A Substrate-and-Gate Triggering NMOS Device for High ESD Reliability in Deep Submicrometer Technology

Chih-Yao Huang1 and Fu-Chien Chiu2

1 Department of Electronics Engineering, Chien Hsin University of Science and Technology, No. 229 Chien Hsin Road, Zhongli, Taoyuan 320, Taiwan
2 Department of Electronic Engineering, Ming Chuan University, No. 5 De Ming Road, GuiShan, Taoyuan 333, Taiwan

Correspondence should be addressed to Chih-Yao Huang; huangcy@uch.edu.tw

Received 1 October 2013; Accepted 2 November 2013

Copyright © 2013 C.-Y. Huang and F.-C. Chiu. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

A substrate-and-gate triggering scheme which utilizes dynamic threshold principle is proposed for an ESD NMOS structure. This scheme enhances the device reliability performance in terms of higher second breakdown current and both reduced holding voltage/triggering voltage as well as elimination of gate over driven effect. The simple resistance and RC substrate-and-gate triggering NMOS structure with various resistance/capacitance values totally exhibit superior ESD reliability than the gate-grounded NMOS (GGNMOS) devices by 18∼29%. The substrate-and-gate triggering scheme in combination with special substrate pickup styles also shows excellent enhancement when compared with the GGNMOS cases of the same pickup styles. The substrate-and-gate triggering NMOS with butting substrate pickup style is better than the general butting case by 28∼30%, whereas the substrate-and-gate triggering NMOS with inserted substrate pickup style is 3.5 times superior to the general inserted case.

1. Introduction

Electrostatic discharge (ESD) issue has become a more serious device reliability problem in semiconductor components and systems. An NMOSFET has been the most popular ESD protection candidate for a long time. Since the shrinking of device size advances continually, the ESD capability of the NMOS device encounters more challenges [1–7]. A gate-grounded NMOS (GGNMOS) can no longer satisfy the ESD protection mission easily. ESD NMOS protection devices usually need the large width size to deal with ESD events. This results in multifinger layout style which is commonly used in practical IC I/O area. But it also has a critical drawback which is not favorable for the ESD protection requirement. The conduction current is usually unevenly distributed along the width direction of the multifingers. Gate-coupling technique using the property that increases the gate bias can reduce the first trigger point of the NMOS device and enable uniform ESD current distribution [4, 8, 9]. Although gate-coupling technique can improve the ESD capability, it still has gate overdriven effect if the gate voltage coupled is much larger than its threshold voltage, and this leads to serious ESD degradation. Furthermore, inserted or butting substrate pickups in the source diffusion region of the ESD NMOS device in deep submicrometer technology also degraded ESD reliability seriously. Such layout style has been strictly prohibited in practical ESD design applications by the technology design rules.

Therefore, in this work, a new substrate-and-gate triggering (SGT) structure that utilizes dynamic threshold MOS (DTMOS) technique based on the previous works [10, 11] is proposed to resolve the above problem. The simple resistance only or resistance-capacitance substrate-and-gate triggering NMOS (SGTNMOS) style in 0.18 μm CMOS technology has been investigated and compared with the GGNMOS in detail. Secondly, the SGTNMOS structures with butting or inserted substrate pickups were also compared with the corresponding GGNMOS pickup style conditions. The SGTNMOS scheme exhibits better performance and it can greatly improve the conventional pickup layout style. Finally, the simulation comparison between the gate-coupling and SGTNMOS case shows that the SGTNMOS has significant
Effect in eliminating the gate overdriven issue that often happens in the gate-coupling NMOS.

2. Substrate-and-Gate Triggering NMOS Scheme

Figure 1 shows the top view of an ESD NMOS device with multifinger layout style. The conduction current is usually unevenly distributed along the width direction of the multifingers. One possible cause of this problem is the nonuniform distribution of individual substrate resistance inherent in the real device layout and this causes two different conceptual devices M1 and M2 for the central part and the peripheral part originated from the layout as shown in the equivalent circuit of Figure 2(a). This equivalent circuit has been studied for nonuniform current distribution in the past [5, 7].

