Materials Challenges for III-V/Si Co-integrated CMOS

D. K. Sadana, C-W. Cheng, B. Wacaser, W. Spratt, K.T. Shiu, and S.W. Bedell
T J Watson Research Center
IBM
Yorktown Heights, NY 10598
dksadana@us.ibm.com

Abstract

This review focuses on material challenges associated with III-V co-integration with Si for future CMOS. There is a huge volume of literature on this topic as implementation of III-V monolithic integration with Si has been the holy grail for last four decades; targeting a wide range of applications including RF devices, LEDs, lasers, photo-detectors and the like. The key drivers have been the cost reduction, scalability with Si wafer diameter, and accessibility to highly scaled integrated circuits next to III-V devices. With the current focus on CMOS the pace of progress on monolithic integration has accelerated by leaps and bounds partly because of its vast impact on CMOS scaling, and partly due to the aggressive CMOS roadmap requirements. The discussion below concentrates on $In_{0.53}Ga_{0.47}As$ channel which is the dominant III-V material being pursued for future technology. Despite the narrow focus, fundamental and engineering challenges posed by this material encompass a broad range of material topics including epitaxial growth, crystallographic defects and their dynamics during growth and subsequent processing, clever device architecture to alleviate adverse impact of defects on device leakage, and innovative engineering for material improvement.

Introduction

Relentless scaling of CMOS transistors has continued despite formidable challenges in lithography and processing. This trend is expected to continue to 5 nm and beyond CMOS (Figure 1). However, scaling to sub-20-nm lithographic dimensions presents unique challenges that will slow the pace of conventional Moore's scaling. The increase in the drive current in scaled devices for faster switching speeds at lower supply voltages has largely occurred at the expense of an exponentially growing leakage current, thus leading to a large standby power dissipation. This strategy has worked well until recently because our computation needs have been dominated primarily by high performance end products (mainframes, desktops, PCs). However, with the recent shift of paradigm in user choice towards low form-factor hand-held gadgets, performance requirements for computing have shifted towards low-power consumption. Current exploration of high carrier mobility channels is being fueled by ever reducing low-power requirements at high performance for future products.

Changing the device architecture (e.g. FinFET, Trigate, Nanowire) can improve the electrostatics of a transistor and achieve lower power high performance devices. However these improvements have not gone far enough and fundamental improvement in device performance are needed. The fundamental improvements arise from the improved carrier mobility achieved by changing the device materials. Novel strain engineering techniques, such as epitaxial growth of embedded –Ge:H. SiGe, and Si:C in source-drain

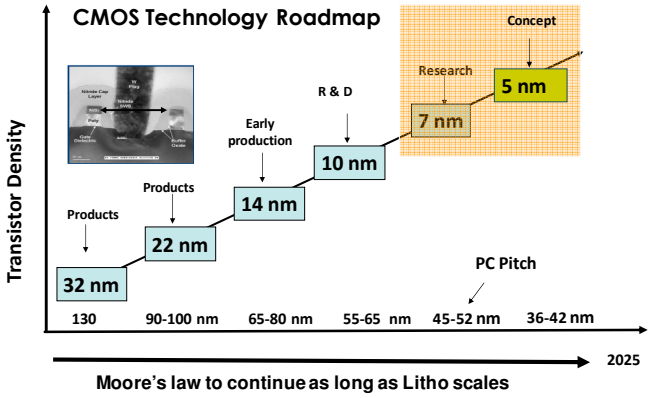


Figure. 1. CMOS technology roadmap extending to 5 nm and showing poly contact (PC) pitch requirements from 32 nm to 5 nm nodes.

regions, and application of compressive and tensile dielectric liners (in particular, Si_3N_4) to n and p-FET devices have already been implemented successfully in CMOS products since 90 nm node for performance

improvement. Further strain scaling would require higher Ge and C concentrations, but these higher concentrations have a negative effect on the strain in SiGe and Si-C due to increased relaxation in these layers. It is clear that further carrier mobility enhancements need to be achieved intrinsically by replacing the channel materials with high mobility materials in conjunction with previous strain enhancement. This is evident in IBM's 22 nm CMOS [1] where record performance p-FETs was demonstrated by introducing a SiGe channel with high hole mobility.

