

## 6.1 Overview

Feature size, pitch, node

## 6.2 State of the Art MOSFETs

Electrical and geometrical parameters

## 6.3 Mobility Enhanced MOSFETs

Strain, SiGe, Liners, Directed Channels, Gate-Last

## 6.4 High-k Metal-Gate MOSFETs

Poly-Depletion, FuSi

## 6.5 Vertical MOSFETs

## 6.6 Halo MOSFETs

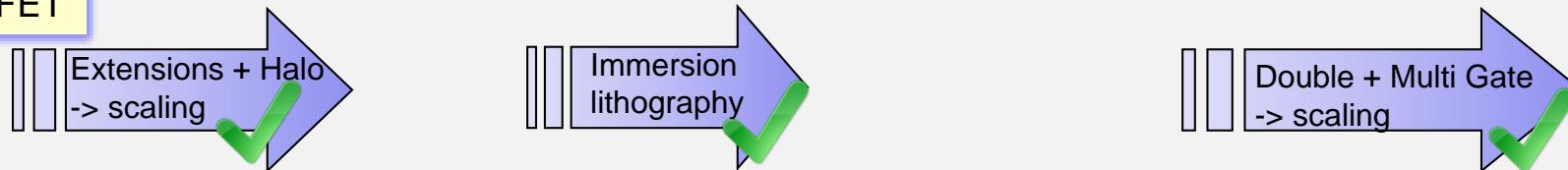
## 6.7 SOI-MOSFETs

## 6.8 Multi-Gate MOSFETs

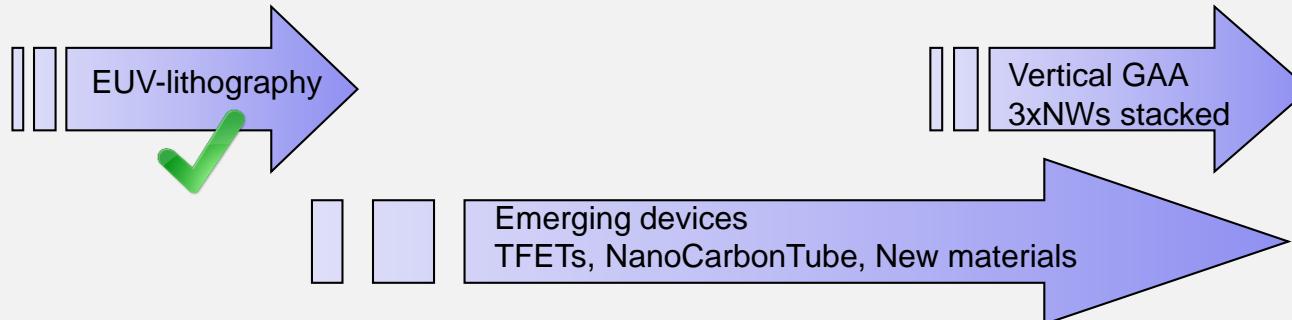
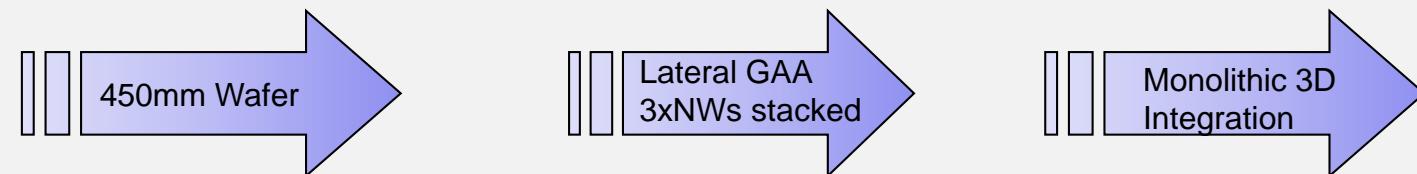
Fully-depleted MOSFETs, FinFETs, Nanowire FETs

## 6.9 Advanced Memory Devices

## Planar MOSFET



## FinFET



From economic view the shrinking of devices reduces fabrication costs dramatically (see chapter 2, economics).

