Extended Abstracts of the 2009 International Conference on Solid State Devices and Materials, Sendai, 2009, pp789-790

Effect of Post Cap-Layer Deposition Annealing Temperature and TiN Thickness on SMDH CMOS Process using TiN Hard Mask

Hirofumi Shinohara, Akira Katakami, Toshinari Watanabe, Manabu Hayashi, Satoshi Kamiyama, Yoshihiro Sugita, Takeo Matsuki, Takahisa Eimori, Jiro Yugami, Kazuto Ikeda, and Yuzuru Ohji

Semiconductor Leading Edge Technologies, Inc. (Selete), 16-1 Onogawa, Tsukuba, Ibaraki, 305-8569, Japan
Phone: +81-29-849-1608, Fax: +81-29-849-1186, E-mail: shinohara.hirofumi@selete.co.jp

1. Introduction

Gate first MG/HK (Metal gate/ high-k dielectric) CMOS processes to achieve lower Vth have been studied intensively for low cost LSTP and LOP devices in 32nm-node technology [1-3]. Although the Single Metal Dual High-k (SMDH) structure is an attractive candidate for MG/HK CMOS because of its process friendly structure, establishing a process for fabricating dual high-k gate dielectrics to independently control the Vth of n- and p-MOSFET is still a problem. N. Mise et al. and T. Schram et al. proposed SMDH CMOS processes using a dual hard or soft mask. In these processes, two lithography processes are needed to fabricate the dual high-k dielectrics [4,5]. In this paper, we demonstrated a gate-first process with which a SMDH structure can be fabricated using a TiN hard mask just once. The dual high-k gate dielectrics were fabricated by using La_2O_3 and Al_2O_3 as capping layers on HfO_2. We found that the higher Post Cap-layer Deposition Annealing (PCA) temperature is effective to obtain lower |Vth|. Furthermore, a thicker TiN hard mask has the benefit of suppressing an increase in EOT and in lowering |Vth|. We also confirmed that the delay time (tpd) measured using a ring oscillator decreased for dual high-k dielectrics with higher PCA temperature.

2. Experimental

Fig.1 shows proposed SMDH CMOS process in which (a) the P-Cap(Al_2O_3) is formed first (Flow-1), and (b) the N-Cap(La_2O_3) is formed first (Flow-2). The process steps for Flow-1 are described as follows; first, the P-Cap Al_2O_3 layer is deposited on SiON/HfO_2 using an ALD technique with 7cycles (1.6E15 Al-atoms/cm^2). Then, PVD TiN and SiN are deposited (Flow-1(1)). After patterning the SiN (Flow-1(2)) by lithography and dry etching, the TiN and Al_2O_3 layer on the P-Well (PW) region are removed by wet etching. The remaining TiN is used as a hard mask because the etch rate selectivity between TiN and the high-k (HfO2) dielectric in the subsequent wet etch process is sufficiently high(Flow-1(3)). The N-Cap La_2O_3 layer is deposited using the ALD technique with 2cycles (5E14 La-atoms/cm^2) and PCA is carried out by RTA to diffuse La into the high-k dielectric on PW and Al into that on NW (Flow-1(4)). The temperature of the PCA is in the range of 850℃ to 1050℃. The TiN and La_2O_3 are removed by wet etching (Flow-1(5)). After this, PVD TiN/CVD Poly-Si are deposited (Flow-1(6)). In the case of Flow-2, the order of deposition of P-Cap and N-Cap is reversed. Fig.2 shows SEM and TEM views after removal of the TiN hard mask on the PW region (Flow-1(3)). There is no recess at the NW-PW boundary, and the physical thickness of the high-k dielectric on PW seems to be almost the same as that on NW. TEM cross-sectional views of n- and p-MOSFETs with 20nm gate lengths fabricated using Flow-1 are shown in Fig.3. ArF wet immersion lithography is used to pattern the gate.