There are several materials to choose from. Figure 3 shows electron and hole mobilities in Si, Ge, and most promising III-V channel materials. Figure 2 shows well known lattice parameter – energy band gap relationship in Si, Ge and III-V compounds. To further scale pFET mobility either high % Ge or pure Ge is being considered as the channel material. For future high performance n-FETs both Si and non-Si channel materials including sSOI, biaxially strained Si on graded SiGe, and $In_{0.53}Ga_{0.47}As$ are being explored. The biggest boost in n-channel performance is expected from $In_{0.53}Ga_{0.47}As$ because of its x6 higher mobility than strain Si (Figure 3). Impressive device results have recently been reported on self-aligned $In_{0.53}Ga_{0.47}As$-channel MOSFETs grown on InP. Peak transconductance of > 2200 µS/µm with an effective channel length of 30 nm and supply voltage of 0.5 V (Figures 4 & 5) was achieved, thus validating scalability of III-V for future CMOS [2]. These

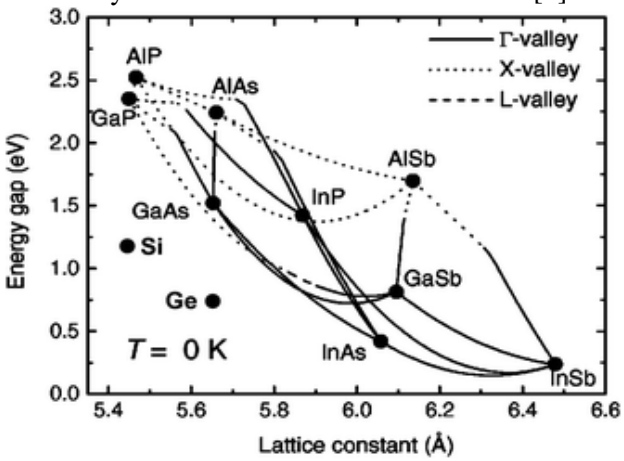


Figure. 2. Energy gap versus lattice parameter of Si, Ge, and III-Vs. Note that all high mobility II-Vs have a low band gap.

Material	μ_e (cm^2/Vs)	μ_h (cm^2/Vs)	E_g (eV)
Si	1350	480	1.12
Ge	3900	1900	0.66
InP	5400	200	1.34
GaAs	8500	400	1.42
InGaAs(53%)	12000	300	0.74
InAs	40000	500	0.35
GaSb	3000	1000	0.73

Figure. 3. Electron and hole mobilities, and energy band gap of IV group and III-V semiconductors.

MOSFETs operate within ~ 20% of the ballistic limit [2]**.** It is clear that co-integration of InGaAs based n-channel with Si has the potential of achieving the ultimate on Si and on the approaches being taken to address the challenges. It should be borne in mind that successful monolithic integration of III-V on Si not only changes the landscape of future CMOS but also opens unparalleled opportunities for integrating opto-electronics, high speed I/Os, and RF devices on Si.

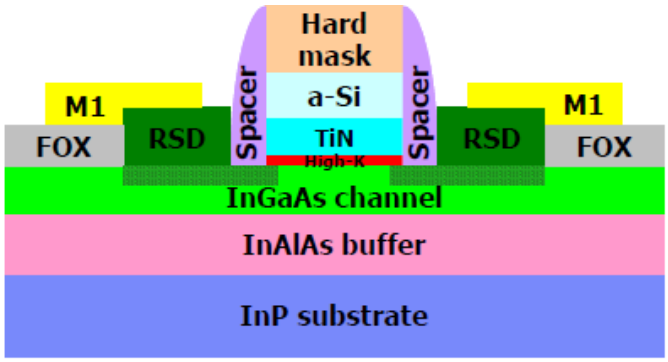


Figure. 4. A schematic diagram of III-V MOSFET with self-aligned contacts. The integration scheme for high-k, metal gate, spacers, and selective epi are analogous to that used for Si MOSFET technology. (ref 2)

