SOFICS
SOLUTIONS FOR ICS

Conference paper

# GGSCRs: GGNMOS Triggered Silicon Controlled Rectifiers for ESD Protection in Deep Sub-Micron CMOS Processes

EOS/ESD symposium 2001

In this paper, design aspects, operation, protection capability and applications of SCRs in deep submicron CMOS are addressed. A novel Grounded-Gate Nmos Triggered SCR device (GGSCR) is introduced and compared to the LVTSCR. Experimental verification, including endurance testing, demonstrates that GGSCRs can fulfill all ESD protection requirements for today's IC applications in different 0.25um, 0.18um and 0.13um CMOS processes.

# GGSCRs: GGNMOS Triggered Silicon Controlled Rectifiers for ESD Protection in Deep Sub-Micron CMOS Processes

Christian C. Russ, Markus P. J. Mergens, Koen G. Verhaege[1], John Armer,
Phillip C. Jozwiak, Girija Kolluri, Leslie R. Avery

Sarnoff Corporation, 201 Washington Road, Princeton, NJ-08543, USA
(phone) 1-609-734-2526 (fax) 1-609-720-4848
e-mail: cruss@sarnoff.com, mmergens@sarnoff.com

(1) Sarnoff Europe, Brugse Baan 188A, B-8470 Gistel, Belgium, kverhaege@sarnoff.com



*Abstract* – In this paper, design aspects, operation, protection capability and applications of SCRs in deep sub-micron CMOS are addressed. A novel Grounded-Gate Nmos Triggered SCR device (GGSCR) is introduced and compared to the LVTSCR. Experimental verification, including endurance testing, demonstrates that GGSCRs can fulfill all ESD protection requirements for todays IC applications in different 0.25um, 0.18um and 0.13um CMOS processes.

## I. Introduction

Silicon controlled rectifiers (SCRs) have long been used because of their superior ESD performance [1]-[4]. The ESD performance of SCRs can provide an area gain factor of typically 4 to 5 over silicide-blocked grounded-gate NMOS protection devices. Recently, concerns about trigger speed and protection capability of SCRs – in deep sub-micron technologies have been raised [5].

Low-ohmic conduction deep in wells and substrate makes SCRs one of the most powerful ESD devices. An alternative approach for extremely high performance by non-SCR devices is found in [6].

Also, little has been published in the recent years about SCRs besides the introduction of new approaches for mature technologies [7], [8].

The purpose of this paper is to review SCR designs for their applicability in current CMOS technologies with minimum feature sizes of 0.25um, 0.18um, 0.13um and below. We introduce a new GGNMOS triggered SCR (GGSCR) where triggering is provided by an external trigger source. We will discuss that the internally triggered SCR (such as the Low Voltage Triggering SCR - LVTSCR) may not be fast enough to protect thin gate oxides. Triggering voltages during fast transients such as CDM have indicated a problem [5]. The physics of the trigger and turn-on speed of both, the internally triggered LVTSCR and the new GGSCR, as they relate to layout design parameters, will be discussed.

## II. SCR Design and Operation

### II.a. SCR Roadmap

Different SCR types for ESD protection in CMOS technologies from above 0.5um down to 0.1um are reviewed in Table 1.

Table 1 SCR Roadmap for CMOS technologies showing the trend in the different SCR types: MLSCR (Modified Lateral SCR), LVTSCR (Low Voltage Triggering SCR), GGSCR (GGNMOS-triggered SCR).

| Applicability of SCRs | | > 0.5 | 0.5 | 0.35 | 0.25 | 0.18 | 0.13 | 0.10 |
|---|---|---|---|---|---|---|---|---|
| MLSCR | n+ trigger diff. (LOCOS) | proven | | not possible | | | | |
| LVTSCR | n+ trigger diffusion | proven | | | | problematic | | |
| GGSCR | NMOS-triggered | expected | | | proven | | | |

One difficulty with SCRs in early sub-micron CMOS was that the intrinsic well-to-well breakdown and thus trigger voltage was too high to efficiently protect circuit nodes (gates) against ESD. The first SCRs which had a trigger voltage reduced from the one purely determined by well-to-well breakdown was the modified lateral SCR (MLSCR) [3]. A trigger diffusion was used to reduce the breakdown voltage. This trigger diffusion was typically determined by the edge of the LOCOS (local oxidation of silicon).

The next 'upgrade' was the now widely used LVTSCR (low voltage triggering SCR), where the trigger voltage is further reduced to that of the NMOS

drain-substrate junction, which became important for protecting thin gate oxides in sub-micron technologies (Figure 1).

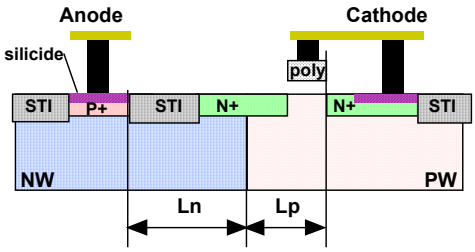


Figure 1 Cross-section of a Low-Voltage Triggering SCR (LVTSCR) with a center N+ trigger diffusion defined by STI and the poly gate of an integrated GGNMOS.

Both, MLSCR and LVTSCR are triggered by the breakdown of an internal junction.

For 0.25um down to 0.1um CMOS the authors have successfully used the GGNMOS-triggered SCR (GGSCR) in which an external NMOS defines the trigger voltage and injects the trigger current into the SCR. Two main types of the GGSCR have been realized: with STI-block (Figure 2) and with STI (Figure 3) as will be discussed later.

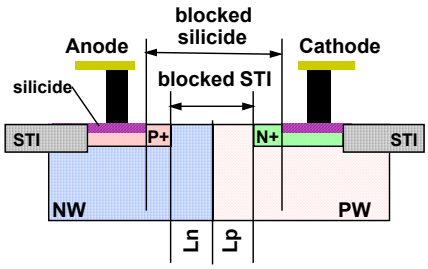


Figure 2 Cross-section for a GGSCR structure without Shallow Trench Isolation (STI) between Anode and Cathode. The trigger current is provided by an external trigger current source. The lateral dimensions Ln and Lp are important parameters for a high and reliable SCR performance.

