CMOS directions in industry

Ernesto PEREA

ST Microelectronics

Crolles2 Alliance

CMOS Directions in Industry

12th Workshop on Electronics for LHC and futute Experiments

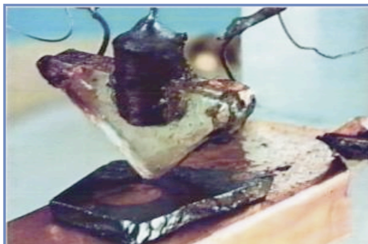
25-29 September 2006, Valencia, Spain

Ernesto Perea
ST Microelectronics
Crolles2 Alliance – France
ernesto.perea@st.com

SUMMARY

- Moore's-Law scaling impact on digital and analog circuits
- Regular reconfigurable structures (Networks on Chip, Multi-core Architectures, and Sampled Analog and RF circuits.
- More than Moore and 3D System Integration
- Fundamental physical limiting effects (oxide tunneling, finite sub-threshold slope, etc)
- New CMOS structures (new dielectric materials, band-gap engineering
- Statistical nature of parameter variabilities
- Device architectures beyond 45nm (new metal-dielectrics gate stacks, stressed-strained channels, etc.
- New tightly-coupled system-architecture-circuit-device design.
- Conclusion: Where is all this going?: The Process, Circuit, Architecture and System Design, and Manufacturing unique synergies and new innovative Cooperation-Competition industrial schemes.

Point Contact Semiconductor Amplifier

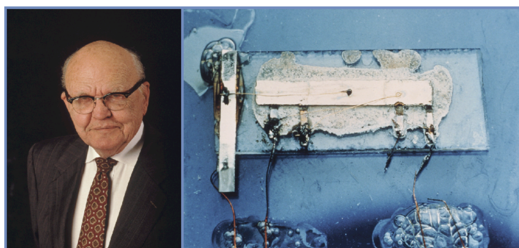


By Bardeen and Brattain, Dec. 16, 1947

1956 Nobel Prize for Transistor Invention at AT&T Bell Labs

J. Bardeen W. Shockley W. Brattain

Texas Instrument, 1st IC in 1958

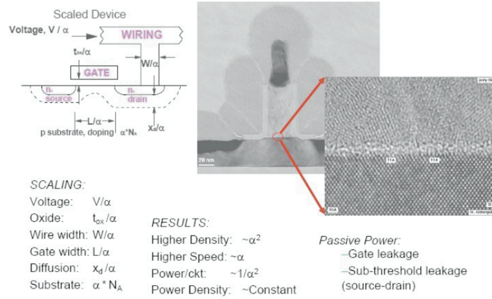


Jack Kilby (Nov. 8, 1923- June 7, 2005) 2000 Nobel
1Tr (Germanium), 1C, 3R, Oscillator, 0.04 inch X 0.06 inch.

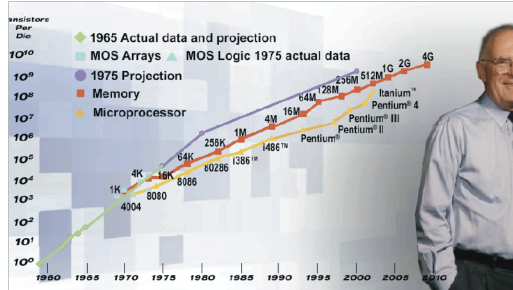
Scaling:

Moore's-Law scaling impact on digital and analog circuits

CMOS Scaling Rules

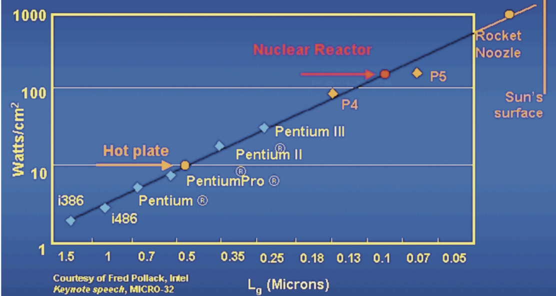


SOC Integration: Scaling and Digital

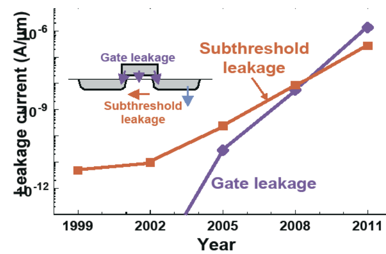


Source: Intel

Source: Intel, Intel Developers Forum, 2/2003



Drain and Gate increase in leakage currents

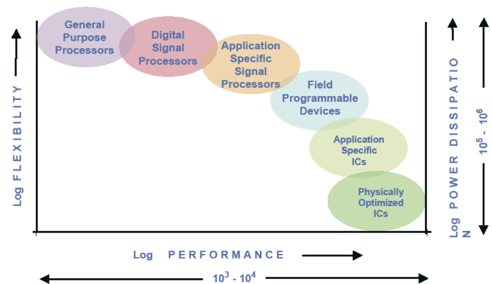


Soft error rates increases 2X for SRAMs at each new proces node and Single event transient below 45nm could escalate 3 decades.