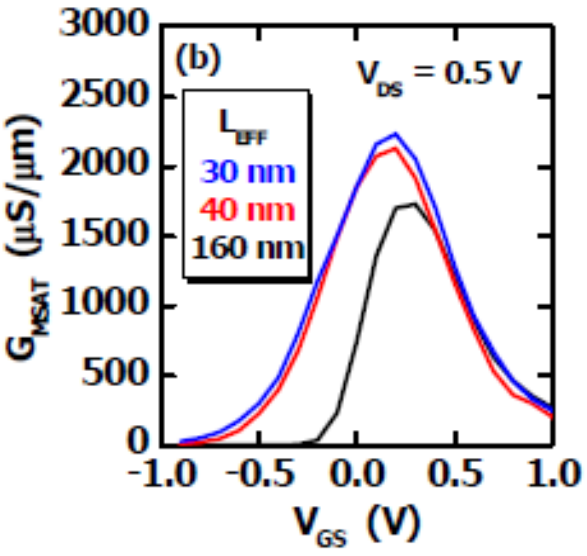


Figure. 5. Saturation transconductance G_{MSAT} vs. V_{GS} characteristics of self-aligned 20-nm-thick $In_{0.53}Ga_{0.47}As$ channel MOSFETs with channel/ barrier doping NA ~ 1 x 10^{17} cm-3, and effective channel Length, L_{EFF} = 30, 40, and 160 nm. The drain bias V_{DS} = 0.5 V. Highest peak G_{MSAT} = 2230 µS/µm achieved at L_{EFF} = 30 nm. (ref 2)

Co-integration of Si and non-Si Channel Materials

An ideal approach to co-integrate III-V with Si would be to bond III-V to Si in an analogous manner to what is routinely done to produce SOI wafers. This is not a manufacturable option due to incompatibility in wafer diameters of III-V and Si starting substrates. Typical product Si wafers are > 200 mm in diameter, these diameters have not yet been reached with III-V materials. This leaves the only known manufacturable option to co-integrate III-V with Si is by epitaxial growth of III-V on Si. However, a large lattice mismatch between $In_{0.53}Ga_{0.47}As$ (8%) and Si leads to a high density of misfit dislocations, stacking faults, micro-twins, and anti-phase boundaries (APDs) at the growth interface. It is interesting to note that in III-V materials, the higher the carrier mobility (Figure 3), the higher its lattice mismatch with Si (Figure 2). Innovative substrate engineering is therefore required to fill the huge gap that currently exists between the required defect density value and that required for CMOS products.

Previously published data both on Si and III-V based devices clearly indicates that misfit dislocations contribute to junction leakage (Figure 10) and degraded device reliability [3]. Success of future high performance CMOS will depend heavily on controlling defect densities such that the target stand by current (I_{off}) values are met. For high performance and low power CMOS products these values are typically ~ 100 nA/µm and a few pA/µm, respectively. Based on previously published results defect density should be controlled to < 10^3 cm^{-2} to meet these I_{off} targets. However, typical defect density for $In_{0.53}Ga_{0.47}As$ on Si is currently at > 10^9 cm^{-2} due to > 8% lattice mismatch [4]. Similarly, defect density for a Ge p-channel grown on Si with 4% lattice mismatch is at > 10^7 cm^{-2} even after post-epi cyclic annealing [5].