An NMOS protection structure which utilizes dynamic threshold MOS (DTMOS) concept is proposed in Figure 2(b) [10, 11]. Figure 2(b) is the equivalent circuit for the resistance SGTNMOS case. A quite small resistance can sufficiently sustain this improvement. This improvement can be attributed to the same current conduction supported by the emitter/base forward bias of the substrate/gate connection mechanism. The trigger point voltages $V_{t1}$ are suppressed to about 4.9 V and the holding voltage $V_{t2}$ to 3 V owing to the gate and substrate triggering connection. This implies that the parasitic BJT turns on more efficiently for the SGTNMOS cases than the GGNMOS.

The substrate connection part is further split into three different conditions: normal substrate pickup, butting substrate, and inserted substrate pickup layout styles. The normal substrate pickup layout connects the substrate triggering node only at the outer pickup diffusion ring around the NMOS body as in Figure 1. The butting substrate pickup layout connects to both the outer pickup ring and substrate pickups whose diffusions are shorted to the NMOS source diffusion regions as shown in Figure 2(c). The inserted pickup layout connects to both the outer pickup ring and the substrate pickup diffusions closely separated from the NMOS source diffusions with a small spacing as in Figure 2(d). These SGTNMOS structures were fabricated in standard 0.18 μm 1.8 V process technology. The main parameters of the NMOS ESD protection devices are the channel width/length $W/L = 360/0.18 \mu m$ with 24 poly gate fingers of 15 μm long and the drain contact-to-gate spacing is 1.72 μm with silicide blocking mask. The whole device structures were stressed by transmission line pulse (TLP) curve tracing. The TLP current pulse width is 100 ns with its rise time equal to 2.5 ns. The failure criterion is defined as the DC leakage current greater than 1 μA after the TLP stress.

3. Results and Discussion

Figures 3(a) and 3(b) show the TLP IV curve comparison among the GGNMOS and resistance SGTNMOS cases with the normal outer pickup ring, butting, and inserted pickups in the source diffusions. In Figure 3(a), the second breakdown current $I_{t2}$ is 3.04 A for the GGNMOS case and its IV curve clearly exhibits very severe nonuniform turn-on action. The failure currents $I_{t2}$ of the SGTNMOS conditions range from 3.59 to 3.93 A which have 18–25% improvement over the GGNMOS case. This quite small resistance can sufficiently sustain this improvement. This improvement can be attributed to more uniform current conduction supported by the emitter/base forward bias of the substrate/gate connection mechanism.

The substrate pickup connection part is further split into three different conditions: normal substrate pickup, butting substrate, and inserted substrate pickup layout styles. The normal substrate pickup layout connects the substrate triggering node only at the outer pickup diffusion ring around the NMOS body as in Figure 1. The butting substrate pickup layout connects to both the outer pickup ring and substrate pickups whose diffusions are shorted to the NMOS source diffusion regions as shown in Figure 2(c). The inserted pickup layout connects to both the outer pickup ring and the substrate pickup diffusions closely separated from the NMOS source diffusions with a small spacing as in Figure 2(d). These SGTNMOS structures were fabricated in standard 0.18 μm 1.8 V process technology. The main parameters of the NMOS ESD protection devices are the channel width/length $W/L = 360/0.18 \mu m$ with 24 poly gate fingers of 15 μm long and the drain contact-to-gate spacing is 1.72 μm with silicide blocking mask. The whole device structures were stressed by transmission line pulse (TLP) curve tracing. The TLP current pulse width is 100 ns with its rise time equal to 2.5 ns. The failure criterion is defined as the DC leakage current greater than 1 μA after the TLP stress.