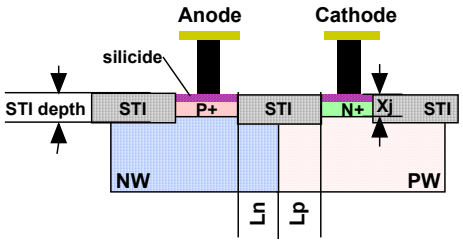


Figure 3 Cross-section for a GGSCR with STI.

Table 2 reviews several technologies that have been considered for SCR design and from which measurement results are presented in this paper.

Table 2 Overview on CMOS technologies considered for SCR design.

| CMOS Process | 0.25 | 0.25 | 0.18 | 0.18 | 0.18 | 0.18 | 0.13 | 0.13 | 0.10 |
|---|---|---|---|---|---|---|---|---|---|
| STI depth [um] | n/a | 0.26 | 0.33 | 0.38 | 0.28 | 0.35 | 0.37 | 0.35 | 0.28 |
| Xjnwell [um] | n/a | 2.1 | n/a | 1.2 | 1.35 | n/a | 1.4 | 1.4 | 1.4 |
| Xjn [um] | n/a | 0.22 | n/a | n/a | 0.18 | n/a | 0.3 | 0.2 | 0.2 |
| Xjp [um] | n/a | n/a | n/a | n/a | 0.15 | n/a | 0.18 | 0.2 | 0.2 |
| substrate | epi | bulk | bulk | epi | bulk | bulk | bulk | bulk | bulk |

# II.b. Operation Principle of the GGSCR

## II.b.1 Review of the LVTSCR

CMOS SCRs are formed by a 4-layer PNPN structure, which can be implemented and triggered in different ways. The most commonly known type of triggered SCR is the LVTSCR [2], [3], [7] (schematic Figure 4, cross-section Figure 1). The LVTSCR uses an integrated grounded gate NMOS (GGNMOS) to trigger SCR action - rather than well-to-well breakdown.

Recently, concerns were raised about the ability of LVTSCRs to protect very thin gate oxides [5]. This will be further assessed below.

The LVTSCR cross-section with its equivalent circuit elements (NPN transistors $Q_{Nsurface}$ and $Q_{Nbulk}$, PNP transistor $Q_P$) of Figure 5 further expands beyond the schematic of Figure 4. The GGNMOS forms a lateral NPN bipolar transistor $Q_{Nsurface}$ close to the surface and with a short base defined by the poly gate. In standalone such parasitic NPN transistor will feature good bipolar characteristics and a fast response time. The NPN transistor in the bulk $Q_{Nbulk}$ has a longer base length Lp because of layout constraints. Thus $Q_{Nbulk}$ will be slower. The N-collector regions of $Q_{Nsurface}$ and $Q_{Nbulk}$ also form a distributed base region of the lateral PNP transistor $Q_P$.

The LVTSCR triggers by avalanche breakdown of the NMOS-drain-bulk junction, i.e. collector-base region of $Q_{Nsurface}$. Hole current will flow in the P-well and electron current will flow through the N+ region into the N-well. Thus, the NPN emitter (N+ cathode) and PNP emitter (P+ anode) will inject electrons and holes, respectively, and initiate SCR action.

The low-doped N-well region normally ensures good bipolar characteristics of $Q_P$. However, this low-doped base is partly bridged by the NMOS-drain N+ diffusion close to the surface resulting in a parallel PNP transistor (not shown in Figure 5). This layer, where the initial current flows after triggering, deteriorates current gain and speed of the effective $Q_P$. Eventually, equilibrium will be reached between the

NPN components (short $Q_{Nsurface}$ and long $Q_{Nbulk}$) and the PNP components of $Q_P$ (long bulk part and very long surface part through the N+ diffusion).

LVTSCR design rules require rather large lateral dimensions for Ln and Lp to accommodate all necessary spacings. These long base lengths Ln and Lp as well as the deteriorated PNP base and spatially extended by the N+ diffusion will reduce the SCR speed. In turn, this can lead to trigger voltage overshoots that may disable the LVTSCR from protecting thin oxides.

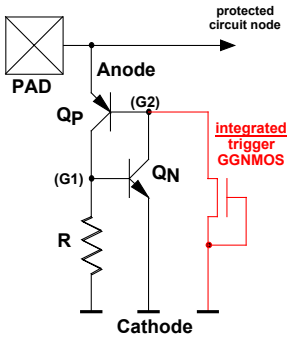


Figure 4 Schematic of a Low-Voltage Triggering SCR (LVTSCR).

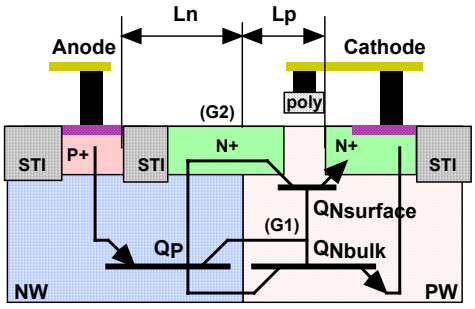


Figure 5 LVTSCR structure showing the presence of two parallel NPN transistors, one close to the surface $Q_{Nsurface}$ and one deep in the bulk $Q_{Nbulk}$.

## II.b.2 Design and Layout of the GGSCR

We introduce the GGSCR – GGNMOS triggered SCR. This device is schematically represented in Figure 6. An NMOS transistor (which resembles a GGNMOS configuration) is used as an external trigger device. It is coupled into the NPN base of $Q_N$ (SCR gate G1). Device cross-sections as shown in Figure 2 or Figure 3 form the intrinsic SCR part which allows minimum design rule geometry for the anode-cathode spacing Ln + Lp. Therefore, GGSCRs always have the smallest possible base lengths for both the PNP and NPN sub-structures.

As shown in the device layout of Figure 7, the SCR cathode is interrupted with P+ trigger tap(s) for the trigger gate G1. At these trigger taps, the trigger current coming from the external trigger NMOS is directly injected into the substrate / base of the NPN next to the SCR cathode / NPN emitter.

