Low Temperature Atomic Hydrogen Treatment for Superior NBTI Reliability—Demonstration and Modeling across SiO_2 IL Thicknesses from 1.8 to 0.6 nm for I/O and Core Logic

J. Franco*, J.-F. de Marneffe, A. Vandooren, H. Arimura, L.-Å Ragnarsson, D. Claes, E. Dentoni Litta, N. Horiguchi, K. Croes, D. Linten, T. Grasser[1], B. Kaczer

imec, Leuven – Belgium, [1]T.U. Wien – Austria, *Jacopo.Franco@imec.be

NBTI remains a primary reliability concern for CMOS technology. Contrary to PBTI, which has been continuously reduced in the last five HKMG technology nodes, NBTI has remained virtually unchanged [1], and is often considered an ineradicable issue. In RMG integration flows, a high-temperature (T~900°C) post-metal anneal is customary to suppress dielectric defectivity. This high thermal budget step is incompatible with novel stacked integration schemes, such as Sequential 3D and CFETs [2]. Furthermore, the so-called 'reliability anneal' typically requires a thick sacrificial TiN/a-Si gate, which may become unsuitable for nanosheets with tight vertical spacing. In [3] hydrogen radicals (H*) generated in a low-T remote plasma were used to passivate hole traps associated with the hydroxyl-*E'* (H-*E'*) SiO_2 defects in 1.2nm thick interface layers (IL). These defects form at stretched Si-O bonds and are abundant in IL's grown at reduced T due to unrelaxed interface strain, causing poor NBTI reliability [3]. In this work, we i) *explore the H* treatment process window*, ii) *optimize the treatment for ultra-thin 0.6nm chemical oxide* IL's (chemOx), focusing on EOT control, and iii) *show the applicability* for 1.8nm thick IL's, of relevance *for I/O devices*.

MOS capacitors were fabricated in an RMG-compatible flow on 300mm wafers. The low-T gate stack fabrication flow is depicted in **Fig. 1**: the IL is exposed to H* right after formation, before HKMG deposition. To further establish the key role of hydrogen for oxide defect passivation, we compare with treatment in He plasma—as shown in **Fig. 2**, only a marginal reduction in NBTI shifts was observed in this case. Various H* exposure T's (100/200/300°C) and times (10"-0.5h) were tested on a 1.2nm IL grown at 600°C (**Fig. 3**), demonstrating improved NBTI using all tested conditions with up to ~100× trap density reduction and a beneficial enhancement of the NBTI field-acceleration exponent γ up to ~7 [4], outperforming a Foundry28 ref. [5] despite the limited thermal budget. The reduction of oxide trap density during H* exposure is well approximated by a first-order H-*E'* passivating reaction model (T+H*→TH, where T denotes the Si dangling bond at the H-*E'* site), with mean activation energy E_a=0.2eV (**Fig. 4**). Since a fraction of the SiO_2 hole trap defect band reaches inside the Si bandgap and therefore behaves as 'fixed' positive charge, a beneficial increase of the pMOS effective work function (eWF) is consistently observed after H* exposure, tightly correlated with the NBTI improvement (**Fig. 5**). Note that the increase in the max operating V_{ov} is much larger than the eWF shift; i.e., the treatment enables larger operating V_g's.

EOT scalability and control are other critical issues for high performance Logic. Different H* exposures on the 1.2 nm IL resulted in EOT changes by as much as ±2Å, as shown in **Fig. 6**. We correlate this non-monotonic trend to the degree of Si surface passivation right after the H* exposure. It is well known that hydrogen interacts with Si surface dangling bonds (so-called P_b-centers). We envision three reactions taking place at P_b sites during the H* exposure: i) H* might passivate the 50% of unsaturated P_b's [6] naturally left after Si oxidation (P_b+H*→ P_bH); ii) at longer exposure times, the reverse reaction (P_bH+H*→ P_b+H_2) [7] will take over, and iii) for even longer times a sufficient concentration of H_2 formed by H* dimerization will induce the P_b passivating reaction exploited in typical (Forming Gas) sintering anneals (P_b+H_2→ P_bH+H*, with E_a~1.66eV [6]). Note that the latter is expected to be the slowest reaction at the low-T range of the H* treatment. Due to the interplay of these reactions, different densities of unsaturated P_b's might exist after different H* treatments, and a larger density will enhance suboxide formation when oxygen becomes available during HKMG deposition, resulting in an EOT increase. (Note: the dedicated P_b passivation anneal in H_2 to ensure low electrically active D_{it} is performed only at the very end of the flow, as usual.)