Reasons:Vdd decrease, Q nodes decrease, higher clock speeds.

- **Signal Integrity, ElectroMagnetic Coupling and RF (SI/EMC/RF) solutions.**
- **Single and Multiple Event Upsets protections at architectural,circuit and device levels.**
- **Single event latch up.**
- **Hardened libraries (process and design levels).**
- **Fail safe architectures.**
- **…**

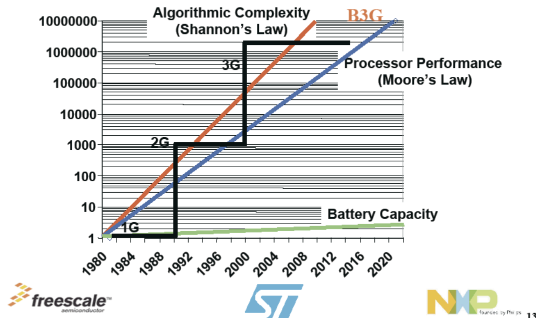
Implementation Space



Source: T.Noll, RWTH Aachen

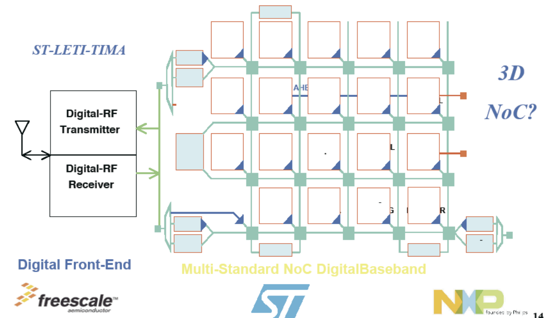
The Algorithmic Driving Force

Shannon asks for more than Moore can deliver...



freescale semiconductor

Reconfigurable Globally Asynchronous-Locally Synchronous Network-on-Chip Transceiver Architecture

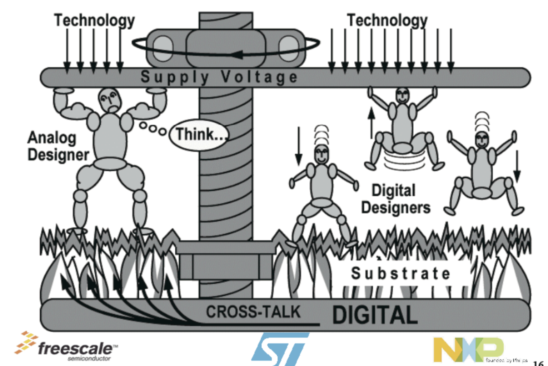


freescale semiconductor

Key Processing Elements (PE's) Required

- Multistandard operation will require:
  - *FIR Filters*
  - *FFTs, IFFTs*
  - *Viterbi and Turbo Decoders*
  - *Rakes*
  - *Synchronizer Blocks*
  - *....etc*
- These parametrable blocks can be arranged in a programmable NoC architecture:
  - *Array of heterogeneous processing elements*
  - *Programmable / parameterizable PEs (reconfigurable logic)*
  - *Routing interconnect*
  - *Soft loadable*

Analog and Digital Scaling Challenges



freescale semiconductor

Panic...

...But it's nice to have Good Digital Neighbors...

- Digital Compensation can be included almost for free:
  - Account for Analog Imperfections through Digital Compensation (Linearity, Offsets, etc)
  - Digital self-calibration and dynamic element matching can be used at small feature size and low voltage.

Analog Systems Considerations/Trends

- Discrete time techniques will work better in scaled submicron CMOS
  - Switched-capacitor (SC) techniques, Bucket-Brigade, CCD's, etc.
  - Switches usually need complementary devices ($V_{DD} > V_{TN} + V_{TP}$)
  - For analog: lower V_T typically offer:
    - ✓ Faster switches
    - ✓ Better amplifier input range
- Dynamically adjust supply and threshold design parameters to center the design in the presence of process variations!
- Digital Compensation can be included almost for free:
  - Account for Analog Imperfections through Digital Compensation (Linearity, Offsets, etc)
  - Digital self-calibration and dynamic element matching can be used at small feature size and low voltage.
- Auto-ranging, Self-Calibration, Self-Test, Active Linearization, Reconfigurability allowing a different performance/cost tradeofs.
  - These functions will have an impact on the Silicon cost, largely compensated by the Total Cost Reduction induced.

Continuous Time versus Discrete Time and Scaling

- Discrete time techniques will work better in scaled submicron CMOS
  - Switched-capacitor (SC) techniques, Bucket-Brigade, CCD's, etc.
  - Switches usually need complementary devices ($V_{DD} > V_{TN} + V_{TP}$)
  - For analog: lower V_T typically offer:
    - Faster switches
    - Better amplifier input range

Reconfigurable Front-End Radios

- The concept of Integrated Reconfigurable Radios (SW Defined Radios = SDR) has been around for many years, but it regains interest for mobile terminals now because advances in silicon technologies make it possible.
- The reason for this regain of interest is also the expectation that a unique radio architecture, reconfigurable by SW to different standards will add flexibility and strongly reduce costs, by simplifying board development, calibration, and BOM, reducing the total cost seen by the phone manufacturer.