(if shrinking can be done using the same equipment.

If new equipment is necessary after some shrinking the fabrication costs rise)

In addition, shrinked devices offer the advantage of:

- \* better dynamic performance (faster, less power consumption)
- \* new applications -> new markets -> more need for chips -> more money

MOSFETs are best suited for shrinking, because:

MOSFETs are Field-Effect devices -> easy scaling

MOSFETs allow CMOS-circuits -> lowest power consumption up today -> best suited for ULSI and mobile

### Moore's Law:

Industry is able to shrink the dimensions of MOSFETs  
a factor  $S = \frac{1}{\sqrt{2}} = 0.7$  every 3 years (one generation)

Using this shrinking factor  $S = 0.7$  / generation a forecast  
can be done relating time and properties of MOSFETs.  
This is done in so-called roadmaps.

DRAM-complexity	historical										US National Roadmaps 1992, 1994, 1997							
	1k	4k	16k	64k	256k	1M	4M	16M	64M	256M	1G	1G *	4G	16G	64G	256G	Scaling S	
Year	1970/71	1973	1976/77	1979/80	1982	1985	1988	1993	1995	1997	1999	2001	2003	2006	2009	2012	All 3 years	
Process	p-chan Si-gate	n-chan Si-gate	n-chan 2-poly	n-chanl 2-poly														
DRAM cell structure	3 T	1 T	1 T	1T	1T	1T	1 T	1 T	1T	1T	1T	1T	1 T	1 T	1 T	1 T		
Min.Feature [µm] (channel length)	8 - 10	7 - 8	6-7 3-4	3.5 2-2.5	2.0 1.5	1.0 0.8	0.7 0.5	0.5 0.3	0.35	0.25	0.18	150	130	100	70	50	S=1/√2=0.7	
junction depth [nm]	2500	2000	1000	700	450	250	150	150		100-50	72-36	60-30	52-26	40-20	30-15	20-10		
gate oxide thick [nm]	120	120	100	70, 70	45	30-20	15-10	15-10		5-4	4-3	3-2	3-2	2-1.5	< 1.5	< 1.0		
gate width µm			5-7	3-4	2	1.2	0.5	0.5										
interconnect width µm			5	4	3	2	1	1										
interconnect thick µm			1	1	0.9	0.75	0.6	0.6										
cell area [µm²]	3700	900	450	170		30	10	3.9	1.5	0.56	0.22	0.14	0.09	0.036	0.014	0.006	S²/1.3	
Chip size [mm²]	13	19	18	21	38	55	80	145	190	280	400	445	560	790	1120	1580	1 / 1.5	
Udd [V]	18/12	12	12,5	5, 5	5	5	5-3.3	3.3		2.5- 1.8	1.8-1.5	1.5	1.2	1.2-0.9	0.9-0.6	0.6-0.5		
Power dissipation [W]					0.2					70	90	110	130	160	170	175		
comments			HMOS	HMOS	HMOS													

! All properties listed in the ITRS are only calculated with a working MOSFET,  
there exists no prediction how to achieve these values !

Table 14 Memory and Logic Technology Requirements

Year of First Product Shipment Technology Generation	1997 250 nm	1999 180 nm	2001 150 nm	2003 130 nm	2006 100 nm	2009 70 nm	2012 50 nm
Min. Logic V <sub>dd</sub> (V) (desktop)	2.5-1.8	1.8-1.5	1.5-1.2	1.5-1.2	1.2-0.9	0.9-0.6	0.6-0.5
V <sub>dd</sub> Variation	≤ 10%	≤ 10%	≤ 10%	≤ 10%	≤ 10%	≤ 10%	≤ 10%
T <sub>ox</sub> Equivalent (nm)	4-5	3-4	2-3	2-3	1.5-2	< 1.5	< 1.0
Equivalent Maximum E-field (MV/cm)	4-5	5	5	5	> 5	> 5	> 5
Max I <sub>off</sub> @ 25°C (nA/µm) (For minimum L device)	1	1	3	3	3	10	10
Nominal I <sub>on</sub> @ 25°C (µA/µm) (NMOS/PMOS)	600/280	600/280	600/280	600/280	600/280	600/280	600/280
Gate Delav Metric (CV/I) (ps)*	16-17	12-13	10-12	9-10	7	4-5	3-4