Materials: Grand Challenge

The grand material challenge is how to control or eliminate crystallographic defects in highly lattice mismatched III-V channel materials for future CMOS. Decades of research on this topic have led to the following five main schemes for defect reduction: (i) growth of a thick (several microns) graded buffer layer to create a virtual substrate of a desired lattice parameter [6], (ii) post-epi annealing processes which exploit both the thermal energy and CTE (coefficient of thermal expansion) mismatch to mobilize and annihilate defects [7], (iii) layer transfer of the low-defect region of the graded buffer layer on an insulator (e.g., sSOI) [8], (iv) aspect ratio trapping [8,9], and (v) growth of elastically relaxed nanowires [10-12]. Examples of each of these methods and recent progress made in defect reduction is discussed below.

Defect Reduction: Graded Buffer Layers

The concept of a graded buffer layer to create a virtual substrate is known since early 70s and pioneering

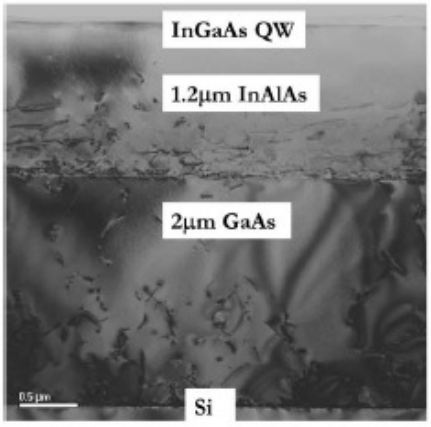


Figure. 6. A cross-section TEM micrograph showing an $In_{0.53}Ga_{0.47}As$ layer grown on Si via buffer layers of GaAs and $InAl_{0.52}As_{0.48}$. Reduction in defect density in the surface region is evident (ref 4).

research was done in early 80s particularly on SiGe graded buffer layers (GBL) to control defects in strain-Si grown above it. A relaxed SiGe GBL creates a larger lattice on Si, i.e., "virtual substrate" and is utilized as the epitaxial template for growing Si in a state of biaxial tension which can have x2 higher electron mobility than conventional Si [6]. A few microns of buffer layer is typically required to reduce defect density. Similar GBL concept has been applied successfully to reduce defect density in $In_{0.53}Ga_{0.47}As$ on Si, Ge, and GaAs substrates (Figure 6) [4]. Typical defect density reduction from an abrupt to graded layer can be several orders of magnitude depending on the thickness of the layer (Figure 9). However, despite a > 3 µm thick layer, the best defect density in $In_{0.53}Ga_{0.47}As$ is still in the 10^8 cm^{-2} range [4].

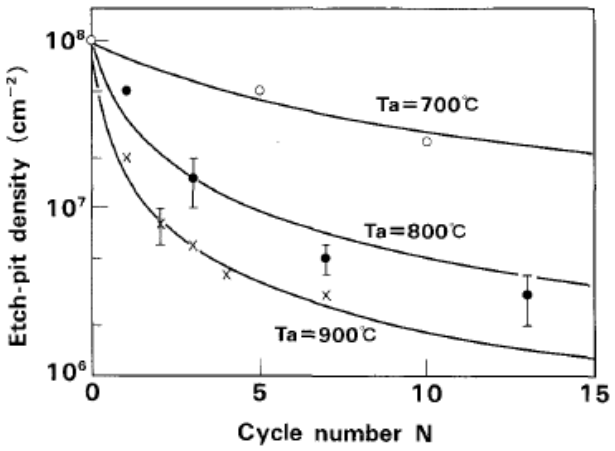


Fig. 7. Defect reduction in GaAs layers grown on Si by cyclic annealing in AsH_3 ambient. Higher the anneal temperature, bigger the defect reduction (ref 7)

Defect Reduction Post-Epi Anneal

Pronounced defect reduction by ~ 3 orders of magnitude (from 10^9 to 10^6 range) by cyclic annealing has been demonstrated on Ge grown on Si despite a 4% mismatch [5]. Typically, a Ge layer of ~ 1 µm is grown followed by multiple cyclic anneals at > 850° C in H_2 under reduced pressure in an epitaxial reactor. A similar approach has been applied to reduce defect density in InP grown on GaAs [7] (Figure 7). It was shown that annealing at > 700°C is essential for defect reduction. However, unlike Ge where post-epi annealing can be performed at near its melting point, this is not possible in III-V materials because of high volatility of the group V element. Consequently, it remains a challenge to significantly reduced defect density in III-Vs via cyclic annealing.