A full Mobile System-on-Chip:

- Full HW/SW integration of
  - Communication + Application + Intelligence

Intelligence should include:

- For digital functions:
  - Fine-grain integrated power management and control, including all available techniques at all levels (from SW to partitioning, design and process). Also Reconfigurability for standards.
  - Design for Test, Self-Test, Fault observability, detection and correction.
- For Front-End (RF and ABB) functions:
  - Auto-ranging, Self-Calibration, Self-Test, Active Linearization, Reconfigurability allowing a different performance/cost tradeofs.

HW/SW overhead:

- These functions will have an impact on the Silicon cost, largely compensated by the Total Cost Reduction induced.

…But, What about the Reconfigurable Digital Baseband ?

More Than Moore's Law (ITRS2005)

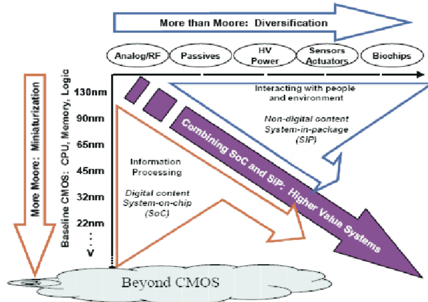
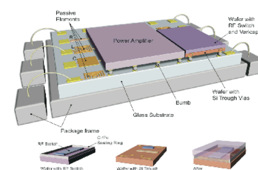


Figure 5 Moore's Law and More

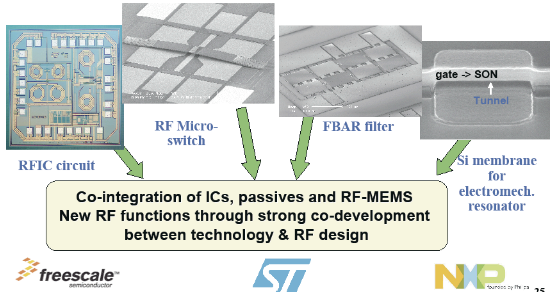
SOC + SIP Integration

- Module integration offers the possibility to integrate different technologies together
  - SOC Transceiver
  - GaAs or Si (Power Amplifier)
  - Passive components
  - RFMEMS components



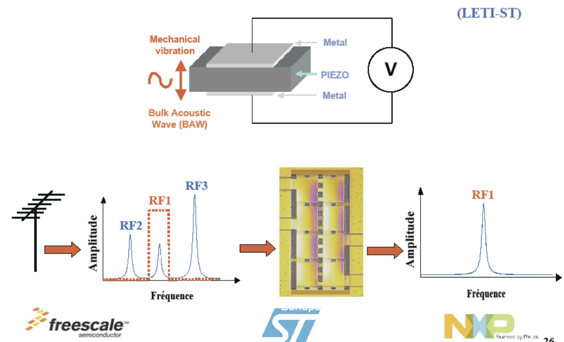
RFMEMS Integration

- Allows System on Chip (SoC) approach
- Opens the door to Innovative RF Front-End Architectures

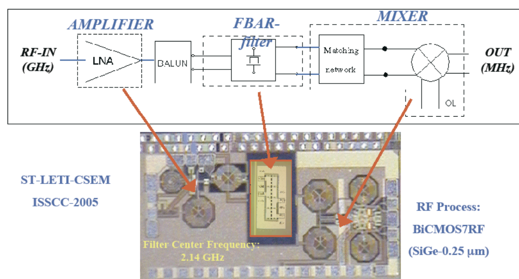


BAW Filter Operating principle

(LETI-ST)



RF Front-End FBAR integration



ST-LETI-CSEM

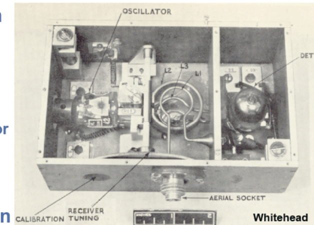
ISSCC-2005

RF Process: BiCMOS7RF (SiGe-0.25 µm)

Challenge: Process Integration and new Architectures

The Old Super-Regenerative Receiver

- Invented by Armstrong in 1922
- Possible radio operation above device f_T, allowing:
  - -Subthreshold biasing and/or
  - -High carrier frequency, increasing integration
- Modern devices/MEMS can solve traditional problems with architecture:



1940s - Super-regenerative

f_c= 500MHz (above f_T)