3. Results and Discussion

Fig.4 shows the dependence of (a) ΔVfb and (b) ΔEOT on PCA temperature of n and pMOS capacitors with Flow-1 and Flow-2. It is obvious that higher temperature PCA lowers Vfb for nMOS and increases it for pMOS (Fig.4 (a)), and increases EOT for both n and pMOS (Fig.4 (b)). In the cases of Flow-2/nMOS and Flow-1/pMOS, in which the TiN hard mask is in contact with the high-k dielectric, the EOT increases significantly for 1050℃ PCA. Another finding is that |ΔVfb| of Flow-1/nMOS and Flow-2/pMOS tend to be larger than that of Flow-2/nMOS and Flow-1/pMOS, respectively. To investigate more detail, we measured the amount of La atoms in HfO_2 after removal of the TiN hard mask (@Fig.1 (5)) using X-ray Fluorescence (XRF) analysis (Fig.5). It is clearly shown that the amount of La atoms in HfO_2 is smaller in Flow-2/nMOS than in Flow-1/nMOS. Hence, Vfb in nMOS should be determined by the amount of La atoms in HfO_2, |ΔVfb| of Flow-2/nMOS became to be small due to decrease of La atoms. This fact must be considered in a CMOS process using a TiN hard mask. Fig.6 shows the dependence of Vfb and EOT on the thickness of the TiN hard mask for the Flow-2/nMOS capacitors in the case of PCA temperature is 1050℃. Vfb is lower and EOT is smaller for thicker TiN. These phenomena for nMOS are possibly explained by the following mechanism schematically shown in Fig.7. For Vfb shift, some of the La atoms in the layer deposited on the high-k dielectric might diffuse into the TiN. Then, |ΔVfb| due to La diffusion into high-k dielectric can be reduced originated in the TiN hard mask contacted with high-k dielectric. Another possibility is that the Al on the TiN hard mask might diffuse to the high-k dielectric through the TiN hard mask during PCA. For the EOT increase, the interfacial layer thickness may increase during PCA due to oxygen or nitrogen diffusion from the TiN. Fig.8 shows Vth (@Lg=1um)-EOT characteristics of n- and p-MOSFETs as a function of PCA temperature and thickness of the TiN hard mask for Flow-1 and Flow-2. The thicker TiN hard mask has the benefit of suppressing the increase in EOT and in lowering |Vth| when the PCA temperature is high and TiN is in contact with the high-k dielectric. To achieve low Vth and, at the same time, avoid an unacceptable increase in EOT, the PCA temperature and TiN hard mask thickness must be optimized. From Fig.8, we selected Flow-1 as the main process flow in this work. Fig.9 shows ring oscillator tpd vs. gate length for Flow-1 as a function of PCA temperature, with the case of a single high-k dielectric as a reference. The tpd is shorter when PCA is at higher temperatures due to lower Vth.

4. Conclusion

We have demonstrated a SMDH gate-first process using a TiN hard mask only once. A higher PCA temperature is effective in lowering |Vth|, and a thicker TiN hard mask has the benefit of suppressing the increase in EOT and in lowering |Vth|. The tpd measured using a ring oscillator was found to decrease when the DH process and higher PCA temperature were used.

5. References

[1] M.Chudzik et al., Symp. VLSI Tech. Dig., p.194, 2007.
[2] X.Chen et al., Symp. VLSI Tech. Dig., p.88, 2008.
[3] F.Arnaud et al., IEDM Tech. Dig., p.633, 2008.
[4] N.Mise et al., IEDM Tech. Dig., p.527, 2007.
[5] T.Schram et al., Symp. VLSI Tech. Dig., p.44, 2008.