Defect Reduction: Layer Transfer

If we assume that non-Si channels will always have higher crystallographic defect density than bulk-Si, what are our options to minimize their impact on device leakage? A well proven approach to control device leakage on a defective substrate is that by transfer the defective region onto a dielectric layer (SiO_2). Figures 8a and 8b defects in a SiGe GBL and strain-Si, and that in sSOI after the layer transfer of strain-Si onto SiO_2. Figures 9a and 9b show XTEM micrographs corresponding to Figures 8a and 8b. When the defect strain-Si is transferred onto SiO_2 (Figure 9b), I_{off} was reduced by > 6 orders of magnitude (Figure 10).

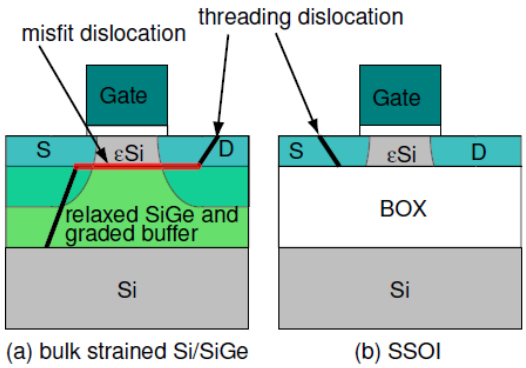


Figure. 8. A schematic diagram showing (a) defect propagation from the SiGe growth interface to the strain-Si surface, (b) layer transfer of the strain Si layer onto SiO_2 grown on Si. (ref 3).

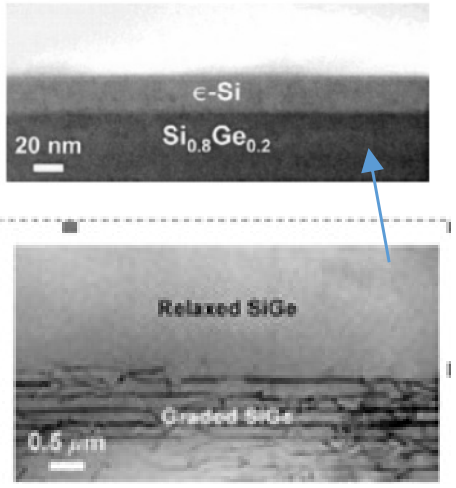
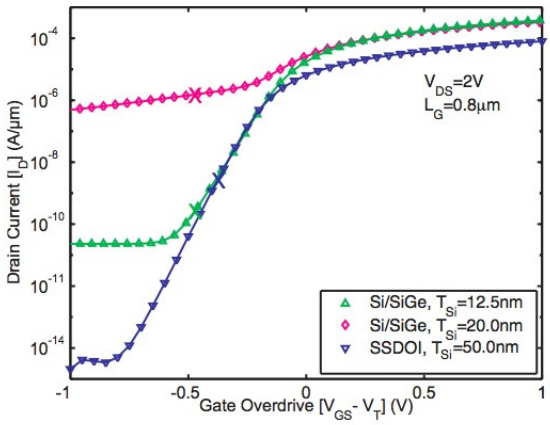


Figure 9. XTEM micrographs corresponding to Figures 8a and 8b diagram showing (a) defects in SiGe GBL/strain-Si and(b) in sSOI (ref 6).