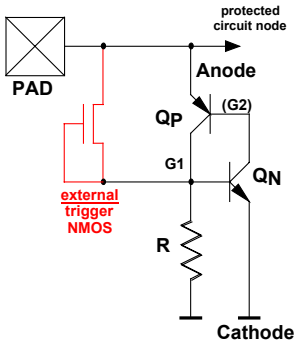


Figure 6 Schematic of the GGNMOS triggered SCR (GGSCR) used in this study.

In contrast to the LVTSCR, the trigger transistor of the GGSCR does not compete with SCR internal sub-components: firstly, because the GGNMOS / lateral NPN is *external*, and secondly, because it is an NPN connected in parallel to the *PNP* transistor $Q_P$ of the SCR.

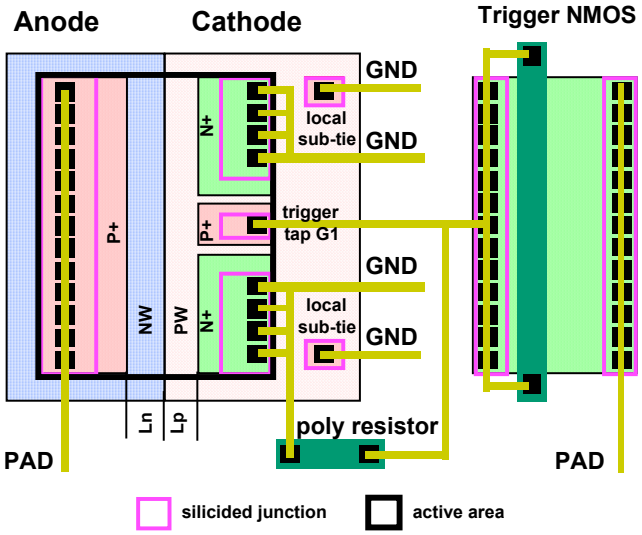


Figure 7 Layout sketch of the GGSCR showing the position of the P+ anode, the N+ cathode and the external trigger NMOS connected to the trigger tap G1 of the SCR. The position of the trigger tap close to the NPN base ensures efficient triggering of the SCR.

The resistance R from G1 to the ground controls the trigger and holding current of the SCR. This resistor R represents the following components: firstly, the substrate resistance from the trigger tap G1 to local substrate ties at a certain distance, and secondly, an external resistor. The substrate resistance is strongly conductivity modulated once the device enters SCR conduction, the external resistor remains constant. The discrete resistor allows to control the trigger and

holding current and to prevent false triggering.

### II.b.3 Characteristics of the GGSCR

All pulsed I-V and DC leakage data presented hereafter have been measured with a Barth 4002 Transmission Line Pulse (TLP) tester with a pulse duration of 100ns and default rise time of 10ns.

The pulsed IV-characteristic and the DC leakage current evolution (as a function of the pulse current) are plotted for a GGSCR in Figure 8. The GGSCR has a width of W=50um. The triggering NMOS has a minimum gate length of L=0.4um using the thicker oxide of this dual gate oxide technology. The NMOS drain and source regions were silicide-blocked.

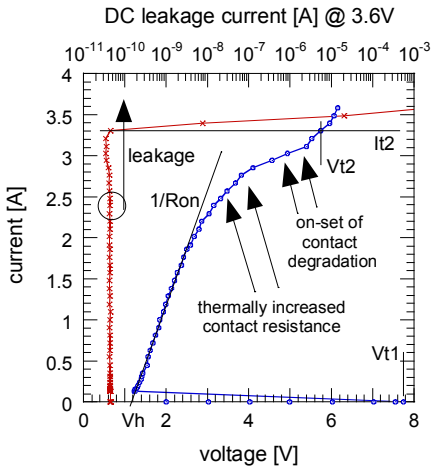


Figure 8 IV-characteristic (Y-axis vs. X1-axis) and leakage evolution (X2-axis vs. Y-axis) of a GGSCR (W=50um, Ltrigger=0.4um, 0.18um CMOS).

As can be seen in the plot, the GGSCR triggers at Vt1=7.8V which is the trigger voltage of the external NMOS. The voltage snaps back instantly to the holding voltage of the SCR (Vh=1.2V). An extremely strong current increase with a very low on-resistance of Ron=0.7 Ohms (for W=50um) is achieved. The DC leakage current stays very constant below 50pA. The destructive pulse current level (It2=3.3A) is indicated by the sudden increase of the leakage current. The voltage level Vt2 at this maximum current level It2 is still below 6V.

A typically observed failure mode in the GGSCR is not failure of the active silicon but failure at the contacts (or even the metal interconnects). This can be seen in a typical bending of the IV-characteristic before reaching It2. The bending is usually caused by excessive self-heating of the contacts leading to a strong increase in the resistivity. For a well-designed SCR, not the active device itself but its interconnects may be the factor ultimately limiting the performance.

## II.c. SCR I-V Behavior

Figure 9 compares the TLP I-V curves of LVTSCRs and GGSCRs implemented in 0.18um and 0.13um CMOS technologies.

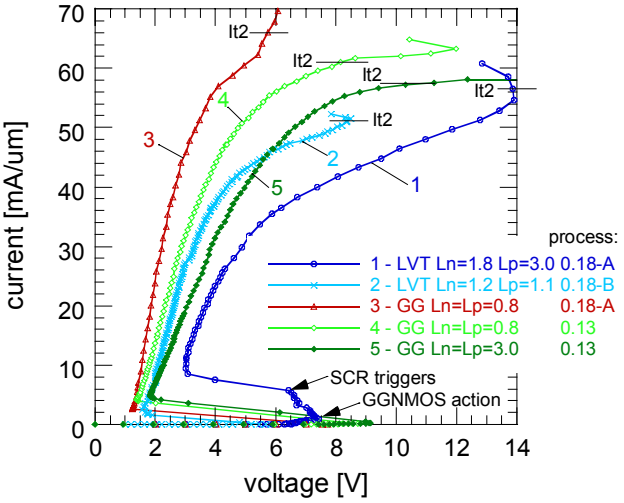


Figure 9 Comparison of LVTSCR and GGSCR TLP curves of devices originating from different processes. Only the LVTSCR shows an intermediate holding state that is defined by the integrated GGNMOS.