As an indicator, whether the technology exists  
to fabricate future devices, colors are used:



no problem

existing technology, but  
development neccessary

no imagination, how to do

## 6.1 Overview

Have a look at:



1997-2016

International Technology Roadmap for Semiconductors



Since 2017

International Roadmap for Devices and Systems

<http://www.itrs.net/>

<http://www.itrs2.net/>

<https://irds.ieee.org/>

INTERNATIONAL ROADMAP FOR DEVICES AND SYSTEMS™

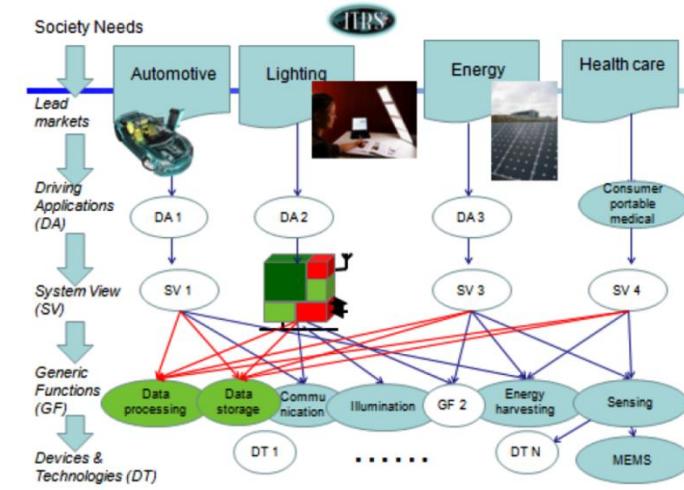
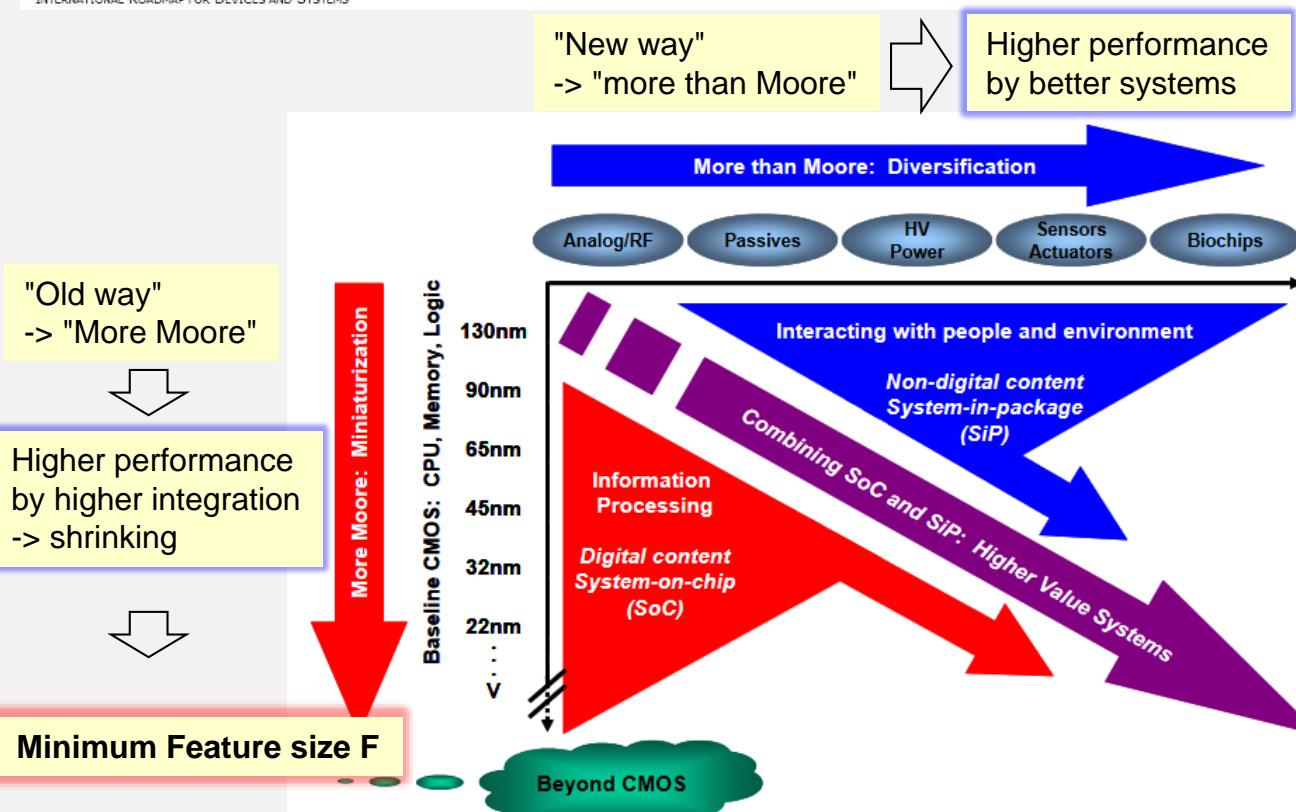
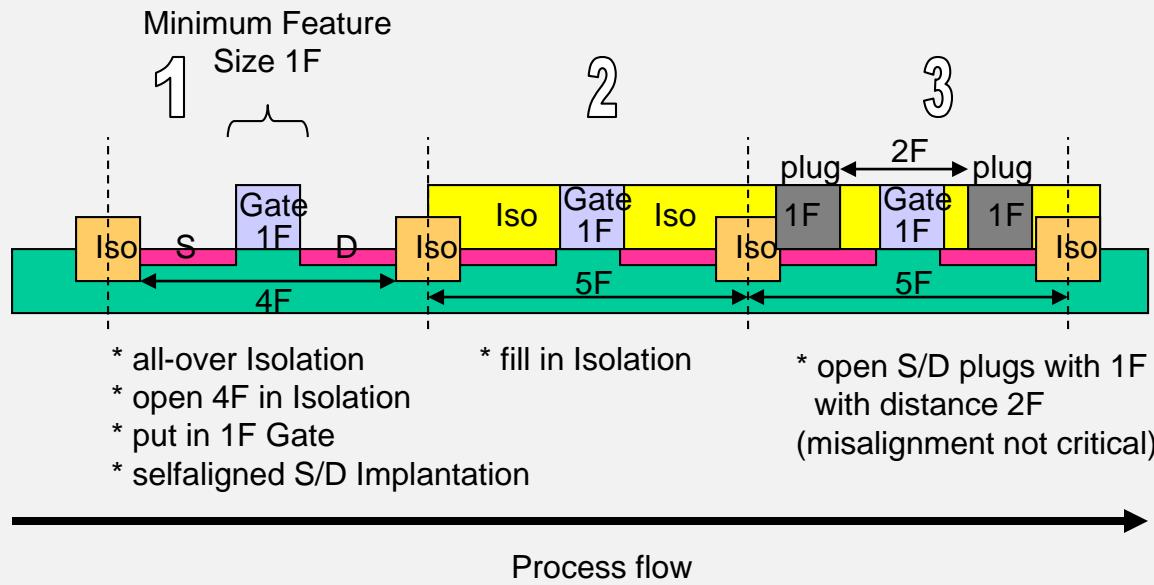