Understanding of the mechanism discussed above becomes crucial when optimizing the H* treatment for an ultra-thin 0.6nm chemOx. As shown in **Fig. 7**, a 300°C-10" H* exposure resulted in excellent NBTI reliability, but with an unacceptable EOT increase of ~7.5Å. Interestingly, such large EOT increase did not yield the corresponding gate leakage suppression expected for an increased SiO_2 thickness (**Fig. 7f**), supporting our hypothesis of suboxide formation. The EOT increase was reduced to ~2Å for a 100°C-100" H* exposure, but this also resulted in a reduced NBTI improvement. A similar EOT control was obtained by performing an additional H_2 sintering anneal (400°C-20') right after the 300°C-10" H* exposure, to ensure optimal P_b passivation (cf. dotted line in **Fig. 6a**) before HKMG deposition. This enables the excellent NBTI benefit of a 300°C-10" H* exposure on the chemOx-based gate stack with a final EOT of ~11Å, *clearly outmatching a Foundry28 ref., at ~3Å thinner EOT* (**Fig. 7e**) *in a low thermal budget flow*. The excellent NBTI reliability is confirmed at elevated stress temperature typically considered for process qualification (**Fig. 8**).

The H* treatment is demonstrated also on a 1.8nm thick IL (**Fig. 9**). A short 10" exposure at 300°C resulted in a marginal NBTI improvement possibly due to i) larger number of defects to be passivated (volume effect), and ii) a reduced number of H* reaching through to the defective near-interface SiO_2 (note: H* transport appears not to be diffusion-limited, as otherwise 10" would be enough to reach through the 1.8nm SiO_2 [8]; H* might instead dimerize or react at other oxide sites along the way). Increasing the exposure time to 10' resulted in the same excellent NBTI reliability previously achieved on the 0.6nm IL's with a 10" exposure. Interestingly, an even longer exposure of 0.5h reduced the NBTI improvement, suggesting that the reverse reaction (TH+H*→T+H_2) eventually takes over for very long exposures (cf. Fig. 6a for the P_b's).

Fig. 10 depicts a benchmark of the NBTI max operating V_{ov} vs. EOT for all the studied stacks: the H* treatment dramatically enhances the tolerable oxide electric field, outperforming i) a Foundry28 ref., ii) a chemOx HKMG stack subject to a conventional high-T reliability anneal, and iii) our previous pMOS low-T gate stack solution based on Al_2O_3 interface dipole [9]. The *excellent NBTI rating of V_{ov} >1V at ~11Å EOT* (E_{ox}~7MV/cm) demonstrated here on Si at low thermal budget *compares well with the best-in-class reliability previously demonstrated in Gate-First Si-capped $Si_{0.45}Ge_{0.55}$ pMOS* [10].

Finally, a physics-based NBTI model including fast and slow near-interface traps is calibrated in the imec/T.U. Wien modeling framework Comphy [5] against the measured NBTI kinetics (**Fig 11**). The model assumes a graded density of traps across the oxide depth, decaying away from the interface in correlation with strain (**Fig 12**). With this feature the model reproduces the NBTI kinetics of the stacks with different IL thicknesses and treatments by adjusting only the defect densities, enabling a direct comparison of all stacks (**Table I**) which further highlights the dramatic reliability improvements achieved with the low-T H* exposures.

References: [1] K. Choi, IEDM 2020; [2] C.-Y. Huang, IEDM 2020; [3] J. Franco, IEDM 2020; [4] J. Franco, IEDM 2013; [5] G. Rzepa, Micr. Rel. 85, 2018; www.comphy.eu; [6] A. Stesmans, Phys. Rev. B 48(4), 1993 [7] E. Cartier, APL 63(11), 1993; [8] J. Stathis, Micr. Rel. 81, 2018; [9] J. Franco, IEDM 2018 [10] J. Franco, TED 60(1), 2013.