Figu 10. The degrading effect of defects in buffer layer on device leakage can be minimized by converting strain-Si on GBL to strain-Si on insulator (sSOI). This scheme reduces device leakage by six orders of magnitude. (ref 3)

Defect Reduction: Aspect Ratio Trapping (ART)

The most extensively studied method to reduce defects in III-Vs on Si is by ART [8, 9]. Figures 11a and 11b show the concept and implementation of the ART methodology, respectively, to grow InP on Si. Promising results have been achieved in reducing defects in $In_{0.53}Ga_{0.47}As$ (and even in InAs) on Si. The InP grown by ART in Figure 11 serves as the lattice matched template for subsequent growth of $In_{0.53}Ga_{0.47}As$. Defect trapping clearly occurs predominantly at the bottom of the trench resulting in higher structural quality surface region. Recent realization of short channel MOSFET devices fabricated with $In_{0.53}Ga_{0.47}As$ by the ART method on a 300 mm Si substrate serves as the testament that III-V integration with Si is quite possible [9].

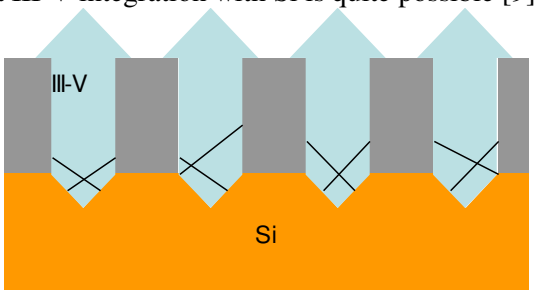


Figure 11a. A schematic diagram showing the concept of ART. Trapping of planar defects occurs at the bottom of the trench with aspect ratio of > 2:1.

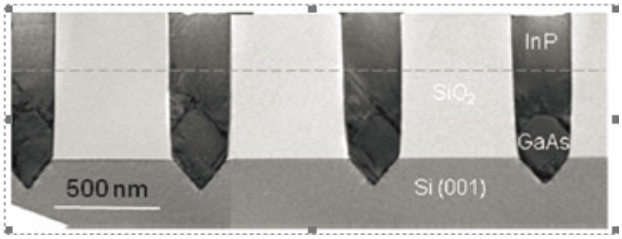


Figure 11b. Demonstration of planar defect trapping at the bottom of a trench when InP is grown on Si with a GaAs buffer layer. (ref 8)

Defect Reduction: Nanowire Growth

Fabrication of III-V nanowires (NWs) grown on Si by both catalyst-free bottom up as well as top down approaches [10-12] is becoming an active area of research because of two reasons: (i) excellent device electrostatics, and (ii) a viable path to dislocation free growth of bottom up $In_{0.53}Ga_{0.47}As$ NWs on Si. Uninhibited elastic relaxation of $In_{0.53}Ga_{0.47}As$ occurs during the NW growth on Si due to availability of free edges thus obviating dislocation formation in the NWs. A varying range of $In_xGa_{1-x}As$ compositions in NWs by both molecular beam epitaxy (MBE) and metalorganic chemical vapor deposition (MOCVD) methods have been successfully demonstrated. Nevertheless, growth induced planar crystallographic defects, such as stacking faults, twins, and antiphase boundary are still present in NWs. Gate-all-around (GAA) NW MOSFETs with excellent device electrostatics have been reported by various groups in the last few years. Figures 12a and 12b show III-V NW MOSFETs from recently published results with well-behaved device electrostatics (DIBL 35 mV V^{-1} and SS slope 75 mV/dec for L_g = 200 nm). Even more impressive electrostatics have been demonstrated on top down InGaAs NWs with 7mV V^{-1} and sub-threshold slope (63mV/dec) [12].