3. Results and Discussion

Figures 3(a) and 3(b) show the TLP IV curve comparison among the GGNMOS and resistance SGTNMOS cases with the normal outer pickup ring, butting, and inserted pickups in the source diffusions. In Figure 3(a), the second breakdown current $I_{t2}$ is 3.04 A for the GGNMOS case and its IV curve clearly exhibits very severe nonuniform turn-on action. The failure currents $I_{t2}$ of the SGTNMOS conditions range from 3.59 to 3.93 A which have 18–25% improvement over the GGNMOS case. This quite small resistance can sufficiently sustain this improvement. This improvement can be attributed to more uniform current conduction supported by the emitter/base forward bias of the substrate/gate connection mechanism. The trigger point voltages $V_{t1}$ are suppressed to about 4.9 V and the holding voltage $V_{t2}$ to 3 V owing to the gate and substrate triggering connection. This implies that the parasitic BJT turns on more efficiently for the SGTNMOS cases than the GGNMOS. Figure 3(b) shows the TLP IV curve comparison among the resistance SGTNMOS and GGNMOS with the normal outer pickup ring, butting, and inserted pickups in the source diffusions. The resistance SGTNMOS with butting pickups exhibits an $I_{t2}$ of 3.78 A which is nearly the same as the resistance SGTNMOS with the normal pickup style. It is also greater than that of the normal pickup GGNMOS case by 24% and superior to the butting pickup GGNMOS by 29%, whereas the inserted pickup GGNMOS shows the worst $I_{t2}$ of 1.09 A. The $I_{t2}$ of 3.92 A of the resistance SGTNMOS with the inserted pickups is much larger than that of the inserted pickup GGNMOS by 3.6 times. Both the triggering and holding voltage of the inserted type SGTNMOS are also about 1 V lower than those of the normal GGNMOS. As a whole, R-SGTNMOS with the butting/inserted pickups in the sources can not only totally overcome the ESD degradation issue of the common inserted/butting pickup layout style, but also better than the normal pickup GGNMOS. These $I_{t2}$ comparisons are listed in Tables 1(a) and 1(c).

Next for the RC SGTNMOS part, Figures 4(a) and 4(b) compare the TLP IV curves among the RC SGTNMOS and GGNMOS cases with the normal outer pickup ring, butting, and inserted pickups in the source diffusions. As shown in Figure 4(a), the RC SGTNMOS with different $R_gC_g$ combinations all have the $I_{t2}$ values of 3.74–3.79 A which are better than that of the normal GGNMOS case by 23–25%. They also demonstrate better improvement for nonuniform current distribution issue. The capacitance and resistance values of all these $R_gC_g$ combinations are selected to have the same
Figure 2: (a) The equivalent circuit for the gate-grounded multifinger ESD devices. (b) The equivalent circuit for the resistance and RC multifinger substrate-and-gate triggering devices. (c) The layout diagram of the RC SGTNMOS device with butting pickups in the NMOS source diffusions. (d) The layout diagram of the RC SGTNMOS device with inserted pickups in the NMOS source diffusions.

Figure 3: (a) The TLP $I$-$V$ curve comparison among the GGNMOS and resistance SGTNMOS devices with normal pickup style. (b) The TLP $I$-$V$ curve comparison among the GGNMOS and resistance SGTNMOS with special pickup styles.
time constant around 200 ns. Further shown in Figure 4(b), the RC SGTNMOS with the butting/inserted pickups in the sources illustrates $I_{t2}$ of both 3.75 A. They are not only larger than that of the GGNMOS case by 23%, but also far superior to those with the butting/inserted pickups by 28% and 3.4 times, respectively. The $V_{t1}$ of the inserted SGTNMOS is reduced from 6.3 V to 4.1 V, and its $V_H$ is reduced from the normal 4 V to 3 V with the only exception of the butting SGT style. These results clearly demonstrate the effectiveness of the SGT scheme on ESD threshold improvement for the butting/inserted pickup layout style. These $I_{t2}$ comparisons are listed in Tables 1(b) and 1(c).