The following can be observed:

- The SCR trigger currents are significantly different. In the LVTSCR, both the NPN and PNP sub-structures have longer bases and additionally a part of the PNP base is bridged with the N+ trigger diffusion. With less efficient bipolar action, in some cases a much higher current is required to turn on the SCR. Before the LVTSCR actually triggers, a GGNMOS current may prevail. The latter becomes visible in an intermediate holding state of the LVTSCR.
- The LVTSCRs show a stronger sensitivity on the layout parameters leading to a more pronounced shift of the characteristics towards higher holding voltages as well as poorer clamping behavior.
- In the GGSCR, the dynamic on-resistance degradation occurs at higher current levels, because of the smaller base geometries.
- The failure current It2 is higher in the GGSCRs, which is attributed to lower power dissipation (lower dynamic on-resistance, less voltage drop).

## II.d. Speed and Transient Behavior

### II.d.1 Theoretical analysis

A physical analysis of the LVTSCR and GGSCR trigger speed is presented. First order geometrical

approximations for the transit time of the SCR $\tau_{SCR}$ can be found in Sze [9]

$$\tau_{SCR} \approx [\tau_{NPN}^2 + \tau_{PNP}^2]^{1/2}$$

with $\tau_{NPN}$ and $\tau_{PNP}$ being the transit time of the NPN and PNP transistor, respectively.

The transit time of a bipolar transistor can be written in general terms as

$$\tau_{BJT} \approx \delta \cdot L^2$$

with $\delta$ being a generic proportionality factor and L the base length. Because of the well-known difference in the mobility of holes and electrons, the ratio of the proportionality factors for NPN and PNP is

$$\delta_{NPN} / \delta_{PNP} \approx 1/3$$

We now substitute the expressions in the formula for $\tau_{SCR}$ and define a ratio of the turn-on times between GGSCR and LVTSCR, that can be simplified to

$$\tau_{GGSCR}/\tau_{LVTSCR} =$$

$$[(L_{P\text{-}GG}^4 + 9\, L_{N\text{-}GG}^4) / (L_{P\text{-}LV}^4 + 9\, L_{N\text{-}LV}^4)]^{1/2}$$

where Lp and Ln indicate the length of the NPN and PNP base, respectively, and -GG and -LV indicate GGSCR and LVTSCR, respectively.

Considering minimum layout dimensions for both LVTSCR and GGSCR, it can be concluded that the LVTSCR will have about 2x larger dimensions Lp and Ln. Considering also an internal symmetry in both the GGSCR and LVTSCR (Ln=Lp), it can be concluded that the turn-on time in a GGSCR should be at least 4 times shorter compared to the LVTSCR case:

$$\tau_{GGSCR}/\tau_{LVTSCR} \approx 1/4$$

In other words: since GGSCR can be made more compact., it will be significantly faster.

### II.d.2 Transient measurements

Figure 10 and Figure 11 report TLP measurements (2ns risetime) of the voltage transients across GGSCR, LVTSCR and GGNMOS devices. The voltage transient across a device is obtained by subtracting a reference waveform into a short at the same current level from the actually measured waveform. Because of this differential measurement, the noise level increases for higher current levels.

For low current levels, the long GGSCR device (Ln=Lp=3.0um) does not trigger and the trigger time strongly decreases for smaller Ln and Lp (2.0 → 0.5um, Figure 10-left). The Vt1 of the shortest GGSCR device and of the GGNMOS are obviously reduced by dV/dt triggering. The latter can only become significant in intrinsically fast devices while slow devices need the triggering junction to go fully into avalanche breakdown. Therefore, dV/dt is not seen in 'long' SCRs.

For high level currents and for longer devices a strong voltage overshoot is observed (Figure 10-right). For short GGSCR devices (Ln=Lp=0.5um), the transient voltage stays entirely below that of the GGNMOS. Clearly, this GGSCR has voltage transient clamping capabilities superior to the GGNMOS.

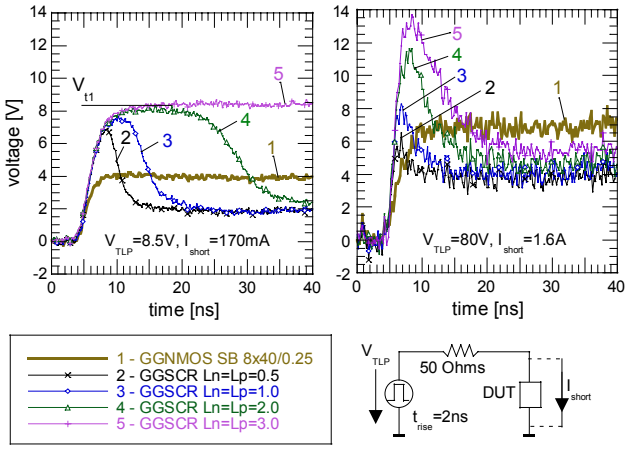


Figure 10 Voltage transients during low and high level TLP stress in GGSCR devices of different geometry in comparison to a GGNMOS. a) at a TLP pulse voltage level of $V_{TLP}$=8.5V and a short circuit current of $I_{short}$=170mA: b) $V_{TLP}$=80V and $I_{short}$=1.6A.

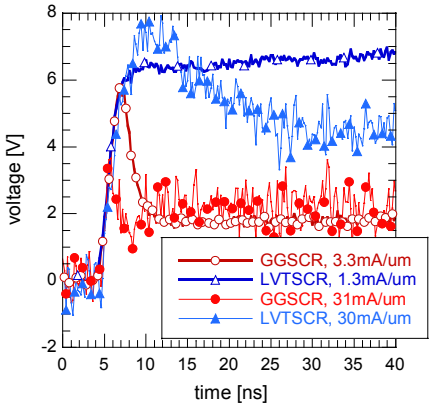


Figure 11 TLP Voltage transients during low and high level TLP stress in GGSCR devices of different geometry in comparison to a GGNMOS. (GGSCR: Ln=Lp=0.8um, LVTSCR: Ln=1.8um, Lp=3.0um).

