

行政院國家科學委員會專題研究計畫 成果報告

超薄氮化矽極氧化層奈米 CMOS 元件可靠性的新方法研究

計畫類別：個別型計畫

計畫編號：NSC91-2215-E-009-040-

執行期間：91年08月01日至92年07月31日

執行單位：國立交通大學電子工程學系

計畫主持人：莊紹勳

報告類型：精簡報告

報告附件：出席國際會議研究心得報告及發表論文

處理方式：本計畫可公開查詢

中 華 民 國 92 年 11 月 17 日

超薄氮化閘極氧化層奈米 CMOS 元件可靠性的新方法研究

New Methodologies for Reliability Evaluation of Ultra-Thin Nitrided Gate Oxide Nano-CMOS Devices

計畫編號：NSC92-2215-E009-040
執行期限：91 年 08 月 01 日至 92 年 07 月 31 日
主持人：莊紹勳 國立交通大學電子工程學系

一、中文摘要

美國 ITRS (SIA) Roadmap 指出,到了西元 2005 年元件通道長度將進展到 75nm,而閘極氧化層厚度將縮為 10~15Å。因此,如何成長高品質的 10~15Å 氧化層是一大關鍵,尤其是可靠性的量測技術,相當欠缺。截至目前,有幾種量測氧化層可靠性的方法,如 DCIV, Gated-Diode (GD)、charge pumping(CP)等。然而,它們受限於量測時導致大量的閘極穿隧漏電流 (gate leakage tunneling current)而產生嚴重誤差。本計劃即著眼於此,將開發出適用於 SION 超薄閘極氧化層結構奈米 CMOS 元件可靠性分析的二種新方法。

第一種方法是改良的電荷幫浦法,稱為差頻電荷幫浦法 (IFCP, Incremental Frequency Charge Pumping),它可量得閘極氧化層厚度 10 至 20Å 元件的界面陷阱。它亦同時可以應用於量測各種不同成長方式閘極氧化層的品質,如 RTNO 及 RPN 等。是一種良好的監測閘極氧化層品質的方法。

第二種方法是改良的閘極二極體方式,稱為低漏電閘極二極體方法 (L^2 -GD, Low Leakage Gated-Diode)。它可量出延著通道的氧化層陷阱及電荷,以電流表示之。進而,可用於研究氧化層氮化,對 p-MOSFET 元件負偏壓高溫度之不穩定性(NBTI)特性的探討。這二種方法,均已成功運用於 10-20Å 氧化層的次 100 奈米 CMOS 元件的量測上。

關鍵詞：超薄閘極氧化層、量子穿隧效應、閘極漏電流、氧化層界面缺陷、氧化層陷阱、奈米 CMOS 元件技術、負偏壓高溫度

之不穩定性、電荷幫浦技術、閘極二極體方法。

Abstract

As device scaling continues, the sub-100nm CMOS device needs a t_{ox} in the range of 10-15Å and with 75nm gate length in 2005, as predicted from the SIA roadmap. How to monitor oxide quality for ultra-thin gate oxide is crucial for the next generation CMOS technology, in particular the monitoring of interface traps (N_{it}) and oxide trapped charges (Q_{ot}) in the gate oxide. So far, various gate oxide reliability diagnostic tools, such as DCIV, GD(Gated-Diode), CP(Charge-Pumping) etc. have been employed for such a purpose. For ultra-thin gate oxide of 20Å and beyond, the above methods are limited by the tunneling leakage through the gate oxide during the measurement since direct tunneling exists.

This project has been focused on developing new techniques for the measurement of ultra-thin gate oxide sub-100nm CMOS devices. We have successfully developed two methods, a new charge pumping method and an improved gated-diode measurement technique.

This first approach is called Incremental Frequency Charge Pumping (IFCP) method. It has been implemented for determining the interface traps in ultra-short gate length CMOS devices with ultra-thin gate oxide in the direct tunneling regime, 12-16Å gate oxide. The leakage current in the gate oxide can be removed from the measured CP current, which enables accurate determination of the interface traps. This method has been demonstrated successfully for various RTNO grown and RPN treated oxide CMOS devices with very thin gate oxide. Moreover, it can be used as a good monitor of ultra-thin gate oxide process.