Figure 3 Proposed Roadmapping Process for More than Moore Technologies

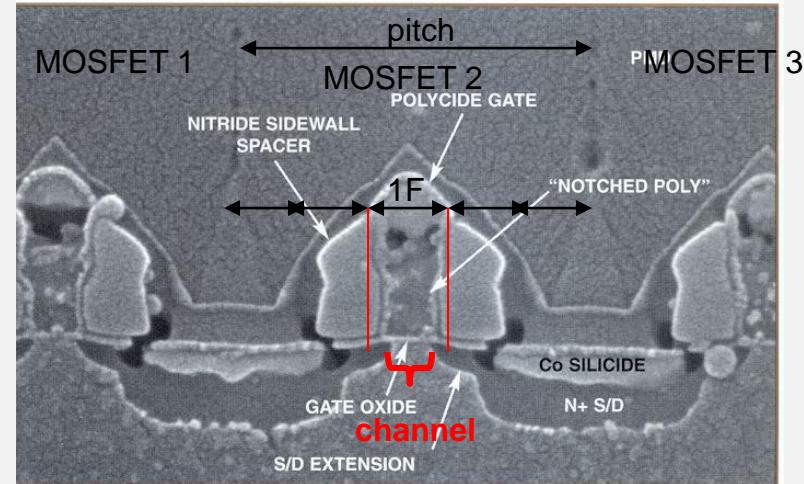
- 1) Generations were characterized by the so-called "**minimum feature size F**", which can be fabricated.  
According to the fabrication process of MOSFETs, this minimum feature size was identically with the channel length.



In the first International Technology Roadmap (2001) two things were recognized:

- 2) Some structures cannot be scaled with F to ensure proper working,  
e.g. shallow Source/Drain with 1F<sup>2</sup> will result in too high series resistance

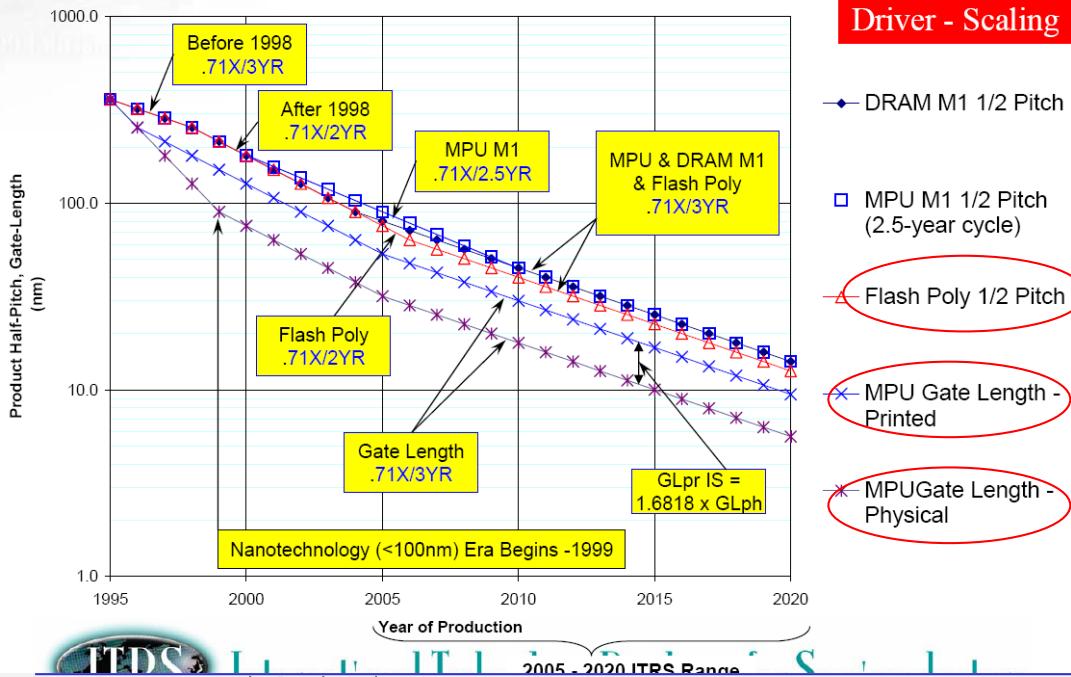
- 3) On chips not single devices are fabricated, but functional structures (DRAM, inverter).  
Therefore, to describe the number of functional blocks, not the channel length is important, but the density of blocks.  
This is described by the "**pitch**" (distance between repeating structures).