Figure 11 compares an LVTSCR and a GGSCR device. At low currents just after triggering, the larger dimensions of the LVTSCR and its deteriorated PNP base region let the SCR not trigger but only conduct by its GGNMOS. The GGSCR triggers at a voltage lower than the LVTSCR and the voltage decays fast to its final holding voltage. For high currents (30mA/um) the LVTSCR shows an overshoot which

grows with increasing current. The voltage decay is very slow and the final holding voltage is much higher (cf. also Figure 9, 0.18-A). Under the same high current condition, the GGSCR shows dV/dt triggering as well as excellent clamping. This clearly demonstrates the benefits of the GGSCR.

### II.d.3 Discussion

We can conclude from this study that short and therefore fast GGSCRs exhibit excellently low clamping characteristics without any trigger voltage overshoot.

The trigger voltage overshoot as observed in [5] is confirmed to the extent that it only occurs for longer geometries (GGSCR and LVTSCR). This finding is in accordance with earlier observations made for single bipolar transistors under very fast transient conditions [10]. Transit time effects are the underlying mechanism.

LVTSCRs may be further compacted and may gain speed as technology down-scaling progresses. However, in any given technology, the GGSCR will always be a more compact layout. Ln and Lp are only determined by the design rules for well overlap over diffusions while in case of the LVTSCR more dimensions are added.

## II.e. SCR Latch Engineering

In order to make SCRs useful ESD protection devices, latch-up must be avoided. Therefore, either the SCR holding current must be sufficiently high or the minimum holding voltage shall exceed the maximum operating voltage of the IC (VDD + 10%). The latter option is achieved by introducing series diodes, as demonstrated in Figure 12. This does not affect the performance (It2) of the SCR. There is no significant impact on the leakage when using up to 2 diodes and also for 3 diodes no impact is expected.

The increase of the on-resistance by series diodes having the same width as the SCR is found to be almost negligible.

It should be noted that for the holding voltage increase by series diodes we found an increase of 1.1V on the average per diode and not by only 0.7V as expected. We attribute this beneficial effect to the fact that any N-well diode features a P+ to N-well junction as well as an N-well to N+ junction. The latter junction has still a built-in voltage but smaller than in a normal P-N junction. Only hole-diffusion is necessary to establish equilibrium in a non-biased state. Under high current conditions, an additional voltage drop occurs across this junction.

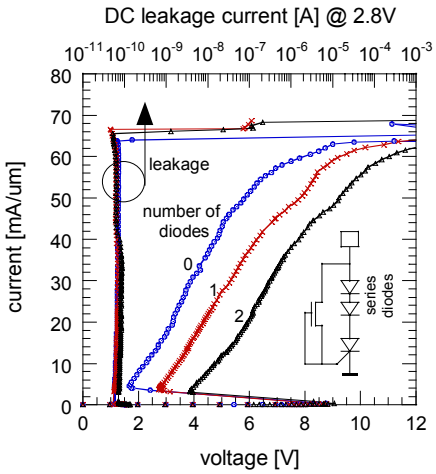


Figure 12 SCR latch control through insertion of series diodes.

## II.f. Anode-Cathode Spacing in STI Technologies

The performance of the SCRs is heavily dependent upon the bipolar capabilities of the NPN and PNP sub-structures of the SCR. In the GGSCRs discussed so far, the STI has been blocked in the center of the structure (cf. Figure 2). In the following we investigate what possible influence a presence of STI (cf. Figure 3) may have on the characteristics.

While in LOCOS processes a P-field implant was used as a channel stopper for the parasitic NMOS, a P-field implant is usually not required under the STI layer. With STI the doping is not locally increased which in turn does not degrade the lateral NPN current gain. Therefore, we may expect a functional SCR in deep sub-micron CMOS. However, the STI still replaces the lowly doped surface of the retrograde wells (Figure 13). The PNP and NPN current gains and the lateral device cross-section may be reduced.

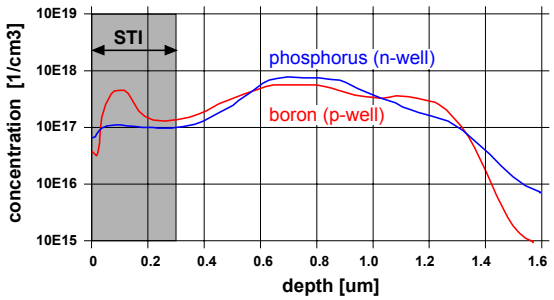


Figure 13 Well profiles (artist's rendering) are relevant for the current gain of the NPN and PNP device in the SCR. The presence of STI blocks the upper part of the profiles for SCR conduction, or if STI-blocked, these surface regions are enabled.

Experimental characterization of 0.18um and 0.13um CMOS technologies revealed that SCRs with STI ('STI-spaced SCR') do not loose their bipolar and high ESD performance, as illustrated in Figure 14. SCRs could be triggered without problem indicating a sufficiently high composite current gain of the NPN and PNP. Moreover the devices show an excellent clamping behavior and overall ESD performance of mostly above 60mA/um. Since the latter is determined by contact failure, a slight possible loss of internal device performance due to a pinched-off cross-section is negligible.

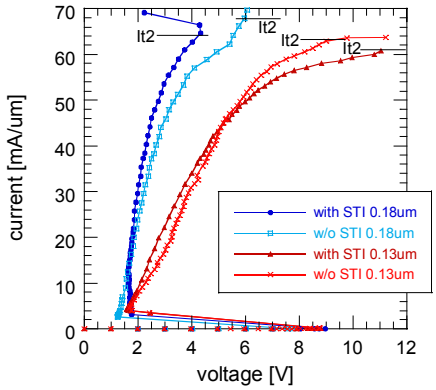


Figure 14 TLP I-V curves of various CMOS STI-spaced SCRs showing triggering by well-to-well breakdown or by a GGNMOS.

Another advantage of the GGSCR with STI is that the intrinsic SCR does not require silicide blocking. In combination with back-end-ballasted [11] NMOS trigger devices, fully silicided high performance ESD protection design can be realized.