The second approach is called Low gate Leakage Gate-Diode (L^2 -GD) method, in which the distribution of interface traps along the channel, can be easily obtained from this L^2 -GD method. This method has been demonstrated successfully for the ultra-thin (10-20Å) gate oxide device. Also, by using this new technique, the localized oxide damage due to NBTI stress effect can be identified simply from the measured drain currents. Therefore, both types of approaches, IFCP and L^2 -GD techniques are well

suitable for the characterization of CMOS reliabilities, and in particular for the sub-100nm CMOS devices with gate oxide in the range 10-20Å.

Keywords: ultra-thin gate oxide, quantum tunneling effect, gate leakage current, oxide interface traps, oxide traps, nano-CMOS device technology, NBTI, charge pumping technique, Gated-diode method.

二、計劃緣由與目的

CMOS 元件製程技術發展迄今，面臨到一個重大的可靠性課題，亦即隨著元件尺寸縮小，元件閘氧化層跟著微縮 (scaling)，美國 ITRS (SIA) Roadmap 指出，到了西元 2005 年元件通道長度將進展到 75nm，而閘氧化層厚度將縮為 10~15Å。此一微縮的結果，使得半導體製程中，最關鍵的閘氧化層技術，變得更加困難，因為傳統閘氧化層存在大的閘極穿隧漏電流 (gate leakage tunneling current)。因此，傳統閘氧化層 (oxide) scaling 以後，如何降低 gate leakage current，成為 CMOS 技術首先要克服的問題。尤其，近年來的諸多研究顯示，要降低這個量子穿隧電流，製程上有二種途徑，例如，採用 high-K 介電質，提高 Physical Oxide 厚度卻獲得較小的 EOT [1-3]，或是高劑量的氮化製程以增加 k 值來降低 leakage current 等 [4-5]。在現階段，high-K 閘氧化層材料特性以及它的製程相容性、材料的穩定性等問題，多屬於研究階段，所以遲遲未能成為工業生產的主流技術。

另一方面，我們則看好傳統的 SiO₂ 技術加上 nitridation 的 SION 結構，將可能成為工業界的主流技術。隨著 CMOS 元件技術的進步，我們已到了 sub-100nm 及更短 channel length 的技術層次，在 20Å 至 30Å 之間的純 SiO₂ 閘極氧化層仍是工業界的主流技術，10Å 至 20Å 的研究，係以 SION 較為可行，其中高劑量氮化製程的使用 [4-5]，有多項優點：如 SiO₂ Radical-Nitridation [5] 或 Remote Plasma Nitrided (RPN) oxide [6] 技術來降低 gate leakage current，抑制 Boron penetration，

提高 ON current 及降低 NBTI (Negative Bias Temperature Instability) [7] 等特性。這當中，Nitrogen 存在於 Si-SiO₂ 界面或氧化層裡面，造成氧化層界面缺陷，Interface trap (N_{it}) 及氧化層陷阱，oxide traps (Q_{ot})，是這類技術成敗的主要關鍵。然而，oxide 的 quality 及可靠性的分析方法，如傳統的 TDDB, CV, lifetime 量測，大致上都有人研究過，唯獨欠缺對於氧化層界面及方法的研究，如 N_{it} 及 Q_{ot} 的 profiling 等，因此對於 Nitrogen 存在於界面，其物理機制 (mechanism) 及它如何影響元件 performance and reliability，則尚待深入探討。

近年來，本研究群在氧化層界面及其可靠性的研究方法 [8-18] (N_{it} 及 Q_{ot} 的 2D profiling)，有一系列探討，在沒有 gate leakage tunneling current 的情況下，大致已發展完成，這當中包含 charge pumping (CP) 及 gated-diode (GD) 量測方法，這些研究，對於 t_{ox} = 25Å 或 30Å 以上的元件可靠性分析，都足以應付。

基本上，對於超薄氧化層 (20Å 以下或更薄) 的傷害及缺陷分析的相關研究上，由於量子化效應 (Quantum-Mechanical Effect) [19] 及量測時出現 gate leakage current，這使得上述 CP 及 GD 傳統分析方法無法再使用，蓋因，閘極漏電流 (gate leakage tunneling current) 已與量得的 charge pumping current (Substrate current) 或 gated-diode measured current 的大小 (order) 相當，使得閘氧化層的量測及理論分析上，變得更加複雜，而需要大翻修。

