Advances, Challenges and Opportunities in 3D CMOS Sequential Integration

P. Batade1, M. Viner1, B. Previtali1, C. Tabone1, C. Xu1, J. Mazarier1, O. Weber1, F. Andreiu1, L. Tosti1, L. Brevard1, B. Sklenard2, P. Coudrain3, S. Bobba4, H. Ben Jamaa1, P.-E. Gaillard2, A. Pousydebasque2, O. Thomas1, C. Le Royer1, J.-M. Hartmann1, L. Sanchez2, L. Baud1, V. Carron1, L. Clavelier1, G. De Michelis1, S. Deleonibus1, O. Faynot1, O. Thomas1, J.-M. Hartmann1, G. De Micheli1, O. Weber3

1CEA-leti, Minatec Campus, F-38054 Grenoble, France / 2ST Microelectronics, F-38926 Crolles, France / 3EPFL, Lausanne, Switzerland.

Abstract- 3D sequential integration enables the full use of the third dimension thanks to its high alignment performance. In this paper, we address the major challenges of 3D sequential integration: in particular, the control of molecular bonding allows us to obtain pristine quality top active layer. With the help of Solid Phase Epitaxy, we can match the performance of top FET, processed at low temperature (600°C), with the bottom FET devices. Finally, the development of a stable salicide enables to retain bottom performance after top FET processing. Overcoming these major technological issues offers a wide range of applications.

Introduction- The 3D sequential integration scheme offers the possibility to fully use the third dimension potential, i.e., to connect two stacked layers at the transistor scale whereas 3D parallel integration is limited to connecting blocks of a few thousand transistors (Fig.1).

In this paper, the challenges of 3D sequential integration, (i.e. 1- stable performance bottom FET, 2- high quality top substrate fabrication, 3- top FET LT processing) will be presented as well as the proposed solutions to achieve such integration. Examples of potential applications are also reviewed.

Device fabrication- The process flow enabling to tackle the above mentioned challenges is presented in Fig.3. P- and N-FDSOI transistors with high-K/metal gate stack are fabricated on the bottom layer and standard high temperature spike anneal (1050°C) is used for dopant activation. Before bonding, thin Inter Layer Dielectric (ILD) is deposited and planarized on top of the patterned bottom transistors. LT (200°C) molecular bonding of SOI substrate enables full transfer of a monocrystalline Si layer. Top MOSFETs are then processed at low temperature (≤600°C). In particular, high temperature dopant activation is replaced by Solid Phase Epitaxy (SPE) at 600°C.

Bottom MOSFET performance- The first challenge in monolithic integration is to preserve bottom FETs performance during top FETs processing. To avoid additional dopant diffusion or interfacial oxide growth on bottom transistor, LT (<650°C) top FETs process is mandatory. Stabilized salicide is also required. Ni based salicide stability is obtained up to 650°C thanks to F & W implantation together with Pt incorporation (Fig.4-left). After complete 3D integration, R, of the stabilized salicide shows no degradation (Fig.4-right).

Fig. 1: Alignment capability versus 3D contact width in parallel and sequential integration schemes. Reported also in the graph: Bulk TSV size and alignment capability limits with correct reliability & throughput respectively/ contact size and alignment capability for planar integration in 65 & 22nm nodes.

However, its implementation faces the challenge of being able to process a high performance top transistor at Low Temperature (LT) in order to preserve the bottom FET from any degradation, as the stacked FET's are fabricated sequentially (Fig.2).

In this case, the stacked wafers are processed separately. In the sequential scheme, the transistors are processed sequentially above each other.

In this paper, the challenges of 3D sequential integration, (i.e. 1-
**Top active fabrication** - Fig. 5 presents the different techniques to obtain a crystalline semiconductor layer above processed transistors. Molecular bonding clearly stands apart from other techniques: first, it suppresses the need for Seed Windows (SW) required in recrystallization techniques [5-10] and thus allows higher integration density. In addition, bonding benefits from pristine crystalline quality and accurate thickness control.

![Image](image1.png)

**Fig. 5**: Description, benchmark and main references for top active realization techniques.

Active layer transfers with semiconductor and interlayer dielectric thicknesses down to 10 and 25nm respectively have been demonstrated (Fig. 6 (a) & 12). Perfect bonding at the wafer scale is evidenced with acoustic and infrared characterization in Fig. 6b&c.

![Image](image2.png)

**Fig. 6 (a)** SEM cross section of a thin Si layer stacked above a transistor layer (b) Infrared and (c) acoustic characterization of bonded top active layer on bottom layer showing full transfer and no bonding defects for 200nm wafer (d) Wafer bonding flow.

**Low temperature process (600°C) for top FET** is achievable thanks to the use of SPE and high-K gate oxide. Indeed, as shown in Fig. 7, SPE anneal at 600°C leads to similar ION/IOFF trade-off than standard 1050°C spike anneal for both n&pFETs. These FDSOI transistors have been processed at CEA-leti.

![Image](image3.png)

**Fig. 7:** Comparison of ION/IOFF trade-off for planar n- and p-FETs with low and high temperature dopant activation anneal.

The LT process does not increases GIDL leakage as shown in Fig. 8. This is explained by the role of buried oxide on the end of range defects dissolution [16] which is enhanced in thin FDSOI devices (Tox=6nm).