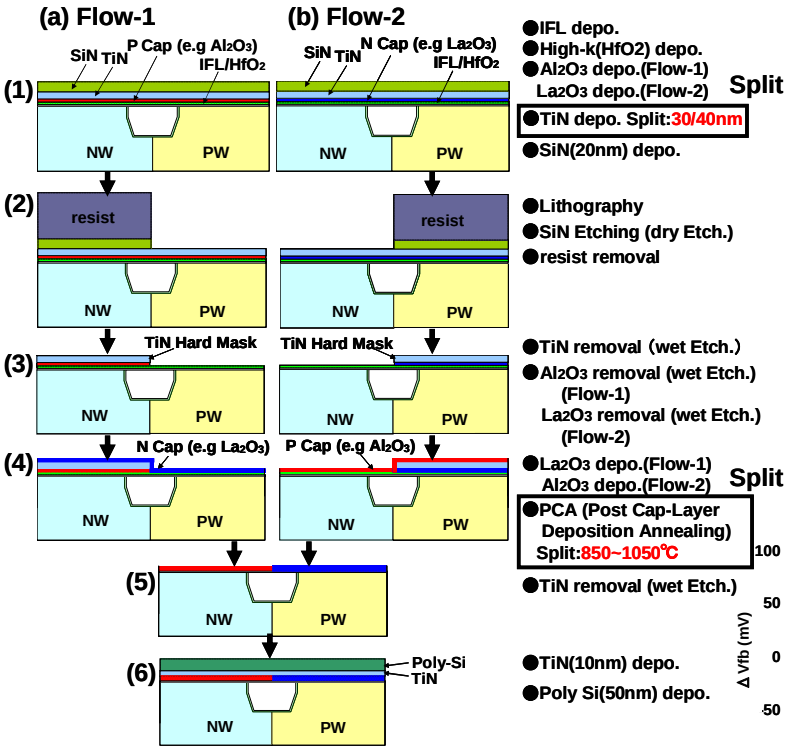

Fig.1 Process Steps for Single Metal Dual High-k (SMDH) Metal Inserted Poly Silicon(MIPS) CMOS flows in which (a) the P-Cap is formed first (Flow-1), and (b) the N-Cap is formed first(Flow-2) .

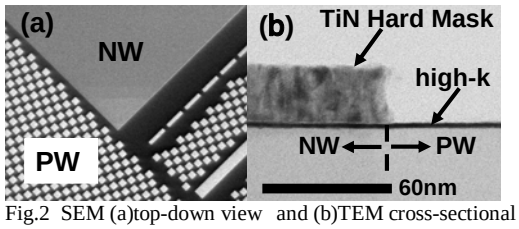

Fig.2 SEM (a)top-down view and (b)TEM cross-sectional view of ring oscillator after removal of the TiN hard mask for the PW region with Flow-1.

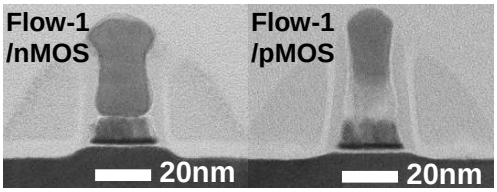

Fig.3 TEM cross-sectional views of n- and p- MOS FETs with Flow-1.

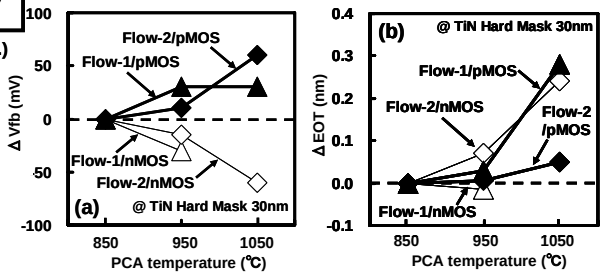

Fig.4 Dependence of (a) ΔVfb and (b) ΔEOT on PCA temperature for n and pMOS capacitors with Flow-1 and Flow-2. ΔVfb is defined as ΔVfb =Vfb-Vfb@850℃, and ΔEOT as ΔEOT =EOT-EOT@850℃.