本計劃即著眼於此，發展下一代 sub-100nm channel length 具有超薄氧化層 CMOS 元件的可靠性分析方法研究，這當中結合了 SION 超薄氧化層 CMOS 元件為測試樣本，以我們已發展的 CP 及 GD 傳統分析方法為基礎，嘗試改善原有及發展新的方法，用於研究 sub-100nm SION 超薄氧化層 CMOS 元件的可靠性研究，以尋找出適合下一代元件使用之超薄閘氧化層

可靠性新的分析方法。

三、結果與討論

本計劃重新修正我們之前已發展完成的氧化層界面分析方法，開發出適用於 SION 超薄閘氧化層結構奈米 CMOS 元件可靠性分析的二種新方法。

第一種方法是改良的電荷幫浦法，稱為差頻電荷幫浦法 (IFCP, Incremental Frequency Charge Pumping)[20]。第二種方法是改良的閘二極體方式，稱為低漏電閘二極體方法 (L^2 -GD, Low Leakage Gated-Diode)[21]。各項結果分述如下：

(1) 目前電荷幫浦技術的發展

基本上,電荷幫浦技術主要是被使用於求延著通道方向 N_{it} 和 Q_{ot} 的 profilng。元件的可靠性和這些界面缺陷和氧化層陷阱的關係可以被決定。當氧化層厚度持續的縮減，漏電流的量隨之增加，將導致在超薄氧化層量測上的錯誤。我們發展出一種新的電荷幫浦方法，可以使用於當氧化層厚度小於 12\AA 時的量測，這新的方法稱做增量頻率電荷幫浦法(IFCP)，適用於超薄閘氧化層可靠性分析量測上。

A. 差頻電荷幫浦(IFCP)的原理

Fig. 1(a)顯示電荷幫浦法的設定和發生在超薄氧化層中閘極(Gate)到基體(Bulk)的漏電成分 I_{GB} 。Fig. 1(b)顯示應用固定 base-level 的閘極脈衝來量出 CP 電流。Fig. 1(c)顯示正常 ($t_{ox} > 30\text{\AA}$) 和異常 ($t_{ox} < 30\text{\AA}$) CP 電流，可以看到當 $t_{ox} < 30\text{\AA}$ 時,穿隧電流發生在閘極和基部之間。這個漏電成分會影響計算界面缺陷量的正確性。

Fig. 1(d)顯示差頻電荷幫浦法(IFCP)的步驟。我們使用這個方法是以兩種不同的頻率來量 CP 電流。舉例來說，首先，我們測量在兩個不同頻率($f_1 = 2\text{MHz}$ and $f_2 = 1\text{MHz}$)下的 I_{cp} 如 Fig. 2，從這兩者我

們可以得到 I_{cp} 的差額。這個 I_{cp} 變成在一個新頻率($f_1 - f_2 = 1\text{MHz}$)下 I_{cp} 的值。Table I 列出這個 CP 技術的方程式，正確的 CP 電流值可以從兩個不同頻率下所得到的 CP 電流相減，來去除漏電的成份。Fig. 3 顯示在不同頻率下量測到的 CP 曲線，這些測量曲線可以乘上任意的係數，舉例來說， $I_{cp,1M-500KHz}$ 和 $5/2 * I_{cp,1M-800KHz}$ 都被視為在頻率 500KHz 下的正確 I_{cp} 值

B. 超薄氧化層的特性測量

對於其他的應用,這個 CP 技術已經被使用在估計以 RTNO 和 RPN 處理的氧化層中的界面缺陷。在這次研究中所用的這些元件是以最先進的 IC 製程所製造的。 $12\sim 16\text{\AA}$ 氧化層是以 RTNO 來形成的。從這些量測的 CP 資料和 Table I 底部的方程式 3(a)和 3(b)，我們可以對 RTNO 和 RPN 處理過的氧化層來計算界面缺陷，如 Fig. 4，可以知道當斜率越陡，氧化層的品質越差。結果可以顯示出(1)RTN 處理後的氧化層比 RTNO 處理過的氧化層有更多的界面缺陷，因為在底部的 base oxide(RTNO)在熱處理時，有重新氧化的機制，顯示在 Fig. 5 (2)pMOSFET's 的氧化層品質並沒有和 nMOSFET's 的氧化層一樣好，因為它產生較多的界面缺陷。