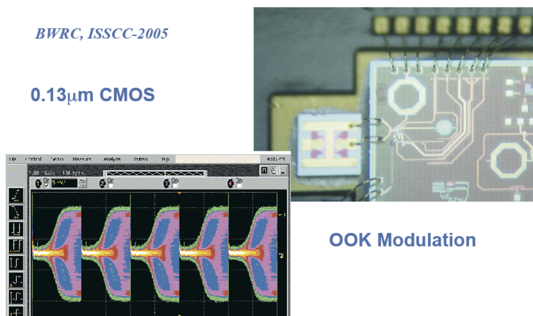
2 active devices

high quality passives – hand tuning required

Receiver Implementation

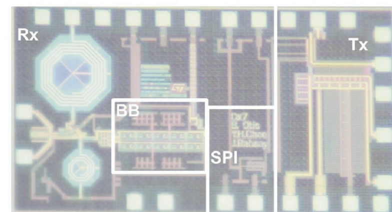
BWRC, ISSCC-2005

0.13µm CMOS



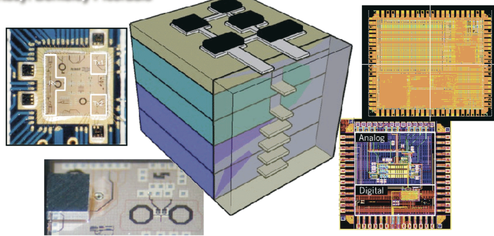
OOK Modulation

Next Generation: (1x1.9)mm² Rx/Tx



- Above IC Components (inductors, crystals, capacitors)
- Full Digital control of analog/RF blocks
  - -Rx Osc Frequency: 3.15MHz x 500kHz
  - -BB Filter BW/Gain
  - -Tx Tank Frequency

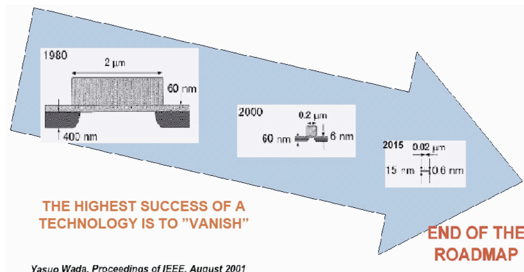
Dealing with Mixed-Everything through Packaging



SIP for Optimum Technology Mix

- Close to the antenna, advanced SiGe:C Bipolars will continue to play an important role for very low NF @ Low-Power, as well as for RF Power Amplification @ High Efficiency, for battery life.
- Medium Voltage / Power functions such as Energy Management and Audio will continue to be handled by adapted MOS/Bip Technologies for reliability and lowest cost. High density/quality capacitors will be key differenciators.
- This Technology Mix will be possible by SIP techniques also allowing Hi-Q Passives and MEMS Integration.

The Ultimate Nanometer MOS

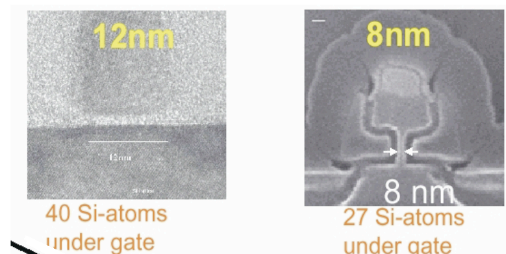


WAYS BACK TO HEALTHY SCALING

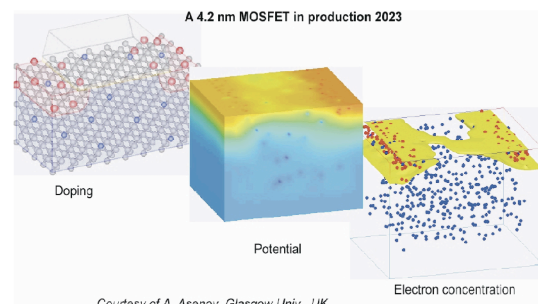
TECHNOLOGY BOOSTERS

- H- K DIELECTRICS (Improves MOS behavior)
- METAL GATES (Improves Ion/Ioff)
- STRAINED Si, SiGe and Ge (Improves Mobility)
- DEVICE STRUCTURE (Reduces Ioff)
- SOI / SON (Improves speed, Reduces Ioff)
- MULTIPLE GATES (Improves MOS, Reduces Ioff)
- BALLISTIC TRANSPORT (Improves Speed)

REAL NANO- MOSFETs



The Discrete Nature of Matter is Visible



FUTURE MOS ARCHITECTURE EXAMPLES

Bulk	PD SOI	FD SOI	FD SON	DG (Delta, FinFET, SON, Vertical, TriGate, Omega, etc., etc.
Scalability ?	Scalability as BULK	Scalability may be better or worse (GP,BOX)	Scalability much improved if GP	Scalability very much improved

But…How do we do any useful design with these devices? Will they buy us any renewed life for Digital / Analog / RF ??

We need to start working on these questions NOW !

