Stacked-Nanowire and FinFET Transistors: Guidelines for the 7nm node

Loïc Gaben1,2,3, Sylvain Barraud2, Marie-Anne Jaud2, Sébastien Martinie2, Olivier Rozeau2, Joris Lacord2, Gaspard Hiblot1, Stéphane Monfray1, Frédéric Beuuf4, Thomas Skotnicki1, François Balestra1, Maud Vinet2

1STMicroelectronics, 850 rue J. Monnet, 38920 Crolles, France, E-mail: loic.gaben@st.com
2CEA-LETI, Minatec campus, 17 rue des Martyrs, 38054 Grenoble Cedex 9, France
3IMEP-LAHC, 3 parvis Louis Néel, BP257, 38016 Grenoble Cedex 1, France

Abstract
This study, based on 3D TCAD simulation, suggests innovative guidelines for benchmarking performances of stacked-nanowires and FinFET architectures. Immunity to short-channel effects (SCE), parasitic capacitances, and switching delays are evaluated. Thin and wide gate-all-around (GAA) stacked-nanowires are found to be the most promising devices for the 7nm node.

1. Introduction
Stacked-nanowire (NW) transistors are considered as the main alternative to FinFET (FF) technology for the 7nm node [1]. In the case of Gate-Last and NW-Last integration, Si fins are replaced by (Si/SiGe)IB fins as illustrated in Fig. 1. The Si channels can be obtained after selective etching of SiGe layers. However, it still remains unclear if, using the same technology constraints as for FF fabrication, stacked-NW devices are able to overcome FF performances in advanced CMOS technologies. In this work, guidelines for accurate benchmarking of stacked-NW and FinFET architectures are suggested. Using TCAD simulation, NW dimensions (width and height) required to overcome FinFET performances for given footprint and height are proposed.

2. Benchmark Guideline for the 7nm node
Significant technological advances have been made in recent years in advanced patterning of FinFET leading to 40nm fin pitch (FP) [2]. As reported by the TEM image of Fig. 1, it is reasonable to expect NW fabrication to benefit from the same technology. This enables stacked architectures with a total height $H_{NW}=H_{FIN}$ and with a space between the stacks $W_{S}=FP$-$W_{FIN}$ as summarized in Fig. 1. $H_{NW}$ includes several stacked-NWs (thickness $T_{NW}$) and the distance between each stacked-NW ($T_{S}=8nm$) in order to have $H_{NW}$=$H_{FIN}$. As device footprint is also decisive in the selection of a technology, we have chosen to compare both technologies within the same footprint. Fig. 2 presents several GAA stacked-NW configurations involving different $W_{NW}$ authorized for a given footprint. TCAD Sentaurus simulation tool [3] was used to perform the 3D simulations which account for quantum confinement effects with a calibrated density-gradient model [4]. Operating voltage was $V_{DD}=0.7V$ and the gate length has been fixed at $L_{G}=16nm$. A basic transport model was used with a constant mobility of 100cm²/V.s and a saturation velocity of $1.07 \times 10^7$cm/s. NMOS delay is defined by $\tau_{CUT} \times V_{DD}/I_{D}$ with $I_{D}$ the effective current extracted at $V_{DD}=0.7V$ from the off-state current $I_{OFF}=100nA/\mu m$ in order to suppress threshold voltage effects. Equivalent capacitance $C_{eq}$ is obtained from [5]:

$$C_{eq} = (M+2 \times FO) \times C_{gdo} \times \frac{3}{4} \times \frac{S_{FIN}L_{P}W_{EFF}}{t_{inv}} \times FO + \frac{C_{BE}}{2}$$

with $M=2$ the coefficient for Miller effect, $FO=3$ the electric fan-out, $W_{EFF}$ the effective width (perimeter), $t_{inv}$ the inversion layer thickness, $C_{BE}$ the back-end capacitance, and $C_{gdo}$ the gate-to-drain parasitic capacitance extracted at $V_{G}=0V$.