總之，在超薄氧化層的可靠性量測上，IFCP 可以適用在小面積的元件。此外，在量測時，漏電流的成分所造成的錯誤可以被抑制。這項技術非常適用於 CMOS 元件可靠性分析上和目前奈米 CMOS 元件上氧化層品質的監控。

(2) Gated-Diode(GD)的測量技巧

GD 這個量測方法是量測汲極端的電流，傳統的 GD 被限制在厚氧化層($> 30\text{\AA}$)的條件下量測汲極電流的第一峰值，因此我們需要改進量測方法以應用在超薄閘極氧化層的元件量測。為了克服薄氧化層導致的漏電問題，一個新的方法稱為低漏電閘二極體方法，(L^2 -GD, Low Leakage

Gated-Diode)。它是利用在汲汲極源極加上較高的順偏偏壓而防止閘極到源極和汲極的邊緣穿隧電流，而傳統的 GD 是將小於 0.3V 的偏壓加在汲極和 N 型井的接面，所以在薄氧化層會有漏電問題(Fig. 6)。Fig. 7 為延著通道可以量到的 gated-diode 電流，這些電流表示 interface trap 的分佈量。Fig. 8 可看出當汲極順向偏壓太小時，邊緣穿隧電流會因閘極偏壓增大時而愈明顯，因而將觀測不出任何峰值。因此增加汲極偏壓的結果：(1)降低閘及到汲極的穿隧電流(2)增加汲極順偏偏壓可明顯觀測出 GD 電流。Fig.7 則量出第二、三 peak 值，分別落於 drain/bulk 附進及 gate-drain 邊緣。

A. L^2 -GD 在可靠性分析的應用

為了證明新的 L^2 -GD 這項方法的實用性，我們先找到 $I_{G,max}$, $I_{B,max}$ 及 $V_G = V_D$ 三種 stress 偏壓下，元件退化的比較(Fig. 10)。Fig. 11 用來證明 $V_G = V_D$ stress 情形下，pMOSFET 有最大的 I_D 衰退。我們發現，在 drain junction 附進，第二 peak 的增加量是 $V_G = V_D$ stress 條件下有最大退化量的原因。

B. L^2 -GD 在 NBTI 研究的應用[22]

於高密度積體電路及較高功率消耗，溫度效應在以雙多晶閘極互補式金氧半場效電晶體(dual poly-gate CMOS)技術，PMOS 已成為主要的可靠性問題。在高溫下並在閘極加負偏壓，也就是 NBTI，是造成可靠性問題最主要的效應。這效應會造成元件起始電壓往負電壓方向飄移，且會造成氧化層電荷及介面缺陷密度的增加而使次臨界區斜率增大。以下的例子指出 NBTI 如何導致 PMOS 的傷害。

Fig.12 利用 Gated-Diode 顯示出在溫度攝氏 125 度， $V_G = -3.3V$ 加壓下 PMOS 因 NBTI 造成的衰減。汲極電流分別顯示出通道區(第一峰值，Fig.11(a))和接面/延展區(第二/第三峰值，Fig.11(b))。這張圖中，我們可以看出：(1) NBTI 效應導致介面缺

陷及氧化層電荷(Q_{ot})的產生，(2) 第一個峰值主要是由在通道產生的 Q_{ot} 所產生，而我們可以看出第一峰值隨著加壓時間增加而往左移動，(3)第二峰值指出在 P 型汲極延展區的界面缺陷密度，我們發現第二峰值隨著加壓時間增加的增加量遠較第一峰值的增加量為高，因此我們相信在 P 型汲極延展區中的介面缺陷是造成 NBTI 主要的機制。這是歸因於閘極-汲極重疊區域的邊緣有硼的堆積。

Fig.13 分別顯示出在源極和汲極端受 NBTI-like 熱載子(在高溫加壓於 $V_G = V_D$)所造成的衰減。這個測試不同於 NBTI 之處是在於在源極處加一偏壓。如 Fig. 13(b)所示，我們認為 NBTI-like 是 NBTI 和熱載子效應的結合。由 Fig. 13 結果所示，在源極端只有因 NBTI 造成的接面傷害(第二峰值)，然而在汲極端出現了因 NBTI 和熱載子效應所造成接面和延展區的傷害。這些結果在 Fig. 13 中為：(1)NBTI 在區域 1 產生的 Q_{ot} 及區域 2 產生的 N_{it} 造成汲極電流的衰減，(2)在 NBTI 和熱載子的雙重效應下會加乘汲極電流衰減的現象。