Conclusions

Finally, a number of manufacturing challenges can arise even when a known and successful material solution is

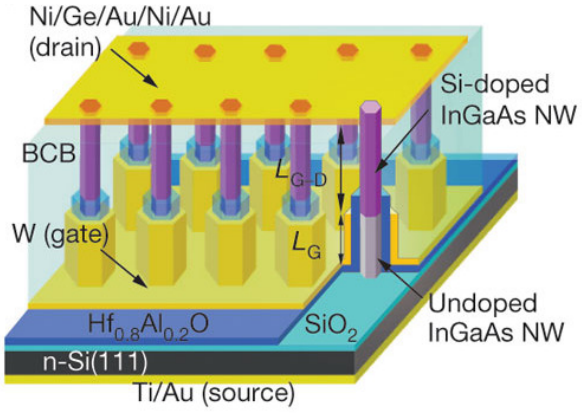


Figure 12a. A schematic diagram showing MOSFETs consisting of an array of bottom up "core-multishell NWs", grown on <111> Si. Layers of InP, InAlAs and InGaAs are grown around an InGaAs core. (ref 10)

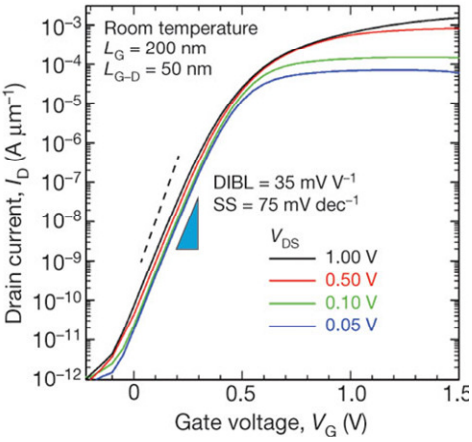


Figure 12b. I_d – V_g characteristics of the MOSFETs of Fig 11a showing excellent device electrostatics. (ref 10)

scaled to a large form factor substrate, i.e., 300 mm Si: (i) compositional uniformity, (ii) thickness uniformity, (iii) variation of defect distribution across the Si wafer, (iii) varying defect density vs pattern dimensions, and (iv) cross contamination of neighboring devices among others. In parallel, rapid progress is required to address formidable device and process challenges associated with III-Vs including demonstration of low leakage devices at < 20 nm channel length, integration schemes for p-FETs which are compatible with III-V processing, high-k/metal gate stack with interface charge of < 10^{11} cm^2 eV^{-1}, contact resistance of < 10^{-9} ohm-cm^2 and so on. Whether Si/non-Si co-integrated technology for future nodes will be ready in the timeframe that is necessitated by the CMOS scaling roadmap remains debatable.

Acknowlegments

The author gratefully acknowledges useful input from many colleagues in IBM. In particular, discussions with Drs Ghavam Shahidi, Effendi Leobandung, and T C Chen, and the III-V integration team are highly appreciated.

References

[1] S. Narasimha P. Chang, C. Ortolland, D. Fried†, E. Engbrecht, K. Nummy, P. Parries, T. Ando2, M. Aquilino, N. Arnold,R. Bolam3, J. Cai2, M. Chudzik, B. Cipriany, G. Costrini, M. Dai, J. Dechene, C. DeWan, B. Engel, M. Gribelyuk, D. Guo,G. Han, N. Habib, J. Holt, D. Ioannou3, B. Jagannathan, D. Jaeger, J. Johnson, W. Kong, J. Koshy, R. Krishnan, A. Kumar2,M. Kumar, J. Lee, X. Li, C-H. Lin2, B. Linder2, S. Lucarini, N. Lustig, P. McLaughlin, K. Onishi, V. Ontalus, R. Robison, C. Sheraw,M. Stoker, A. Thomas, G. Wang, R. Wise, L. Zhuang, G. Freeman, J. Gill, E. Maciejewski, R. Malik, J. Norum & P. Agnellol, "22nm High-performance SOI technology featuring dual-embedded stressors, Epi-Plate High-K deep-trench embedded DRAM and self-aligned Via 15LM BEOL," IEDM 2014, pp 3.3.1 – 3.3.4.