![Image](image4.png)

**Fig. 8:** Similar junction quality for HT and LT planar devices is demonstrated by the plotting of Imax cumulative distribution (Vth=0.9V).

Additionally, LT activation presents variability values in line with state of the art results for FDSOI devices (AVT=1.35mV.µm) [17] as shown in Fig. 9.

![Image](image5.png)

**Fig. 9:** Pelgrum plot of planar LT & HT devices. LT AVT value is in line with state of the art variability values on FDSOI [17]

Finally, LT process presents an improved gate leakage versus EOT trade-off than HT process (Fig. 10-left). The EOT reduction is explained by a 4 Å thinner interfacial oxide (Fig. 10-right).

![Image](image6.png)

**Fig. 10-Left:** Figure of merit of top (LT) and bottom (HT) oxides underlining the improved gate stack quality with LT process.

**Fig. 10-Right:** TEM cross section of low temperature and standard high temperature process showing a reduction of interfacial SiO₂ oxide.

**3D structures demonstration** - Fig. 11 benchmarks the technological options used in state-of-the art 3D sequential demonstrators. It highlights the interest of molecular bonding together with the 600°C process scheme, enabling the integration of bottom salicide.
Fig.12: TEM cross-sections of stacked transistors with record Lg=50nm and ultra thin interlayer dielectric TILD=23nm, TSi=10nm.

Performance benchmark of the top pFETs with the state of the art is presented in Fig.13. For the same I eff of 100nA/µm, the 600°C top pFET reaches comparable I on values (taking into account the smaller V th) than the top pFET of [3, 6] (processed at 650°C + spike anneal activation >1000°C). Note that using HT spike anneal for top dopant activation is detrimental for bottom FET performance (Ni salicide agglomeration and additional dopant diffusion).

Fig.13: Benchmark of top FET with 3D sequential integration literature with Lg<100nm.

Fig.14 presents the 3D inverter transfer voltage characteristics with such scaled gate length. Functional 3D 6T SRAMs cells have also been demonstrated, as shown in Fig.15.

Application and perspectives- Sequential integration offers 3D contacts pitch close to planar contact pitch (Fig.1). This enables circuit partitioning at a fine granularity (i.e. at transistor/gate scale), which yields new potential applications. For example, such high density 3D contacts can be helpful for FGAs, highly miniaturized imagers [19] and CMOS gates. Gain in performance is possible through the integration of the best suited technologies for different functions on distinct levels. Fig.16 summarizes the possibilities of co-integrations adapted to the different split functions.
Such heterogeneous co-integration in a planar scheme would lead to complex and thus costly process. It is worth noting that the partitioning of the different functions described in Fig.16 can only be achieved in sequential integration thanks to its low 3D contacts pitch (in opposition to TSV technology with its 10μm pitch) as highlighted in Fig.18.

As these proposed applications are built from matrix of full custom cells, their design can be achieved without using 3D place and route tools. By extending this concept to complex digital ASIC, it is envisaged to stack logic cells in a 3D arrangement. The issue of 3D place and routing is settled thanks to a new 2D to 3D transformation technique [20], which is based on smaller standard cell stacking on top of bigger cells (Fig.19). This enables the use of standard 2D place and route algorithm.

Using this tool, a 15% reduction in the average interconnect length and a x1.8 improvement in overall power-delay-area product are predicted for the 45nm node (Fig.20).

Finally, because of their regular and dense architectures, memories would largely benefit from 3D sequential integration. Indeed, it appears that, to continue to decrease bit cost, stacking will be more efficient than scaling [22].

**Conclusion**—Thanks to its ability to offer fine-grain circuit partitioning at the transistor scale, 3D sequential integration opens up a new field of applications and design. It enables both increasing the density and performances without resorting to aggressive scaling. Its key technological enablers are molecular bonding and low temperature top FET process which lead to design 3D transistors targeting the targets of advanced nodes thanks to low access resistance, salicide, scaled EOT, optimized threshold voltage and mobility boosters.

**Acknowledgments**—This work has been carried out in the frame of the ST/IBM/CEA-Leti development alliance.

**References**

1. P. Giroux et al., Handbook of 3D integration, vol 1 (Wiley Ed)
2. P. Giroux et al., Handbook of 3D integration, vol 2 (Wiley Ed)
3. S-M. Jung et al., VLSI 2005, pp220
4. P. Batude et al., ECS journal 2008, VO16, pp47
5. Y-H. Son et al., VLSI 2007, p80
6. S-M. Jung et al., VLSI 2007, pp82
9. T. Cheng et al., IEDM 2009, pp179
10. T. Cheng et al., IEDM 2010, pp936
11. T. Naito et al., VLSI 2010, pp219
12. L. Xue et al., Trans. on Electron Dev, VOL. 50, NO. 3, pp601 (2003);
14. P. Batude et al., VLSI 2009, pp166
15. P. Batude et al., IEDM 2009; pp345
17. O Weber et al., IEDM 2010, pp38
18. P. Batude et al., VLSI 2011, pp158
19. P. Coudrain et al., IEDM 2008, pp1
20. S. Bobba et al.; ASPDAC 2011, pp336
21. www.opencores.org
22. S-M. Jung et al., IEDM 2006, pp1