C. 結果與討論

由前述的結果可知，高濃度電漿氮化製程對 0.1 微米以下超薄氧化層元件的設計上，仍有諸多的缺點。即使元件的漏電流大幅度的被降低外，但是付出的代價除了因為含了 Si_3N_4 多氧化層缺陷的材料特性外，另外也造成晶圓生產上因為電漿製程傷害造成的良率/品質/電特性的分佈不均；如何開發出一套新的高濃度電漿氮化製程設備，可以有效的降低不勻稱的結果，將會是未來半導體設備的研發重點，也是未來 0.1 微米以下元件設計之關鍵技術。

四、計劃成果自評

現階段，sub-100nm 及更短 channel length 的 CMOS 技術，high-K 閘氧化層其材料特性以及它的製程相容性、材料的穩定性等問題，多屬於研究階段，所以遲

未能成為工業界的主流技術。我們則看好傳統的 SiO₂ 技術加上 nitridation 的 SION 結構較為可行，將可能成為工業界的主流技術。

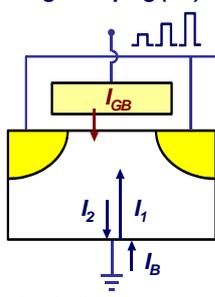
本計劃之完成，發展出一套適合 sub-100nm 超薄閘氧化層分析方法，用以提供未來 (0.1 微米以下) CMOS 元件閘氧化層技術的可靠性分析，以建立下一代元件技術採用 SION gate dielectric CMOS 元件之設計準則。

研究的成果已發表一篇 VLSI、一篇 IRPS 及一篇 IEDM 論文[20-22]，二項國內外專利審核中。而在工業界，可應用於最新 0.1 微米以下之新世代 CMOS 元件之可靠性設計分析上，有實質的經濟效益。所以，本計劃的完成，將對學術研究水準的提昇及工業界的實際應用上都有莫大的助益。

五、參考文獻

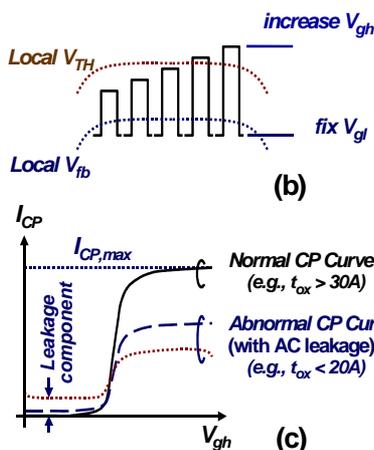
- [1] K. Yoshikawa, in *IEDM Tech. Digest*, pp. 11-14, 2000.
- [2] H.-J. Cho et al., in *IEDM Tech. Digest*, pp. 655-658, 2001.
- [3] X. Guo et al., in *IEDM Tech. Digest*, pp. 137-140, 1999.
- [4] K. Eriguchi et al., in *IEDM Tech. Digest*, pp. 323-326, 1999.
- [5] M. Togo et al., in *Symposium on VLSI Tech.*, pp. 81-82, 2001.
- [6] M. Rodder et al., in *IEDM Tech. Digest*, pp. 623-626, 1998.
- [7] N. Kimizuka et al., in *Symposium on VLSI Tech.*, pp. 73-74, 1999.
- [8] S. S. Chung et al., in *Tech. Digest of International Electron Device Meeting (IEDM)*, pp. 295-298, 1997.
- [9] S. S. Chung et al., in *Symposium on VLSI Tech.*, pp. 111-112, 1997.
- [10] S. S. Chung et al., in *Symposium on VLSI Tech.*, pp. 19-20, 1999.
- [11] S. S. Chung et al., in *Symposium on VLSI Tech.*, pp. 103-104, 1995.
- [12] S. S. Chung et al., *Proc. of International Reliability Physics Symposium (IRPS)*, pp. 249-252, 1999.
- [13] S. S. Chung et al., *Proc. of International Reliability Physics Symposium (IRPS)*, pp. 389-393, 2000.
- [14] S. S. Chung et al., *Proc. of International Reliability Physics Symposium (IRPS)*, pp. 419-424, 2001.
- [15] S. S. Chung et al., *Proc. of SSDM*, pp. 841-843, 1993.
- [16] S. -J. Chen et al., *Proc. of SSDM*, pp. 16-17, 2000.
- [17] S. S. Chung et al., *IEEE Trans. on Electron Devices*, Vol. ED-46, pp. 1371-1377, 1999.
- [18] G. H. Lee et al., *IEEE Trans. on Electron Devices*, Vol. ED-43, pp. 898-903, 1996.
- [19] C. Choi et al., in *Symposium on VLSI Tech.*, pp. 63-64, 1999.
- [20] S. S. Chung et al., in *Symposium on VLSI Tech.*, pp. 74-75, 2002.
- [21] S. S. Chung et al, *IEDM Tech. Digest*, pp.513-516, 2002.
- [22] S. J. Cheng et al, *Proc. IRPS*, pp. 203-207, 2003.