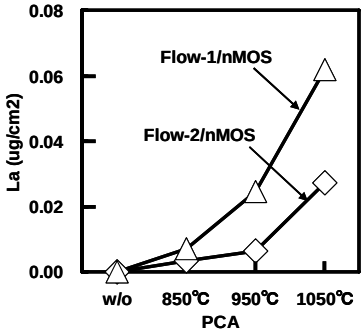

Fig.5 Dependence of the amount of La atoms in HfO_2 on PCA temperature for the nMOS capacitors with Flow-1 and Flow-2. The atomic weight of La is evaluated by XRF analysis.

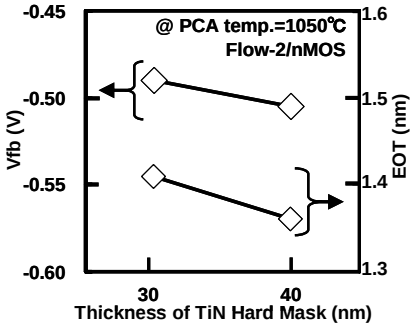

Fig.6 Dependence of Vfb and EOT on the thickness of the TiN hard mask for the Flow-2/nMOS capacitors.

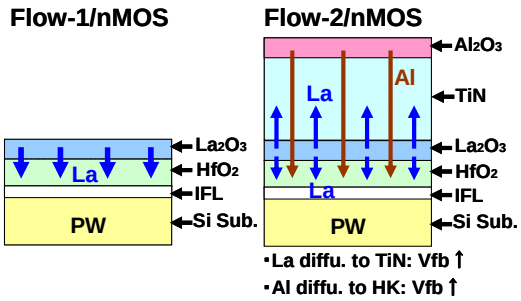

Fig.7 Schematics of proposed model for La and Al diffusion in Flow-1/nMOS and Flow-2/nMOS.

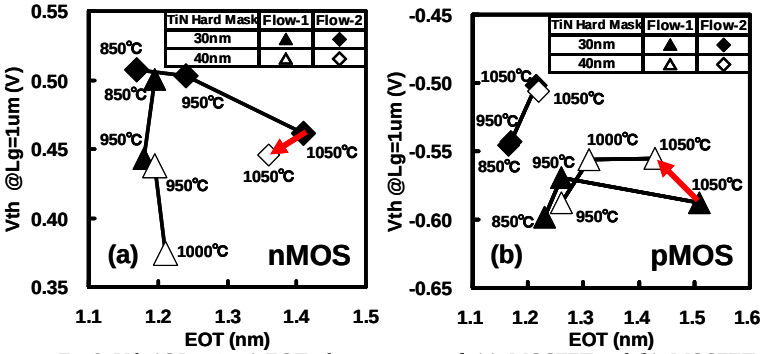

Fig.8 Vth (@Lg=1um)-EOT characteristics of (a)nMOSFET and (b)pMOSFET as a function of PCA temperature and thickness of the TiN hard mask for Flow-1 and Flow-2.

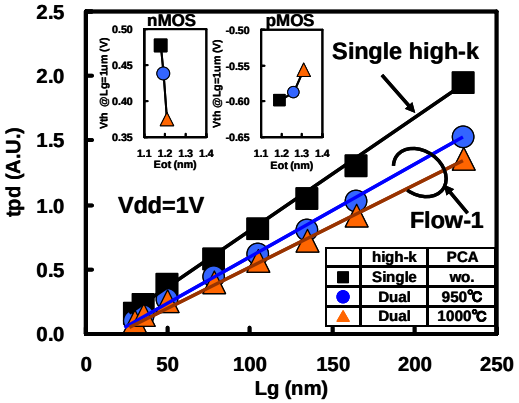

Fig.9 Ring oscillator delay time vs. gate length for Dual high-k dielectrics (Flow-1) as a function of PCA temperature and also for a process with a single high-k dielectric.