- \* MOSFETs in INTEL-PIII MPU (IEDM2000)  
(test-pattern of MOSFETs, not real circuit-design)
- \* here: pitch 5F, but missing isolation -> 6F
- \* channel length < 1F, but 1F is pattern size (litho)

## 2005 Definition of the Half Pitch

2005 ITRS Product Technology Trends -  
Half-Pitch, Gate-Length

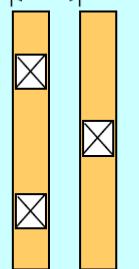


### Semiconductor Roadmap Driver - Scaling

- DRAM M1 1/2 Pitch
- MPU M1 1/2 Pitch (2.5-year cycle)
- Flash Poly 1/2 Pitch
- MPU Gate Length - Printed
- MPUGate Length - Physical

DRAM ½ Pitch  
= DRAM Metal Pitch/2  
MPU/ASIC M1 ½ Pitch  
= MPU/ASIC M1 Pitch/2

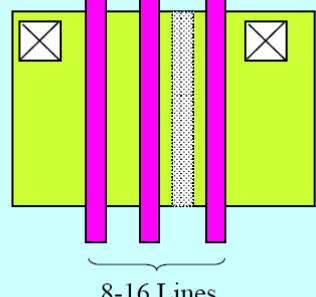
Metal  
Pitch



Typical DRAM/MPU/ASIC  
Metal Bit Line

FLASH Poly Silicon ½ Pitch  
= Flash Poly Pitch/2

Poly  
Pitch



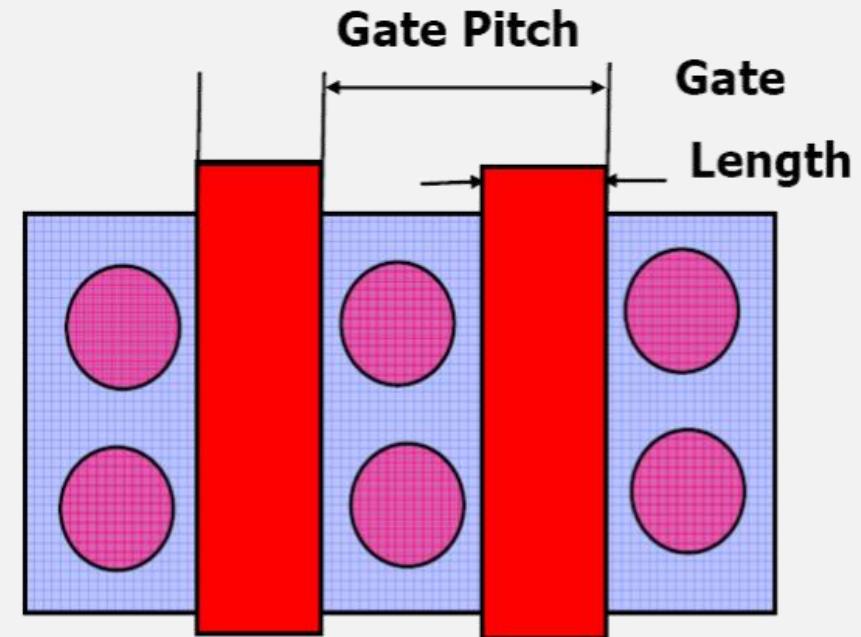
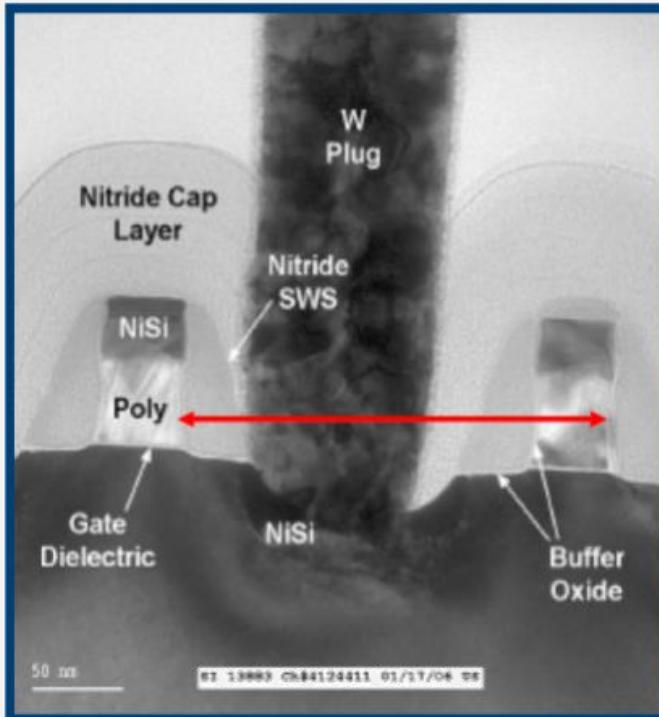
Typical flash  
Un-contacted Poly