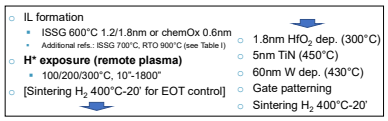

Fig.1: Schematic of the fabrication flow.

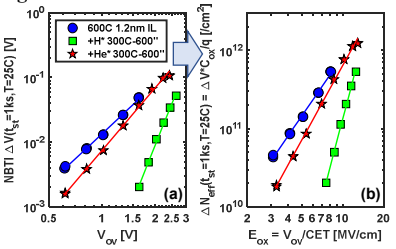

Fig.2: NBTI-induced (1ks stress at 25°C) **(a)** $\Delta V(V_{ov})$, converted into **(b)** $\Delta N_{eff}(E_{ox})$, as measured on MOS capacitors (1.2nm 600°C IL) w/o and w/ exposure to He* or H* (300°C-600") through a remote plasma.

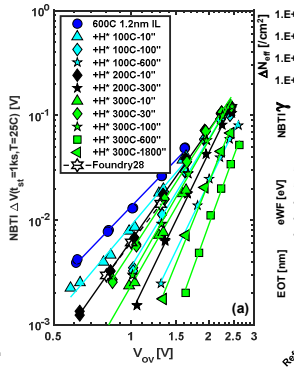

Fig.3: Impact of different H* exposures on a 1.2nm 600°C IL: **(a)** NBTI-induced voltage shift vs. stress V_{ov}, **(b)** equivalent trapped charge density at E_{ox}=4.3MV/cm, **(c)** NBTI field-acceleration exponent γ, **(d)** effective Work Function, **(e)** EOT.

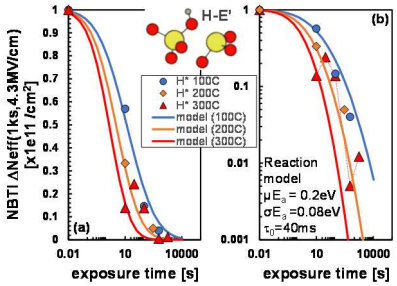

Fig.4: Measured NBTI-induced ΔN_{eff} with different H* exposures (100-300°C, 10-1800"; ref. w/o H* is plotted at 0.01s). The ΔN_{eff} reduction is approximated well by a first-order passivating reaction with mean E_a=0.2eV. **(a)** lin-log, **(b)** log-log to accentuate low values. The deviation from the simple model is possibly due to other concurrent reactions (cf. Fig. 6).

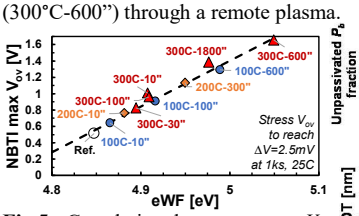

Fig.5: Correlation between max V_{ov} improvement and eWF increase with different H* exposures on a 1.2nm IL.

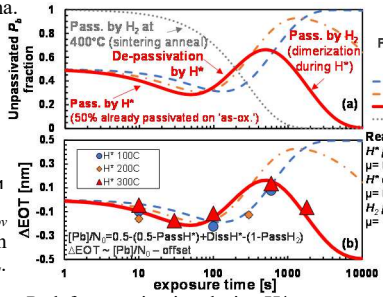

Fig.6: (a) Qualitative model of interface P_b defect passivation during H* exposure, considering three reactions (assumed to be independent for simplicity): [1]passivation/[2]de-passivation by H*, and [3]passivation by H_2 formed by H* dimerization or supplied by subsequent sintering anneal (400°C-20'). **(b)** The measured EOT evolution correlates well with the P_b kinetics.