Gate, Source, Drain

Tied Gates side-wall conduction: DELTA, FINFET, OMEGAFET,…

GATE, SOURCE, DRAIN, Si-substrate, STI

Tied gates planar conduction: DG SON

ITRS Roadmap 2005 Edition

Year of Production	2005	2006	2007	2008	2009	2010	[illegible]	2012	2013
DRAM ½ Pitch (nm) (contacted)	80	70	65	57	50	[illegible]	40	36	[illegible]
MPU/ASIC Metal 1 (M1) ½ Pitch (nm)(contacted)	90	78	68	59	52	45	40	36	32
MPU Physical Gate Length (nm)	32	28	25	22	20	18	16	14	13
L_g: Physical L_{gate} for High Performance logic (nm) [1]	32	28	25	22	20	18	16	14	13
EOT: Equivalent Oxide Thickness [2]									
Extended planar bulk (Å)	12	11	11	9	7.5	6.5	5	5	
UTB FD (Å)				9	8	7	6	5	5
DG (Å)							[illegible]	7	6
Gate Poly Depletion and Inversion-Layer Thickness [3]									
Extended Planar Bulk (Å)	7.3	7.4	7.4	2.9	2.8	2.7	[illegible]	2.5	
UTB FD (Å)				4	4	4	[illegible]	4	4
DG (Å)							[illegible]	[illegible]	[illegible]
EOT_{elec}: Electrical Equivalent Oxide Thickness in inversion [4]									
Extended Planar Bulk (Å)	19.3	18.4	18.4	11.9	10.3	9.2	[illegible]	7.5	
UTB FD (Å)				13	12	11	[illegible]	9	9
DG (Å)							[illegible]	11	10
$J_{g,limit}$: Maximum gate leakage current density [5]									
Extended Planar Bulk (A/cm²)	1.88E+02	5.36E+02	8.00E+02	9.09E+02	[illegible]	1.56E+03	[illegible]	2.43E+03	
FDSOI (A/cm²)				7.73E+02	9.50E+02	1.22E+03	[illegible]	2.07E+03	2.23E+03
DG (A/cm²)							[illegible]	7.86E+02	[illegible]
$I_{sd,leak}$: Source/Drain Subthreshold Off-State Leakage Current [8]									
Extended Planar Bulk (µA/µm)	0.06	0.15	0.2	0.2	0.22	0.28	0.32	0.34	
UTB FD (µA/µm)				0.17	0.19	0.22	0.22	0.29	0.29
DG (µA/µm)							0.1	0.11	0.11
$I_{d,sat}$: effective NMOS Drive Current [9]									
Extended Planar Bulk (µA/µm)	1020	1130	1200	1570	1810	2050	2490	2300	
UTB FD (µA/µm)				1486	1625	1815	[illegible]	[illegible]	2198
DG (µA/µm)							1899	1932	2220
Mobility Enhancement Factor for $I_{d,sat}$ [10]									
Extended Planar Bulk	1.09	1.09	1.08	1.09	1.10	1.10	1.12	1.11	
UTB FD				1.06	1.06	1.06	1.06	1.05	1.05
DG							1.05	1.04	1.05

Products ready in 2010

High-k + Metal gate

Bulk or FDSOI

Mobility booster
- Process induced strain
- Crystal orientation
- sSOI
- Dual channel

CMOS Worlwide Trends

(100) (110) (100) nMOS pMOS

Tensile compressive SiC SiGe nMOS pMOS

nMOS pMOS

Mobility enhancement
Orientation for pMOS

Strain
Tensile strain for nMOS
Compressive strain for pMOS

Dual MG + Higk-K
for reduced gate tunneling

IEDM-2005 Trends

- **Standard CMOS platform**
  - Main 65nm and 45nm general plaform stay in bulk conv.
- **Strained Silicon**
  - Intensively full strain sources re-use
    - NMOS= SMT+high tensile stress
    - PMOS= eSiGe+high compressive stress
  - Very interesting IBM paper using hybrid subtsrate types (bulk)

Strained-Silicon Performance Enhancement

PMD liner, Capped poly, Raised S/D, Recess S/D epi, Implant S/D strain, STI liner, Substrate orientation, Strained channel, Active area epi

Normal Si lattice, Strained Si lattice, Improved mobility

Significant Ion/Ioff improvement

35 % improvement, 40 nA/µm, 680 µA/µm, Reference, SiGe

Ioff log (A/µm): -6, -7, -8, -9, -10

Ion (µA/µm): 300, 500, 700

Defect in Silicon

Ni silicide, Poly

1. Strained Si can produce an increase in I_{on} without a degradation in I_{off}.
2. Stress-induced defects can occur in the silicon if the stress is not applied properly.

Impact of High-k on Gate-Dielectric Leakage

Gate Leakage

SiO_2 Trendline

HfSiON

△ nMOSCAP HfSiON
○ pMOSCAP HfSiON
— NMOS SiO_2 Trendline
● nMOSCAP HfON

J_g @±(|V_{fb}|+1) (A/cm²): 10^{4}, 10^{3}, 10^{2}, 10^{1}, 10^{0}, 10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}, 10^{-5}

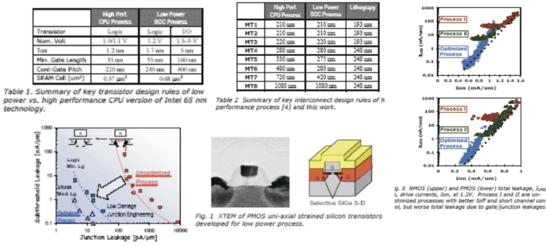
EOT (Å): 5, 10, 15, 20, 25

$$EOT = \left(\frac{k_{SiO_2}}{k_{high-k}}\right) \bullet (thickness)_{high-k}$$

High-k thickness can be kept relatively thick, to control leakage, while EOT can be made quite low.