For the primary DTMOS operation, the gate and substrate nodes are coupled together. As the gate and substrate voltage increases, the threshold voltage decreases dynamically according to the body effect principle, resulting in a much higher current drive than the regular MOSFET. On the other hand, when the gate/substrate bias is zero the corresponding threshold voltage is still high as normal, so the leakage current is low during off state. With the gate and substrate shorting, the gate oxide voltage drop disappears and the voltage stress for hot carriers is eliminated. As far as ESD application is concerned, DTMOS concept has also been applied to SOI ESD protection design quite successfully [11]. When the gate and substrate are coupled to a high bias during ESD stress, this substrate acts as the base region of the parasitic BJT under the gate. The BJT turns on to conduct ESD current as the substrate bias exceeds base-emitter junction cutin voltage 0.7 V. In the gate-coupling part, the substrate body effect enables the threshold voltage reduction of the MOSFET, and larger drain current will be generated. Hence the avalanche part of the substrate current formation will also be enhanced and self-bias BJT turn-on performance of this scheme will be more robust. Both gate-coupling and substrate triggering together contribute to BJT turn-on action. This is reflected by $V_{t1}$ and $V_{t2}$ lowering phenomenon in the TLP $IV$ snapback curves. From the ESD point of view, the special feature of the SGT structure may be mainly dominated by the substrate triggering scheme and aided by the gate-coupling structure.

Next the discussion focuses on the $V_{t1}$ and $V_{t2}$ tendency. The SGT structure uses a resistance or resistance capacitance pair to trigger the parasitic BJT substrate. The $V_{t1}$ value is reduced by this triggering mechanism as compared with the GGNMOS $V_{t1}$ value 5.8 V. Figure 5(a) shows that all the resistance can reduce $V_{t1}$ to about 4.9 V in the R-SGTNMOS scheme. It also shows that the original butting/inserted pickup GGNMOS layout results in the highest $V_{t1}$ around

Table 1: (a) Resistance SGTNMOS $I_{t2}$ with different pickup styles. (b) RC SGTNMOS $I_{t2}$ with different pickup styles. (c) GGNMOS $I_{t2}$ with different pickup styles.

<table>
<thead>
<tr>
<th>Normal pickup</th>
<th>Butting RC-SGT</th>
<th>Inserted RC-SGT</th>
<th>GGNMOS</th>
</tr>
</thead>
<tbody>
<tr>
<td>$R_g$ ($\Omega$)</td>
<td>500</td>
<td>2 k</td>
<td>5 k</td>
</tr>
<tr>
<td>$R_{SGT}$ $I_{t2}$ (A)</td>
<td>3.8</td>
<td>3.8</td>
<td>3.59</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Pickup style</th>
<th>Normal RC-SGT $I_{t2}$ (A)</th>
<th>Butting</th>
<th>Inserted</th>
<th>(Ω)</th>
<th>1 k</th>
<th>10 k</th>
<th>1 k</th>
<th>10 k</th>
</tr>
</thead>
<tbody>
<tr>
<td>R-SGT $I_{t2}$ (A)</td>
<td>3.78</td>
<td>3.79</td>
<td>3.92</td>
<td>3.91</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Normal RC-SGT $I_{t2}$ (A)</th>
<th>3.75</th>
<th>3.74</th>
<th>3.79</th>
<th>3.77</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pickup style</td>
<td>Butting</td>
<td>Inserted</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RC-SGT $I_{t2}$ (A)</td>
<td>3.75</td>
<td>3.75</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$I_{t2}$ (A)</th>
<th>Normal</th>
<th>Butting</th>
<th>Inserted</th>
</tr>
</thead>
<tbody>
<tr>
<td>GGNMOS</td>
<td>3.04</td>
<td>2.92</td>
<td>1.09</td>
</tr>
</tbody>
</table>
6.3 V due to shorting effect between the substrate and source diffusion. Besides, the SGTNMOS with the inserted pickups can also settle its \( V_{t1} \) to 4.9 V effectively, whereas only that of the butting pickup style still remains at 6.3 V. Figure 5(b) further exhibits the role which the RC network plays on \( V_{t1} \) reduction. In this figure, the RC-SGT structure leads to a better \( V_{t1} \) lowering amount of 2 V. It is apparently more effective in reducing \( V_{t1} \) than the R-SGT structure by 1 V, because the gate capacitance is a more direct and controllable coupling link than the parasitic gate-drain overlap capacitance in the R-SGT structure. Furthermore, Figures 5(c) and 5(d) illustrate the \( V_{t1} \) tendency with respect to the resistance or resistance-capacitance values. It is noted that the capacitance and resistance values of all these \( R_gC_g \) combinations are selected to have the same time constant around 200 ns. Figure 5(c) shows that resistance SGT scheme can reduce the \( V_{t1} \) from 4 to 3 V with the only exception of the butting SGT style. In Figure 5(d), the RC-SGT structure can also reduce its \( V_{t1} \) smaller than the normal GGNMOS case by 1 V with still the only exception of the butting condition. Thus the SGT structure can much improve both the holding and trigger voltage characteristics of the NMOS device.