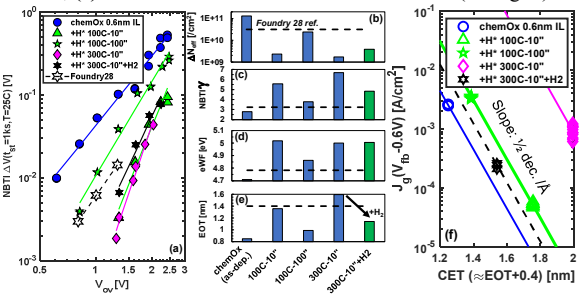

Fig.7: (a-e) Same as Fig. 3, now for a chemOx 0.6nm IL. A 300°C-10" H* exposure followed by a 400°C-20' H_2 sintering yields optimal NBTI reliability with just ~2Å EOT increase. **(f)** Gate leakage density at V_{fb}-0.6V plotted vs. CET.

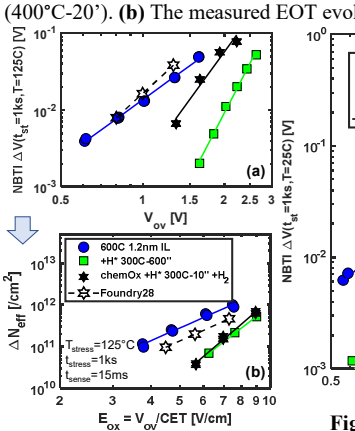

Fig.8: (a) $\Delta V(V_{ov})$ converted into **(b)** $\Delta N_{eff}(E_{ox})$, measured after 1ks stress at 125°C on the 1.2 and 0.6 nm IL samples, confirming the excellent reliability after H* exposure.

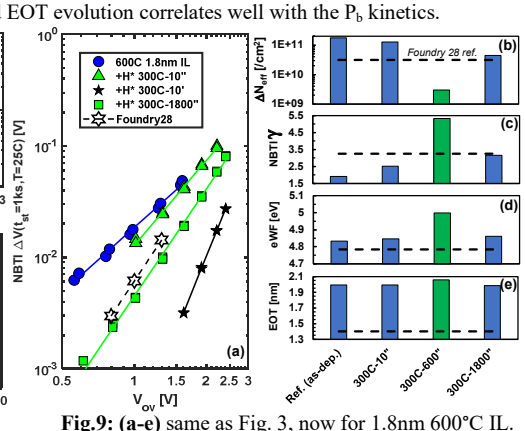

Fig.9: (a-e) same as Fig. 3, now for 1.8nm 600°C IL.

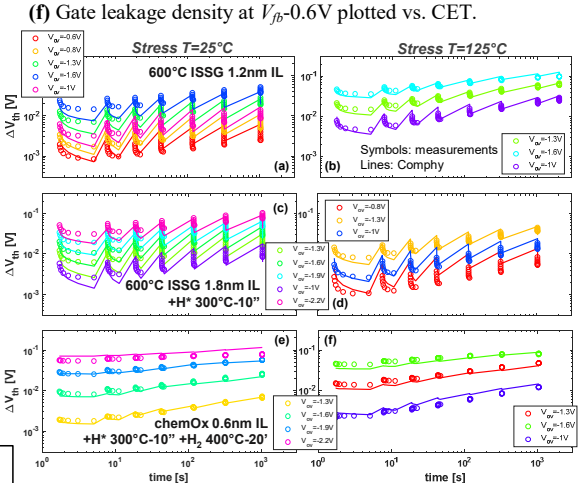

Fig.11: NBTI stress-recovery traces measured at different stress V_{ov} and T (25/125°C) are well reproduced by Comphy simulations (lines). The model calibrated on **(a-b)** a 1.2nm 600°C IL ref. reproduces well the kinetics on all other samples (different IL thicknesses and treatments, cf. Table I) by adjusting only the defect densities. Shown as examples: **(c-d)** 1.8nm IL w/ H* 300°C-10", and **(e-f)** 0.6nm chemOx w/ H* 300°C-10"+400°C-20' H_2.

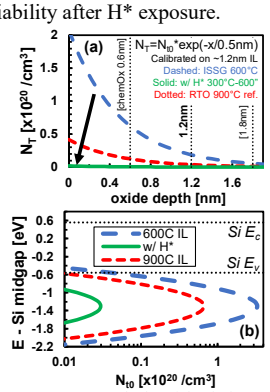

Fig.12: (a) A graded oxide defect profile, representing strain relaxing away from the interface, allows to reproduce the NBTI kinetics of all IL's by adjusting only the peak density. **(b)** Hole trap band within the SiO_2 bandgap: the low-T H* exposure outperforms a 900°C ref. IL.