• Charge Pumping (CP) Setup



- $I_B = I_1 - I_2$
= Recombination Current
= CP Current (I_{CP})
- I_{GB} = Gate-Bulk Leakage
- Equation for extraction N_{it}
 $I_{CP,max} = f q W L N_{it}$

• Gate Pulses for the Fixed Base Level CP Method



• Flow Chart of the Low Leakage CP Methodology

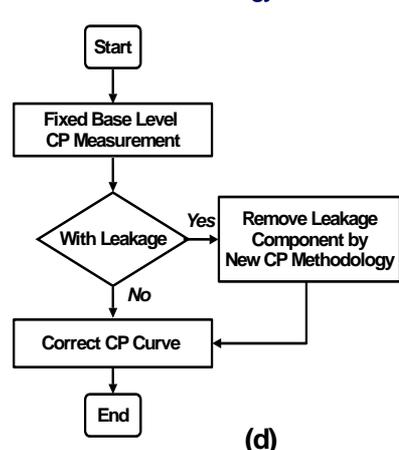


Fig. 1 (a) Charge pumping (CP) setup and leakage component in CP measurement. (b) Local threshold voltage (V_{TH}) and local flat band (V_{fb}) distribution in relating to low level and high level gate pulse (V_{gl} and V_{gh}). (c) Normal (without leakage) and abnormal (with leakage) CP curves. (d) The flow chart of the IFCP technique.

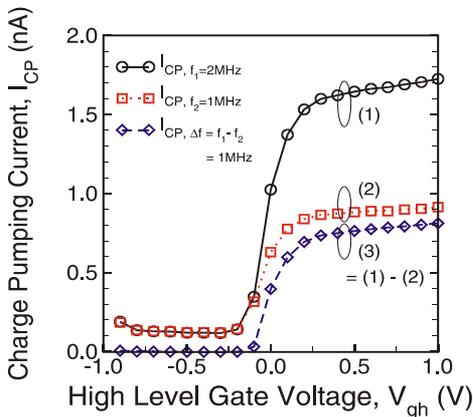


Fig. 2 Incremental charge pumping (CP) methodology- Using with-leakage CP curves, $I_{CP,f1}$ and $I_{CP,f2}$, to obtain a correct CP curve, $I_{CP,\Delta f=f_1-f_2}$.

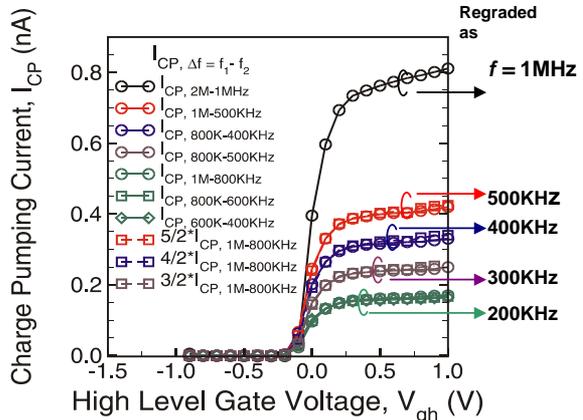


Fig. 3 Corrected CP curves for the incremental CP technique. Note that these CP curves have linear relationship between the $\Delta f = f_1 - f_2$, and can be multiplied by an algebraic factor.

