perfectly understood at the moment, this interaction does not cause any failures.

The observed on-resistance in regulation mode of Fig. 14 can be explained by the total interconnect resistance between the IO and VSS pads. Note that the "local" stress voltage (i.e. the voltage across driver Mn) in the zapped IO pad is assumed to stay fairly constant (at ~V_{IOmax}) until the I-V curve describing the regulation mode merges with the unpowered I-V curve (e.g. as seen in the Vdd=1.5 V curve at ~13 A). At this point the clamps cannot be turned on any harder and the IO pad stress voltage starts to increase.

All of these TLP cases experience a change in IO pad leakage (damage) starting at ~12 A. This corresponds well with our 8 kV PESD target with a ~12 A equivalent TLP current level at 50 ns pulse length.

VII. Conclusions

The design method introduced in this work has extended the active MOSFET rail clamp approach from the component level ("un-powered stress") to the system level ("powered stress"), directly competing with on-chip SCR based protection techniques. With a new clamp triggering scheme that regulates the pad voltage rather than simply switching the clamps on or off, a collapsing VDD supply, and hence potential system resets, can be avoided. A new well pumping technique for the clamp device has also been introduced for enhancing the multi-finger turn on of the parasitic bipolar of the clamp. Microcontroller products built with this protection scheme have demonstrated excellent powered and un-powered transient immunity performance, meeting a 8kV direct pin contact PESD stress target on a test board.

Acknowledgements

The authors would like to thank Brian Winstead for TCAD simulation support and Ann Concannon for mentoring this paper.

References

[1] Industry Council on ESD Target Levels, "White Paper 3: System Level ESD Part I: Common Misconceptions and Recommended Basic Approaches," Rev. 1, 2010, at http://esda.org

[2] M. Stockinger et al., "Boosted and Distributed Rail Clamp Networks for ESD Protection in Advanced CMOS Technologies," ESD Symp. 2003

[3] M. Stockinger et al., "CDM Protection Design for CMOS Applications Using RC-Triggered Rail Clamps," ESD Symp. 2009

[4] C. Torres J. Miller, M. Stockinger, M. Khazhinsky, J. Weldon, "Modular, Portable, and Easily Simulated ESD Protection Networks for Advanced CMOS," ESD Symp. 2001

[5] K.-M. Chang , S.-T. Kang, J. Yater, "An advanced embedded flash technology for broad market applications", Solid-State and Integrated Circuit Technology Conf. 2012

[6] M. Stockinger, J. Miller, "Characterization and Modeling of Three CMOS Diode Structures in the CDM to HBM Timeframe," ESD Symp. 2006

[7] K.-L. Chen, G. Giles, D. Scott , "Electrostatic Discharge Protection for One Micron CMOS Devices and Circuits", IEDM, vol. 32, pp. 484-487, 1986

[8] G. Notermans, A. Heringa, M. van Dort, S. Jansen, F. Kuper, "The Effect of Silicide on ESD Performance", Int. Reliability and Physics Symp. 1999

[9] Scott Ruth, M. Stockinger, J. Miller, V. Whitney, M. Kearney, S. Ngo, "A CDM Robust 5V Distributed ESD Clamp Network Leveraging Both Active MOS and Lateral NPN Conduction," ESD Symp. 2011

[10] M. Khazhinsky, J. Miller, M. Stockinger, J. Weldon, "Engineering Single NMOS and PMOS Output Buffers for Maximum Failure Voltage in Advanced CMOS Technologies," ESD Symp. 2004

[11] C. Duvvury, S. Ramaswamy, A. Amerasekera, R. Cline, B. Andresen, V. Gupta, "Substrate Pump NMOS for ESD Protection Applications," ESD Symp. 2000

[12] G. Langer, B. Deutschmann, G. Auderer, "Characterizing the Immunity of Integrated Circuits against Electrical Fast Transient Disturbances", http://www.langer-emv.de/en/literatur/technical-reports/, 2004