3. Results and Discussion
First of all, let us assume a FinFET technology characterized by $W_{FIN}=7nm$ and $FP=30nm$. A 157nm footprint (corresponding to 6 Si fins) is shown here as an example. Immunity to SCE is investigated for FinFET and two families of stacked-NW devices (cf. schematics in Fig. 3): (i) a “FULL-GAA” structure with gate-all-around (GAA) NWs and (ii) a “MIXT-GAA” structure composed of a trigate (bottom wire) and two GAA NWs. In both cases, the use of wide NWs leads to increase $W_{eff}$ as compared to FF which is a key element for improving performances. Similarly to FF, the “FULL-GAA” configuration shows a DIBL~60mV/V but with higher $W_{eff}$. However, the “MIXT-GAA” configuration is less interesting due to the trigate structure of bottom NW [6]. $C_{gdo}$ capacitance includes most of the parasitic capacitances involved in the inverter switching frequency. Fig 4 suggests that an increase of $W_{eff}$ through $H_{FIN}$, $W_{NW}$ and $H_{NW}$ might not penalize the switching delay since normalized $C_{gdo}$ grows slower than $W_{eff}$. If performance improvement of stacked-NW for a given height can be obtained by increasing $W_{NW}$, several fins are usually added in parallel for FF technology. This is why our benchmark has been done within a constant footprint (Fig. 1 and 2). Fig 5 shows the delay reduction obtained by using a NW technology (FULL-GAA) instead of FF as a function of the footprint. The highest improvements are obtained for the largest footprints and $W_{NW}$. It can be noticed that the NW delay improvement is mostly dependent on $W_{NW}$. This effect is highlighted in Fig. 6. The gain also appears barely affected by the number of vertically stacked GAA ($H_{NW}$). As a consequence, this plot is then used as a figure of merit in Fig. 7 to evaluate the influence of $W_{S}$ and $T_{NW}$. If FF technology FP is reduced to 25nm ($W_{S}=18nm$) the delay reduction over FF technology is less significant due to a lowered ratio $W_{ff}/W_{FF,NW}$. The same reduction is observed if $T_{NW}$ is increased due to higher SCE.

4. Conclusions
Stacked-NW and FF devices geometries are evaluated for the 7nm node. Our benchmark methodology compares FF and GAA stacked-NWs for identical footprint and height. Wide and thin GAA NWs (i.e. nanosheets) are found more competitive than FinFET thanks to their increased $W_{eff}$ and limited parasitic capacitances within given footprint and height.
Si Fins are replaced by (Si/SiGe/Si/SiGe/Si) Fins. SiGe is then selectively etched leaving suspended Si-channels. The TEM cross section image of such (Si/SiGe), fin is presented as a proof of concept. Main parameters selected for 7nm-node design are summarized in the tables.

**Fig. 1:** Process adapted from FinFET fabrication for stacked-NW patterning. Si Fins are replaced by (Si/SiGe/Si/SiGe/Si) Fins. SiGe is then selectively etched leaving suspended Si-channels. The TEM cross section image of such (Si/SiGe), fin is presented as a proof of concept. Main parameters selected for 7nm-node design are summarized in the tables.

**Fig. 2:** Guideline for benchmarking architectures. Several stacked-NW configurations match the envelope defined by the FinFET configuration. In this example labeled “FULL-GAA” with 2 GAA vertically stacked (Htot = 2 * Ttot + 2 * T), several WNW meet the condition Footprint=157nm for Htot=35nm.

**Fig. 3:** DIBL vs WFIN in the case of footprint=157nm, several WNW are allowed increasing WFIN over the reference FinFET configuration. The “FULL-GAA” configuration offers a better electrostatic control (DIBL=80mV/V at LFIN=16nm, TFIN=7nm, HFIN=Htot=30nm) than “MIXT-GAA” configuration due to the bottom trigate (LFIN=16nm TFIN=6.3nm, HFIN=Htot=35nm).

**Fig. 4:** Normalized Cpar as a function of HFIN and WFIN within a 157nm footprint. Parasitic capacitances are less increased than WFIN when HFIN and WFIN are increased for (a) FinFET and (b) stacked-NW architectures, respectively. WFIN can also be increased for stacked-NW in the vertical direction by stacking more channels: once again WFIN increases more than Cpar as HFIN is increased.

**Fig. 5:** Delay reduction over FinFET technology for different footprint. Each stacked-NW configuration is compared to FinFET having the same HFIN and footprint. Here the comparison is based on FinFET devices with WFIN=7nm, HFIN=30nm and FP=30nm (W=23nm). Only the configuration “FULL GAA” is shown in this case with 2 vertically stacked GAA NWs. NW thickness is TFIN=7nm for a total height HFIN=Htot=30nm. 40% delay reduction is observed for the widest NWs.

Configurations involving NW having the same WFIN present similar values: the 3 configurations with WFIN=37nm show a delay reduction of -26%, -25% and -24% for 37nm, 97nm and 157nm footprints, respectively.

**Fig. 6:** Delay reduction over FinFET technology for different WFIN independently of the number of stacked GAA channels (TFIN=7nm, HFIN=Htot=30nm) or footprint: delay reduction over FF depends mainly on WFIN.

**Fig. 7:** Figure of merit: delay reduction over FinFET technology. Delay improvement is lowered as WFIN is shrunk or if channel thickness TFIN is relaxed from 6nm to 8nm.

**References**