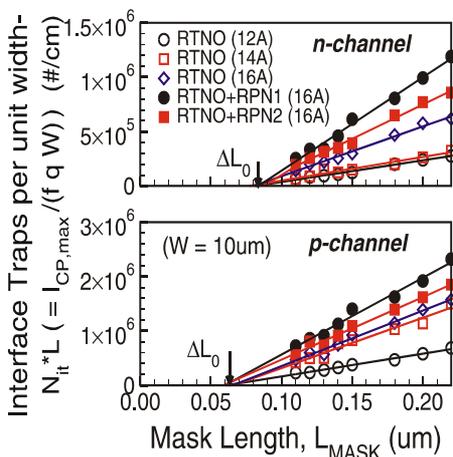


Fig. 4 Calculated N_{it} from $I_{CP,MAX}$ in Fig. 3.

Gate Oxide	Type	N_{it} (#/cm ²)
RTNO 16A	nMOSFET	$\sim 3.86 \times 10^{10}$
	pMOSFET	$\sim 7.37 \times 10^{10}$
RTNO 16A + RPN	nMOSFET	$\sim 5.86 \times 10^{10}$
	pMOSFET	$\sim 9.49 \times 10^{10}$

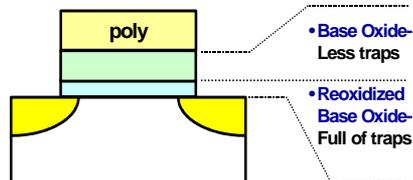


Fig. 5 (b) RPN treated RTNO gate oxide. The bottom layer has been reoxidized which causes more oxide traps.

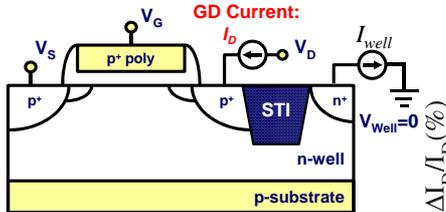


Fig. 6 The schematic diagram of the GD measurement method. Note that the drain current I_D is the so called gated-diode current.

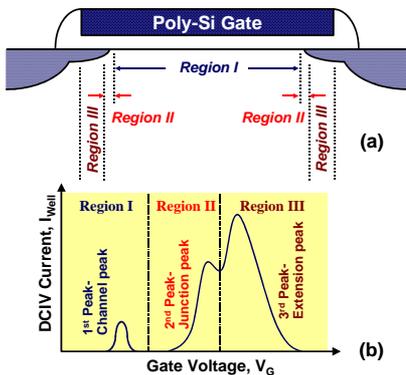


Fig. 7 Three peaks of drain current as a measure of N_{it} or Q_{it} from GD measurement. (a) Three peak regions. (b) Three-peak curves.

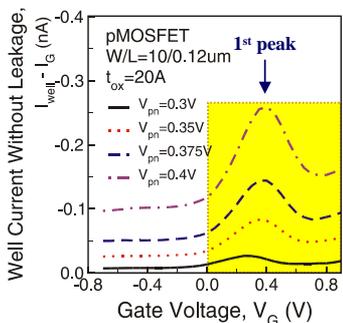


Fig. 8 An L^2 -GD method by raising the drain bias to see the first peak of the measured gated-diode current.

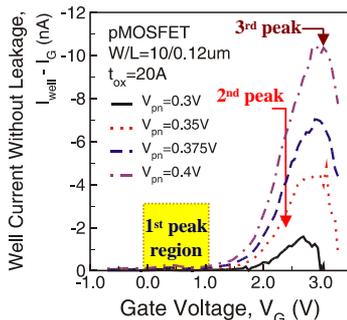


Fig. 9 An L^2 -GD method by raising the drain bias to see the second and third peaks of the measured gated-diode currents.

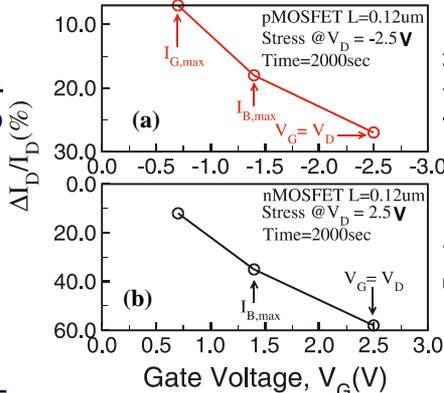


Fig. 10 Drain Current (I_D) degradation after 2000sec HC stress at different gate biases. (a) pMOSFET. (b) nMOSFET.