## II.g. Technology Overview

ESD testing results from a large variety of deep sub-micron CMOS technologies are compared.

The results clearly demonstrate the universality of the GGSCR principle (Figure 15). Functional SCRs were obtained on both, bulk and epi substrates.

Linear width scaling of the ESD performance has been observed consistently for SCRs throughout all technologies. This easy scaling is a result of the double-injection and regenerative conduction mechanism in the SCR. Moreover, due to the extremely high performance, only much shorter finger widths are required. Also multi-finger arrays with SCRs have been demonstrated successfully by guaranteeing turn-on through electrical coupling of the SCRs.

Table 3 lists the normalized TLP-It2 per um width. Consistently excellent and scalable performance has been obtained for all GGSCRs. HBM test data ranged always well above the 100V/um range for GGSCRs while it was found to be in order of 70V/um for the LVTSCR. The SCRs with STI-block have proven functionality in all three investigated 0.18um technologies and the trend continues for 0.13um. No problems in SCRs with STI were observed either.

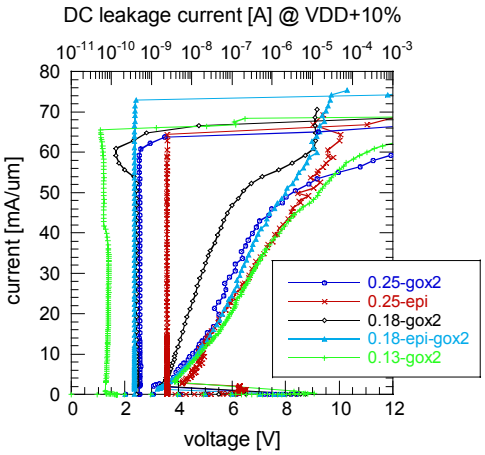


Figure 15 Typical IV characteristics and leakage evolution plots of the GGSCR demonstrate its operation in all investigated technologies.

Table 3 Overview on ESD test data from selected technologies. GGSCRs provide a consistent picture of functionality and high performance throughout a very wide range of technologies with both bulk and epi substrates.

| CMOS Process | 0.25 | 0.25 | 0.18 | 0.18 | 0.18 | 0.13 |
|---|---|---|---|---|---|---|
| | ESD Results TLP: It2 [mA/um] | | | | | |
| LVTSCR | n/a | 56 | 52 | 54 | 40 | n/a |
| GGSCR STI-block | 64 | 64 | 74 | 66 | n/a | 63 |
| GGSCR STI-spaced | n/a | n/a | 66 | 60 | 46 | 60 |

Furthermore, there is a general tendency visible that the LVTSCR does not provide the same high performance as seen for the GGSCR.

# III. SCR Protection

The oxide protection capability of GGSCRs is experimentally verified with gate oxide monitors (Figure 16). For comparison, large silicide-blocked multi-finger GGNMOS devices as reference with minimum gate length and thus shortest bipolar base width were also implemented with the same gate monitors.

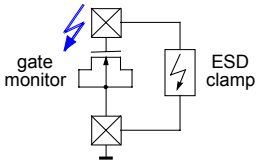


Figure 16 Gate monitor structure and ESD protection clamp. Two different clamp types are compared: GGSCR and GGNMOS.

## III.a. ESD Gate Oxide Reliability

### III.a.1 Transient Oxide Breakdown Measurements

Gate oxide reliability of deep sub-micron technologies is addressed by how the oxide may reduce the effective voltage level permissible during an ESD event. Stand-alone gate monitor structures are used for this purpose.

In Figure 17, we show time dependent dielectric breakdown (TDDB) measurements performed with 100ns TLP pulses. In the left plot, the quasi-static IV-characteristic is depicted. Around voltages of 11V, the tunneling current in the oxide (tox=6nm) becomes large enough that it can be resolved with the TLP (0.5mA resolution). The current increases further at higher voltage, while the DC leakage current (stress induced leakage current) is still as low as in the fresh device samples (right plot).

Oxide breakdown is defined at the transient voltage level where the stress induced leakage current abruptly increases. This occurs in the presented example between 14 and 15V with a very abrupt increase in leakage.

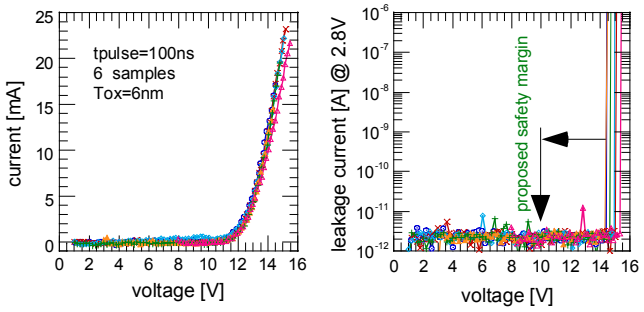


Figure 17 Time dependent dielectric breakdown (TDDB) measurements performed with 100ns TLP pulses for NMOS-type gate monitor structures. The left plot shows the IV-characteristics, the right depicts the leakage evolution as function of the pulse voltage.

To cover both intrinsic and extrinsic gate oxide failure a large safety margin must be employed to define the maximum permissible voltage limit under ESD conditions (suggested value: 10V). This value also corresponds very well to the value of 9.6V from 'static' TDDB measurements extrapolated to a time of 100us (thus 3 orders of magnitude longer than the 100ns TLP pulses).

It is also understood that it is impossible with this method to provide the full statistical significance for TDDB characteristics. However, with the gate monitor testing and the large safety margin we proposed an easy and pragmatic methodology for ESD protection engineering purposes.

### III.a.2 Gate Oxide Endurance Testing

To test the gate oxide reliability, endurance tests are conducted at the defined maximum permissible gate oxide voltage (10V ESD design rule) (Figure 18). The same type of samples were used as for the breakdown experiments (tox=6nm) reported in Figure 17. The endurance experiments show that the leakage current (left y-axis) does not increase even after 3000 zaps. The transient voltage response (right y-axis) has been also recorded as function of the number of pulses.