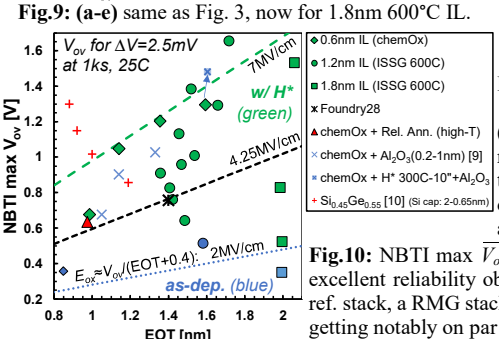

Fig.10: NBTI max V_{ov} vs. EOT for all the IL thicknesses and treatments studied. The excellent reliability obtained w/ the low-T H* treatment outperforms a Foundry 28nm ref. stack, a RMG stack with high-T rel. anneal, and our previous dipole-based stack [9], getting notably on par with best-in-class Si-capped $Si_{0.45}Ge_{0.55}$ Gate-First devices [10].

(a) Fast traps (NMP model [5] calibrated on eMSM)		(b) Slow traps (Double Well model [5] calibrated on *CpG-V* [3])	
N_{t0} [/cm³]	see stack list (c)	Precurs. [/cm²]	see stack list (c)
x_0 [nm]	0.5	$\varepsilon_1\mu$ [eV]	2.14
$E_T\mu$ [eV]	-1.30	$\varepsilon_1\sigma$ [eV]	0.41
$E_T\sigma$ [eV]	0.25	$\varepsilon_2\mu$ [eV]	2.20
$S\mu$ [eV]	1.69	$\varepsilon_2\sigma$ [eV]	0.46
$S\sigma$ [eV]	0.20	γ [eV mV]	8.69e-10
R (removed)	1	k_0 [/s]	1e13

(c) IL	IL treatment	sims. t_{IL}[nm]	N_{t0} [/cm³]	DW prec. [/cm²]	N_T integrated across IL + DW prec. [/cm²]	Total normalized to 600°C 1.2nm ref.
ISSG 600°C 1.2nm	none	1.35	2.1E+20	3.8E+13	1.8E+14	1.00
ISSG 600°C 1.2nm	H* 300°C-10"	1.35	3.1E+19	3.1E+13	5.2E+13	0.29
ISSG 600°C 1.2nm	H* 300°C-600"	1.35	1.9E+18	7.7E+12	9.0E+12	0.05
ISSG 600°C 1.8nm	none	1.8	2.1E+20	3.8E+13	4.1E+14	2.24
ISSG 600°C 1.8nm	H* 300°C-10"	1.8	1.9E+20	3.2E+13	3.6E+14	2.00
ISSG 600°C 1.8nm	H* 300°C-600"	1.8	3.7E+18	1.6E+13	2.3E+13	0.13
ISSG 700°C 1.2nm	none	1.15	1.2E+20	3.0E+13	8.4E+13	0.46
ISSG 700°C 1.8nm	none	1.8	1.2E+20	3.0E+13	2.4E+14	1.34
RTO 900°C 1.2nm	none	1.15	4.2E+19	2.7E+13	4.5E+13	0.25
chemOx 0.6nm	none	0.6	9.4E+20	4.5E+14	5.6E+14	3.08
chemOx 0.6nm	Rel. ann. (850°C-1.5" after sacr. gate)	0.6	2.5E+20	4.4E+13	7.3E+13	0.40
chemOx 0.6nm	H* 300°C-10" + H2 400°C-1200"	0.8	1.4E+19	2.3E+13	2.6E+13	0.14

Table I: (a-b) Comphy model parameters [5] for fast and slow NBTI traps as calibrated on the 1.2nm 600°C ref. IL. **(c)** The model reproduces well the NBTI kinetics of all samples studied by adjusting only the defect densities (columns w/ yellow background). Note the dramatic reduction of the total defect density achieved by low-T H* exposure.