Paper 3.5

A 65nm Ultra Low Power Logic Platform Technology using Uni-axial Strained Silicon Transistors Intel Corporation



freescale semiconductor · ST · NXP

Paper 3.4

Design of High Performance PFETs with Strained Si Channel and Laser Anneal,

IBM System & Technology Group, 1Chartered Semiconductor Manufacturing, Ltd,
2Infineon Technologies AG, 3Samsung electronics Co., Ltd,
4Toshiba America Electronic Components, Inc.,

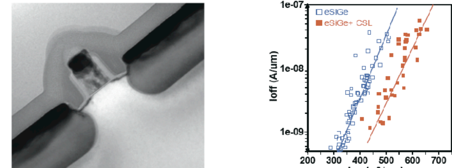


Fig. 2 Cross-section of a typical transistor with e-SiGe and CSL

Fig. 5 Ion vs. Ioff at |Vdd|=1V of PFETs with e-SiGe and PFETs with e-SiGe + CSL.

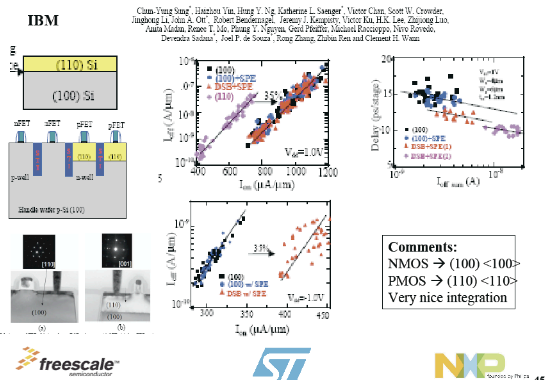
freescale semiconductor · ST · NXP

Paper 10.3

IBM

High Performance CMOS Bulk Technology Using Direct Silicon Bond (DSB) Mixed Crystal Orientation Substrates

Chun-Yung Sung*, Haizhou Yin, Hung Y. Ng, Katherine L. Saenger*, Victor Chan, Scott W. Crowder, Jinghong Li, John A. Ott*, Robert Bendernagel, Jeremy J. Kempisty, Victor Ku, H.K. Lee, Zhijiong Luo, Anita Madan, Renee T. Mo, Phung Y. Nguyen, Gerd Pfeiffer, Michael Raccioppo, Nivo Rovedo, Devendra Sadana*, Joel P. de Souza*, Rong Zhang, Zhibin Ren and Clement H. Wann



Comments:
NMOS → (100) <100>
PMOS → (110) <110>
Very nice integration

freescale semiconductor · ST · NXP

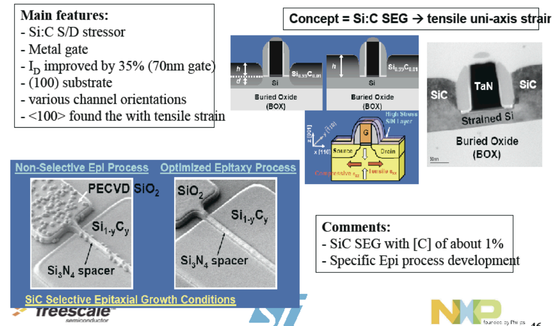
Paper 20.3

Thin Body Silicon-on-Insulator N-MOSFET with Si:C S/D Regions for Performance Enhancement

Singapore University

Main features:
- Si:C S/D stressor
- Metal gate
- I_D improved by 35% (70nm gate)
- (100) substrate
- various channel orientations
- <100> found the with tensile strain

Concept = Si:C SEG → tensile uni-axis strain



Comments:
- SiC SEG with [C] of about 1%
- Specific Epi process development

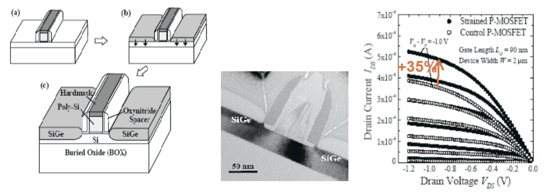
freescale semiconductor · ST · NXP

Source/Drain Germanium Condensation for P-Channel Strained Ultra-Thin Body Transistors

National University of Singapore

This paper reports a novel technique to fabricate uniaxial compressive strained P-channel transistors with Silicon-Germanium (SiGe) source and drain (S/D) stressors.