Resist trimming

- printed gate length
- physical gate length
- electrical channel length

## MOSFET scaling



Node	Device Pitch (nm)
45	170-180
32	120-130
22	90-100
15	70-80
11	55-65

Starting in 2002 the terminus "**node**" was used to characterize the smallest "**half-pitch**" in devices (mainly the DRAM Gate half-pitch and Metal-1 in MPUs)

Since 2012 there is not a consensus on the "**node**" naming across different foundries and integrated device manufacturers (IDMs).  
IRDS 2018

# Advanced MOSFET Development

## - Minimum Feature Size, Pitch and Node

YEAR OF PRODUCTION	2018	2020	2022	2025	2028	2031	2034
G54M36	G48M30	G45M24	G42M21	G40M16	G40M16T2	C40M16T4	
<i>Logic Industry "Node Range" Labeling (nm)</i>	"7"	"5"	"3"	"2.1"	"1.5"	"1.0 eq"	"0.7 eq"
<i>IDM-Foundry node labeling</i>	i10-f7	i7-f5	i5-f3	i3-f2.1	i2.1-f1.6	i1.6-f0.7e	i1.0e-f0.7e
Logic device structure options	FinFET	finFET	LGAA	LGAA	LGAA-3D	LGAA-3D	VGAA
Mainstream device for logic	finFET	finFET	finFET	LGAA	LGAA	LGAA-3D	LGAA-3D
<b>LOGIC DEVICE GROUND RULES</b>							
Mx pitch (nm)	40	36	32	24	20	16	16
Mx pitch (nm)	36	32	30	21	20	20	20
<i>Mx pitch (nm)</i>	36	30	24	21	16	16	16
Gate pitch (nm)	54	48	45	42	40	40	40
L <sub>x</sub> Gate Length - HP (nm)	20	18	15	14	12	12	12
L <sub>x</sub> Gate Length - HD (nm)	22	20	18	14	12	12	12
Channel aspect ratio - line sides	0.20	0.20	0.20	0.20	0.20	0.20	0.20
Spacer width (nm)	8	7	6	6	6	6	6
Contact CD (nm) - finFET, LGAA	18	16	17	16	16	16	16
Contact C.D. (nm) - VGAA							
Device architecture key ground rules							
FinFET pitch (nm)	32.0	28.0	24.0				
FinFET Fin width (nm)	8.0	7.0	6.0				
FinFET Fin height (nm)	40	50	60				
Footprint drive efficiency - finFET	2.75	3.02	5.25				
Lateral GAA lateral pitch (nm)				22.0	20.0	20.0	20.0
Lateral GAA vertical pitch (nm)				18.0	16.0	14.0	14.0
Lateral GAA (nanosheet) thickness (nm)				7.0	6.0	5.0	5.0
Number of vertically stacked nanosheets				3	3	4	4
LGAA width (nm) - HP				25	20	15	10
LGAA width (nm) - HD				15	11	6	6
LGAA width (nm) - SRAM				7	6	6	6
LGAA total height (nm)				53	48	57	57
Footprint drive efficiency - lateral GAA - HP				4.80	4.59	5.52	5.00
Device effective width (nm) - HP	88.0	107.0	125.0	192.0	156.0	160.0	120.0
Device effective width (nm) - HD	66.0	107.0	125.0	132.0	102.0	88.0	88.0
Device lateral pitch (nm)	32	28	24	22	20	20	20
Device height (nm)	40.0	50.0	60.0	53.0	48.0	57.0	57.0
Device width (nm) - HP	8	7	6	25	20	15	10
Device width (nm) - HD	8	7	6	15	11	6	6
Device width (nm) - SRAM	8	7	6	7	6	6	6