After this 'multi-zap' experiment, these endurance-stressed oxide samples were driven into breakdown and thus damaged as done before with the other samples. The gate oxide breakdown occurred as for the other samples above 14V. This confirms again the safe definition of the 10V ESD design rule for this 6nm oxide.

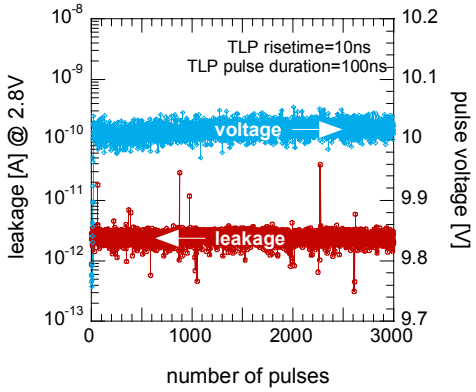


Figure 18 Gate oxide endurance testing. No increase in gate leakage is observed after 3000 TLP stress events with the defined ESD design limit voltage of 10V and a gate oxide tunneling current of 0.4mA.

## III.b. ESD-Protected Gate Monitor Structures

The TLP test results of Figure 19 give strong evidence that GGSCRs (Ln=Lp=0.8um) can protect a gate oxide equally effective than GGNMOS devices, while being over 8 times more width-efficient. The leakage current of the complete structures (protection + gate monitor) stays absolutely constant up to very high current levels. Neither of the two compared protection devices gives evidence of gate oxide degradation.

To provide a "best-case" for the GGNMOS, this device used a floating substrate to avoid multi-finger turn-on issues. This in turn led to a reduced trigger voltage. The GGSCR, however, uses a grounded substrate and therefore defines the most severe test condition: the trigger voltage is slightly higher than

the avalanche breakdown voltage.

Yet as discussed above, the GGSCR experiments give evidence of excellent ESD gate oxide protection capability. This finding is in perfect accordance with the transient characteristics of short GGSCR devices.

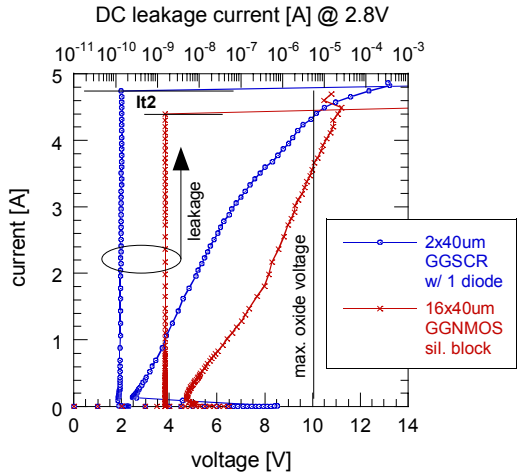


Figure 19 Gate monitors protected by GGSCR and large GGNMOS demonstrate both, very high failure currents while protecting successfully the gate oxide.

## III.c. Endurance Testing of ESD-Protected Gate Monitor Structures

The final experimental proof of suitability of the GGSCR for deep sub-micron CMOS technologies is provided by TLP endurance tests. The results of the gate monitors of the 0.13-um technology protected by GGSCR and GGNMOS (silicide blocked and fully silicided) is shown in Figure 20. The encircled areas indicate the large number (>1000) of only slightly scattered I-V-data points of this endurance test. The pulse current amplitude was kept constant below the failure level expected for each device type. The DC leakage current was monitored as a function of the number of pulses (Figure 21). The DC leakage current in these areas stayed absolutely constant over the duration of the endurance test (GGSCR and silicide blocked GGNMOS) while a slight shift was observed in the silicided GGNMOS. The silicide blocked GGNMOS was tested at a clamping voltage of 11V which was already slightly above the maximum defined oxide voltage of 10V, thus reconfirming the need for the safety margin. Finally, the regular TLP test was continued until regular device failure occurred. From the final failure value, it was determined that the endurance testing had been performed at 70 to 80% of It2.

Testing of several GGSCR samples revealed consistent results of stable and low leakage currents.

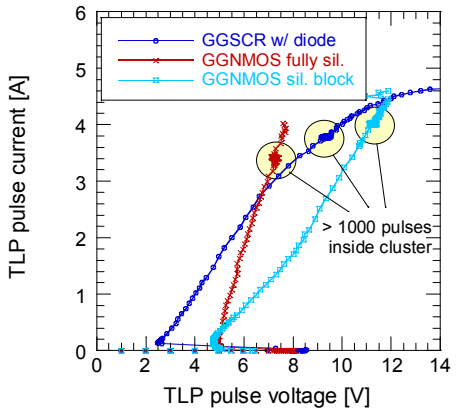


Figure 20 TLP endurance tests at very high current levels (encircled areas) with large number of repetitive pulses (>1000).

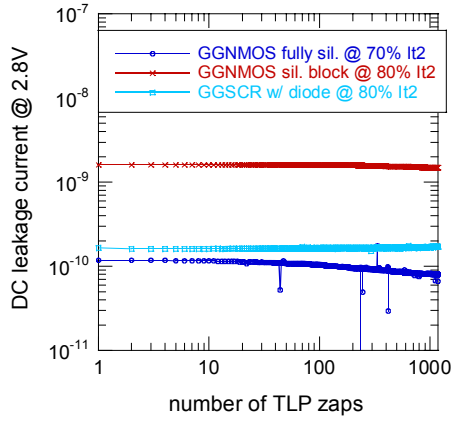


Figure 21 TLP endurance tests versus the number of stress pulses at 70 to 80% of It2.

## III.d. Endurance Testing of STI-SCRs

The STI-spaced SCRs were subjected to same severe endurance testing procedure to provide experimental proof that there is no long-term reliability concern in these devices.