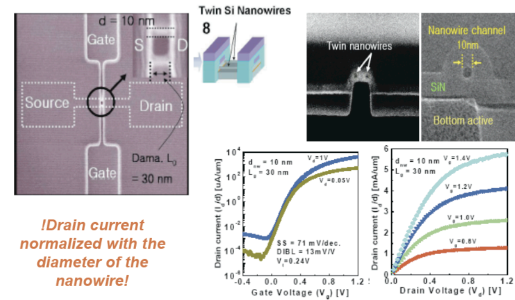
The process involves local Ge condensation of a selectively grown SiGe region, thus driving Ge into and enriching the Ge concentration in the source and drain regions adjacent to the transistor channel.



freescale semiconductor · ST · NXP

High Performance 5nm radius Twin Silicon Nanowire MOSFET(TSNWFET) : Fabrication on Bulk Si Wafer, Characteristics, and Reliability

Samsung Electronics Co.



!Drain current normalized with the diameter of the nanowire!

freescale semiconductor · ST · NXP

Paper 36.7

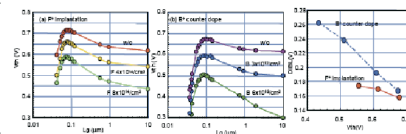
Vth-tunable CMIS Platform with High-k gate Dielectrics and Variability Effect for 45nm node

T. Hayashi, M. Mizutani, M. Inoue, J. Yugami, J. Tsuchimoto, M. Anma, S. Komori, K. Tsukamoto, Y. Tsukamoto, K. Nii, Y. Nishida, H. Sayama, T. Yamashita, H. Oda, T. Eimori, and Y. Ohji

Renesas Technology Corp., Wafer Process Engineering Development Dept.

Renesas

F implant for Vth correction (poly/ HfSiON!)



615/221 µA/µm @ 20pA/µm (V_{dd}= 1.2V)

	This work	ref.[9]	ref.[10]	ref.[11]	ref.[12]
Lg(nm)	55	55	60	90	65
Vdd(V)	1.2	1.2	1.2	1.2	1.2
dielectrics	HfSiON	HfSiON	HfO_2-Al_2O_3	HfSiON	HfSiON
Ion n/p (µA/µm)	615 / 221	510 / 220	335 / 115	430 / 197	430 / 160
Ioff n/p (pA/µm)	20 / 20	20 / 20	0.9 / 2.0	25 / 10	16 / 16
Ig n/p (pA/µm)	17 / 1.9	2.5 / 0.84	-	6.4 / 0.9	1 / -

Table 1. Performance comparison with other papers.

30.1 **High Performance Multi-Gate pMOSFETs using Uniaxially-Strained SGOI Channels**

MIRAI-ASET+, MIRAI-AIST++, Toshiba Ceramics*

• "Ultimate" PMOS device combining Multi-gate, SiGe channel, uniaxial strain & (110) conduction
• Strong performance improvement demonstrated

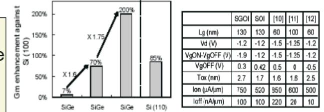


Fig.13 g_m enhancement against Si (100) resulting from each performance enhancement booster. Note that all the boosters included in uniaxially-strained SGOI FinFETs are additive.

	SGOI	SOI	[10]	[11]	[12]
Lg (nm)	130	130	60	100	60
Vd (V)	-1.2	-1.2	-1.5	-1.25	-1.2
VgON-VgOFF (V)	-1.9	-1.2	-1.5	-1.25	-1.2
VgOFF (V)	0.3	0.42	0.5	0	-0.5
Tox (nm)	2.7	1.7	1.6	1.6	2.5
Ion (µA/µm)	750	520	850	600	500
Ioff (nA/µm)	100	100	220	20	10

Table1 Comparison of I_{on} in FinFETs with previously reported data. Current is normalized by 2*H_{fin}. Taking into account the T_{ox} difference, 45% enhancement of I_{on} against control SOI is obtained.

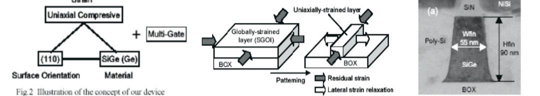


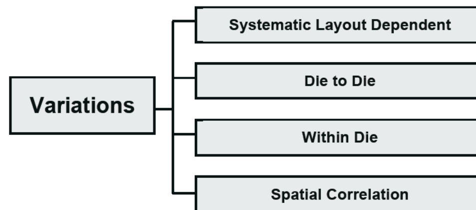
Fig.2 Illustration of the concept of our device structure. Uniaxially-strained SGOI multi-gate MOSFET can be the best solution for pMOS in terms of strain, surface orientation and material.

Fig.4 Schematics illustrating lateral strain relaxation of a biaxially-strained substrate. Uniaxially-strained layers are realized by using this technique.

Variability matters !