An RC network gate-driven ESD clamp circuit frequently leads to gate overdriven ESD degradation problem [9]. The electrothermal device simulation was further performed to investigate the operation details of the SGT structure. Figure 6(a) illustrates the simulated gate voltage transient response of an RC gate-coupling NMOS with a gate resistance 15 kΩ. A tendency is apparent that the gate voltage rises during 100 ns range, while the gate-coupling capacitance
increases. The gate voltage reaches 9.3 V corresponding to the RC time constant about 200 ns. The high voltage at the gate introduces a serious damage effect that is possibly due to overheating located under the gate channel and thus the ESD failure threshold degrades significantly and the primary purpose of the RC gate-coupling technique may not achieve. In contrast, the RC SGTMOS performs well in its gate voltage transient response with different gate capacitance parameters as shown in Figure 6(b). No matter the increase of the gate capacitance, the gate voltage always settles around 0.8 V during pulse rise time because the substrate connection can help eliminate the influence of the gate bias effect. The substrate-source forward bias effectively clamps the substrate/gate node at diode forward voltage drop and thus the SGT structure can eliminate the gate overdriven effect of the gate-coupling scheme effectively. The voltage rises to 2–10 V during the TLP pulse rise time transient but then settles down to around 0.8 V during pulse width period. This may be due to longer response time delay of the PN junction after the RC response. Hence, the SGTMOS scheme combines the advantages of the gate-coupling and substrate triggering styles for the performance enhancement of ESD NMOS devices.

4. Conclusions

A new substrate-and-gate triggering structure which combines the advantages of the gate-coupling and substrate triggering technique is presented. This SGT device structure has the advantages of reduced threshold voltage by the substrate bias, enhancement of drain current, and higher avalanche breakdown efficiency. This structure can maintain the minimum channel length with normal off-state leakage current and without hot-carrier effect concern either. The ESD robustness comparisons among the GGNNMOS and SGTMOS cases further show that this SGTMOS scheme with simple resistance/resistance-capacitance connection can relieve the nonuniform turn-on issue of the multifinger layout by 18–29%. Both the holding and triggering characteristics are also improved at the same time. It also has excellent effectiveness in resolving the butting/inserted pickup layout problem by 28% to 3.5 times, and more importantly, the gate overdriven ESD degradation effect can be eliminated by this SGTMOS structure.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

The authors would like to thank the National Chip Implementation Center of National Applied Research laboratories for providing the process information and testkey foundry, the National Center for High-Performance Computing for computer time and facilities, Mr. Guo-Xiang Su and Mr. Chou-Wei Chang for their hard work in layout design/measurement/simulation study, and finally the research grant from the National Science Council of Taiwan government under Contract NSC 102-2221-E-130-015-MY2.

References