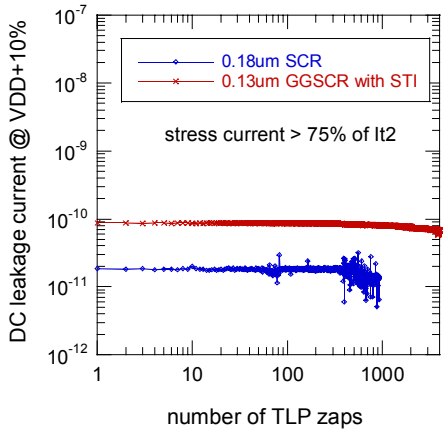


Figure 22 TLP endurance tests of STI-spaced SCRs demonstrate no leakage increase over a large number of stress pulses.

The data shown in Figure 22 report the leakage current of SCR samples from 0.18 and 0.13um CMOS technologies. Regular IV-measurements of those devices have been already shown in Figure 14.

Clearly, there is no leakage increase found even after an extremely large number of stress pulses. The SCR with STI remains absolutely stable.

# IV. SCR Applications

SCRs are becoming the workhorses for ESD protection in high performance product applications. They are used either to achieve extremely high protection levels or very low parasitic capacitance.

The presented GGSCR devices have been successfully implemented in:

- High volume applications, like e.g. Chip Card ICs, requiring high protection levels (6kV HBM) combined with the smallest area consumption.
- Pad limited designs with 80um pitch and single finger SCRs to provide more than 8kV HBM, 600V MM in 0.25um and below CMOS on epitaxial substrates
- Narrow pad pitch libraries with pitches of 50 or 40um by coupled SCR.
- 0.13um CMOS 3GHz regular and over-voltage tolerant RF pads with less than 120fF total added parasitic capacitance due to the ESD protection structures, with a capability of over 7kV HBM and over 350V MM.

SCRs are currently widely used in many commercial products worldwide. The presented GGSCR device has replaced other triggered SCRs in our advanced CMOS designs.

# V. Conclusions

This paper introduced a novel GGSCR device that has been silicon proven in a broad range of (sub)-quarter micron IC technologies.

The experiments demonstrate that with proper SCR design, the turn-on time is not an issue, because clearly, the thin gate oxides of deep-sub-micron technologies can be successfully protected.

The comparisons with GGNMOS-protected devices reveal that the fast GGSCR protection devices can compete with GGNMOS device in protecting the oxides while using much smaller silicon area consumption. Gate oxide endurance tests with both protection structure types also showed that the gate oxides are absolutely safe for repetitive stress with thousands of pulses.

Finally, it is reported that the GGSCR structure has been successfully implemented across a wide range of applications and advanced CMOS processes.

# Acknowledgements

The authors gratefully acknowledge their Sarnoff co-workers Dana Brock for administrative support, Bill Mayweather for management support, and Pradyumna Swain for technical discussions. Jon Barth and John Richner (Barth Electronics, Inc.) are cordially thanked for advanced TLP testing. The authors also acknowledge the critical review and valuable suggestions by the mentors of this paper Guido Groeseneken (IMEC) and Charvaka Duvvury (TI).

# References

[1] L. R. Avery, “Using SCRs as Transient Protection Structures in Integrated Circuits”, EOS/ESD 1983, pp. 177.

[2] A. Chatterjee, T. Polgreen, “A Low Voltage Triggering SCR for On-Chip ESD Protection at Output and Input Pads”, IEEE EDL-12, 1991, pp. 21.

[3] A. Amerasekera and C. Duvvury, *ESD in Silicon Integrated Circuits*, Wiley and Sons, 1995, ISBN 0 471 95481 0.

[4] C. Diaz, G. Motley “Bi-Modal Triggering for LVSCR ESD Protection Devices”, EOS/ESD 1994, pp. 106.

[5] J. Wu, P. Juliano, E. Rosenbaum, “Breakdown and Latent Damage of Ultra-Thin Gate Oxides under ESD Stress Conditions”, EOS/ESD 2000, pp. 287.

[6] M. Mergens, K. Verhaege, C. Russ, J. Armer, P. Jozwiak, G. Kolluri, L. Avery, “Multi-Finger Turn-on Circuits and Design Techniques for Enhanced ESD Performance and Width-Scaling”, to be published at EOS/ESD 2001.

[7] M.D. Ker, H. Chang, “How to Safely Apply the LVTSCR for CMOS Whole-Chip Protection without being Accidentally Triggered On”, EOS/ESD 1998, pp. 72.

[8] A. Wang, C.H. Tsay, “On a Dual Polarity On-Chip Electrostatic Discharge Protection Structure”, IEEE TED-48, No. 5, 2001, pp. 978.

[9] S. M. Sze, *Physics of Semiconductor Devices*, John Wiley, New York, 1981.

[10] H. Wolf, H. Gieser, W. Wilkening, “Analyzing the Switching Behavior of ESD-Protection Transistors by Very Fast Transmission Line Pulsing”, EOS/ESD 1999, pp. 28.

[11] K. Verhaege, C. Russ, “Wafer Cost Reduction through Design of High Performance Fully Silicided ESD Devices”, EOS/ESD 2000, pp. 287.

# About Sofics

Sofics (www.sofics.com) is the world leader in on-chip ESD protection. Its patented technology is proven in more than a thousand IC designs across all major foundries and process nodes. IC companies of all sizes rely on Sofics for off-the-shelf or custom-crafted solutions to protect overvoltage I/Os, other non-standard I/Os, and high-voltage ICs, including those that require system-level protection on the chip. Sofics technology produces smaller I/Os than any generic ESD configuration. It also permits twice the IC performance in high-frequency and high-speed applications. Sofics ESD solutions and service begin where the foundry design manual ends.

# Our service and support

Our business models include

- Single-use, multi-use or royalty bearing license for ESD clamps
- Services to customize ESD protection
  - Enable unique requirements for Latch-up, ESD, EOS
  - Layout, metallization and aspect ratio customization
  - Area, capacitance, leakage optimization
- Transfer of individual clamps to another target technology
- Develop custom ESD clamps for foundry or proprietary process
- Debugging and correcting an existing IC or IO
- ESD testing and analysis

# Notes



# Version

May 2011

**Sofics BVBA**
Groendreef 31
B-9880 Aalter, Belgium
(tel) +32-9-21-68-333
(fax) +32-9-37-46-846
bd@sofics.com
RPR 0472.687.037