- Identify main sources
- SRAM test structures
- Analog test structures (Diff. Pair matching, Switches, etc ..+ passives…)
- Digital basic cells (RO's, Adders, Latches, etc)
- Environment measurements (Vdd noise, couplings, etc)
- Wiring RC delay characterization
- Using variability information
  - Exhaustive Data Mining approach
  - Variability modeling
  - Yield models
- Impact on Circuit / Device / Process design choices

Categorizing Variations



Systematic and Random Device Variations

Parameter	Random	Systematic
Channel dopant concentration Nch	Affects Δvth	Non-uniformity in the process dopant implantation dosage diffusion
Gate Oxide thickness Tox	Gate- dielectric-channel interface roughness	Non-uniformity in the process oxide growth
Gate length L	Line edge roughness	Lithography Proximity effects Resist development etching, etc

The Importance of Variability

- Absolute amount of variability for a particular design
  - Often captured in design corners
  - Determines power penalty for guaranteeing performance
  - Limits benefits of technology scaling
  - Avoid pessimistic Over-Design
- Nature of variability
  - Within-die (WID), Die-to-die (D2D), Wafer-to-wafer (W2W), Fab-2-Fab
  - Spatial vs. Temporal (incl. intermittent)
  - Systematic vs. random
  - Correlated vs. uncorrelated

Digital Systems Considerations/Trends

- Process and design interaction: DFM
  - Physical design impacts yield
  - Power management and Analog/RF integration require more extensive modeling
  - Distributed test and debug functions
- Fine-grain integrated power management and control necessary, including all available techniques at all levels (from SW to partitioning, design and process). not all transistors need min Lg or LVt.
- Scaling: Intrinsic chip speed grows faster than I/O bandwidth: Data will be trapped on-chip.
- Layout/Routing techniques will outperform Low-K approaches.
- Multi-Core Trend: Simpler cores, smaller clock zones, shorter wires, lower latencies: Network-on-Chip, higher IP-reuse, higher parallelism/multi-threading, smaller cores/redundancy: improved yields, self-repair.
- Needed: more design for testability, better yield models, higher I/O bandwidth, better system-level power reduction.

Conclusion

Where is all this going?:
Process, Circuit, Architecture and System Design, and Manufacturing unique synergies and new innovative Cooperation-Competition industrial schemes.

…Evolution of Integration…

- *SSI*
- *MSI*
- *LSI*
- *VLSI*
- *SoC (System on a Chip)*
- *NoC (Network on a Chip)*
- *LoC (Laboratory on a Chip)*
- *CiH (Chip in a Human)*
- *…HiC ?... maybe not…*

There is plenty of room at the bottom

(Richard Feynman, 1960)

- "Can we write the Encyclopedia Brittanica on the head of a pin?"
- "Can we write all books into the volume of a pin?" (3D)
- "How do we writ it?"
- How do we read it?"
- "Can we see atoms?"
- "Can we move and manipulate individual atoms?"
- "Can we build machines at the atomic level?"
- "Can we synthesize substances with physics, not chemistry?"
- "Can we build an atomic-scale computer?"

The Importance of Variability

- Absolute amount of variability for a particular design
  - Often captured in design corners
  - Determines power penalty for guaranteeing performance
  - Limits benefits of technology scaling
  - Avoid pessimistic Over-Design
- Nature of variability
  - Within-die (WID), Die-to-die (D2D), Wafer-to-wafer (W2W), Fab-2-Fab
  - Spatial vs. Temporal (incl. intermittent)
  - Systematic vs. random
  - Correlated vs. uncorrelated

Digital Systems Considerations/Trends

- Process and design interaction: DFM
  - Physical design impacts yield
  - Power management and Analog/RF integration require more extensive modeling
  - Distributed test and debug functions
- Fine-grain integrated power management and control necessary, including all available techniques at all levels (from SW to partitioning, design and process). not all transistors need min L_g or LV_t.
- Scaling: Intrinsic chip speed grows faster than I/O bandwith: Data will be trapped on-chip.
- Layout/Routing techniques will outperform Low-K approaches.
- Multi-Core Trend: Simpler cores, smaller clock zones, shorter wires, lower latencies: Network-on-Chip, higher IP-reuse, higher parallelism/multi-threading, smaller cores/redundancy: improved yields, self-repair.
- Needed: more design for testability, better yield models, higher I/O bandwidth, better system-level power reduction.

Conclusion

Where is all this going?:
Process, Circuit, Architecture and System Design, and Manufacturing unique synergies and new innovative Cooperation-Competition industrial schemes.

…Evolution of Integration…

- *SSI*
- *MSI*
- *LSI*
- *VLSI*
- *SoC (System on a Chip)*
- *NoC (Network on a Chip)*
- *LoC (Laboratory on a Chip)*
- *CiH (Chip in a Human)*
- *…HiC ?... maybe not…*

There is plenty of room at the bottom

(Richard Feynman, 1960)

- "Can we write the Encyclopedia Brittanica on the head of a pin?"
- "Can we write all books into the volume of a pin?" (3D)
- "How do we writ it?"
- How do we read it?"
- "Can we see atoms?"
- "Can we move and manipulate individual atoms?"
- "Can we build machines at the atomic level?"
- "Can we synthesize substances with physics, not chemistry?"
- "Can we build an atomic-scale computer?"