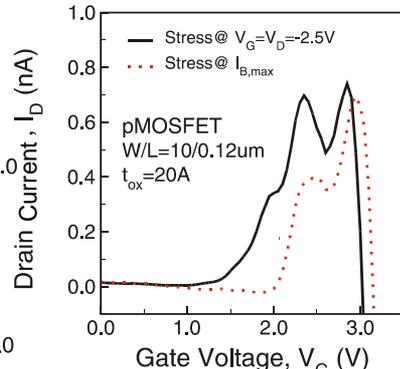


Fig. 11 Comparison of the measured gated-diode currents at $V_D = 0.325V$ for pMOSFET's after 2000 second $I_{B,max}$ and $V_G=V_D$ stresses.

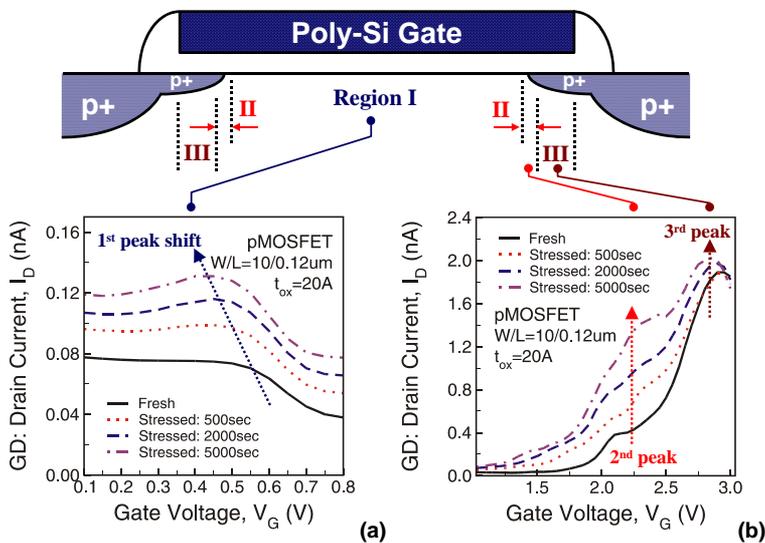


Fig. 12 Time dependence of L^2 -GD curves under NBTI stress ($V_G = -3.3V$, $T = 125^\circ C$). (a) Measured results in the channel region. (b) Measured results on the drain-side.

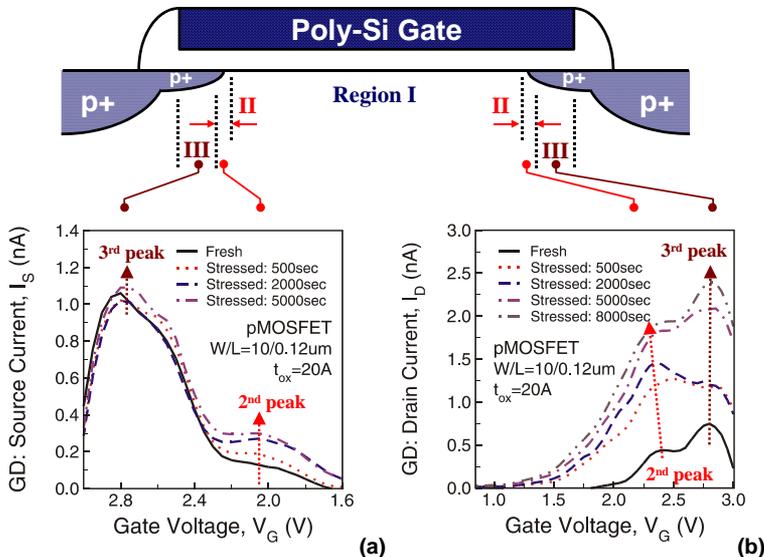


Fig. 13 Time dependence of L^2 -GD curves under NBTI-like HC stress ($V_G=V_D = -2.5V$, $T=125^\circ C$). (a) Measured results on the source-side. (b) Measured results on the drain-side. Note that the NBTI and HC induced oxide damage is asymmetrical.