[2] Y. Sun, Y. Sun, A. Majumdar, C.-W. Cheng, R. M. Martin, R. L. Bruce, J.-B. Yau, D. B. Farmer, Y. Zhu, M. Hopstaken, M. M. Frank, T.Ando, K.-T. Lee, J. Rozen, A. Basu, K.-T. Shiu, P. Kerber, V. Narayanan, R. Mo, D. K. Sadana, and E. Leobandung et al, "CMOS compatible self-aligned In0.53Ga0.47As MOSFETs with *G*MSAT over 2200 mS/mm at *V*DD = 0.5 V, " IEDM 2014, pp 25.3.1 – 25.3.4.

[3] I. Lauer, T.A. Langdo,Z-Y. Cheng, J.G. Fiorenza, G. Braithwaite, M.T. Currie,C.W. Leitz, A. Lochtefeld, H. Badawi,M.T. Bulsara, M. Somerville, D.A. Antoniadis, "Fully depleted n-MOSFETs on supercritical thickness strained SOI," EDL, vol25, 83, 2004.

[4] S. Dutta, G. Dewey, J. M. Fastenau, M. K. Hudait, D. Loubychev, W. K. Liu, M. Radosavljevic, W. Rachmady, and R. Chau, "Ultrahigh-Speed 0.5 V Supply Voltage InGaAs Quantum-Well Transistors on Silicon Substrate," EDL, Vol 28, 685, 2007.

[5] H-C. Luan, D. R. Lim, K. K. Lee, K. M. Chen, J. G. Sandland, K. Wada[1] and L. C. Kimerling, Appl. Phys. Lett. 75, 2909, 1999.

[6 M.L. Lee, E. A. Fitzerald, M.T. Bulsara, M. T. Currie, and A. Lotchefeld, "Strained Si, SiGe, and Ge channels for high-mobility metal -oxide-semiconductor field-effect transistors," J. App. Phys. 97, 011101, 2005

[7] Norio Hayafuji, Tatsuya Kimura, Naohito Yoshida, Nobuaki Kaneno, Mari Tsugami, Kazuo Mizuguchi, Toshio Murotani and Sumiaki Ibuki *et al, Jpn. J. Appl. Phys.* 28, L1721, 1989.

[8] J. G. Fiorenzaa, J.-S. Parka, J. M. Hydricka, J.Lia, J. Z. Li b, M. Curtinb, M. Carrollb, A. Lochtefeld, "Aspect Ratio Trapping: a Unique Technology for Integrating Ge and III-Vs with Silicon CMOS," ECS Transactions, 33 (6) 963-976, 2010.

[9] N. Waldron, C. Merckling, W. Guo, P. Ong, L. Teugels, S. AnsarI, D. Tsvetanova, F. Sebaai, D.H. van Dorp, A Milenin, D. Lin, L. Nyns, J. Mitard,A Pourghaderi, B. Douhard, O. Richard, H. Bender, G. Boccardi, M. Caymax, M. Heyns, W. Vandervorst, K. Barla, N. Collaert and AV-Y. Thean, "An InGaAs/InP Quantum Well FinFet Using the Replacement Fin Process Integrated in an RMG Flow on 300mm Si Substrates," 2014 Symposium on VLSI Technology Digest, pp1-2.

[10] Katsuhiro Tomioka, Masatoshi Yoshimura & Takashi Fukui, "A III–V nanowire channel on silicon for high-performance vertical transistors," Nature 488, 189–192, 2012.

[11] J. J. Gu, O. Koybasi, Y. Q. Wu, and P. D. Ye, " III-V-on-nothing metal-oxide-semiconductor field-effect transistors enabled by top-down nanowire release process: Experiment and simulation," App Phys Lett, 99, 112113, 2011.

[12] J. A. del Alamo, "Nanometre-scale electronics with III-V compound semiconductors," Nature, vol. 479, no. 7373, pp. 317–323, Nov. 2011.