Note the different dimensions for node, pitch, channel length

<https://irds.ieee.org/>



INTERNATIONAL ROADMAP FOR DEVICES AND SYSTEMS™

**INTERNATIONAL  
ROADMAP  
FOR  
DEVICES AND SYSTEMS™**

2018 UPDATE

YEAR OF PRODUCTION	2018	2020	2022	2025	2028	2031	2034
G54M36	G54M36	G48M30	G45M24	G42M21	G40M16	G40M16T2	G40M16T4
Logic industry "Node Range" Labeling (nm)	"7"	"5"	"3"	"2.1"	"1.5"	"1.0 eq"	"0.7 eq"
IDM-Foundry node labeling	i10-f7	i7-f5	i5-f3	i3-f2.1	i2.1-f1.5	i1.5e-f1.0e	i1.0e-f0.7e
Logic device structure options	FinFET	finFET	finFET LGAA	LGAA	LGAA VGAA	LGAA-3D VGAA	LGAA-3D VGAA
Mainstream device for logic	finFET	finFET	finFET	LGAA	LGAA	LGAA-3D	LGAA-3D
DEVICE ELECTRICAL SPECS							
Mx pitch (nm)	40	36	32	24	20	16	16
M1 pitch (nm)	36	32	30	21	20	20	20
M0 pitch (nm)	36	30	24	21	16	16	16
Gate pitch (nm)	54	48	45	42	40	40	40
L <sub>o</sub> : Gate Length - HP (nm)	20	18	16	14	12	12	12
L <sub>g</sub> : Gate Length - HD (nm)	22	20	18	14	12	12	12
Power Supply Voltage - Vdd (V)	0.75	0.70	0.70	0.65	0.65	0.60	0.55
Subthreshold slope (mV/dec) - HP (mV/dec)	68	78	82	80	75	75	70
Subthreshold slope (mV/dec) - HD (mV/dec)	68	72	75	75	68	65	65
Inversion layer thickness (nm) [1]	1.10	1.10	1.00	1.00	0.90	0.90	0.90
V <sub>t,sat</sub> at I <sub>off</sub> =10nA/um - HP (mV)	190	222	237	236	226	226	211
V <sub>t,sat</sub> (mV) at I <sub>off</sub> =100pA/um - HD (mV) [2]/[3]	323	345	363	371	341	326	326
Effective mobility (cm <sup>2</sup> V.s)	150	125	125	100	100	100	100
R <sub>sd</sub> (Ohms.um) [4]	300	285	271	257	244	232	221
Ballisticity/Injection velocity (cm/s)	1.32E-07	1.39E-07	1.46E-07	1.46E-07	1.46E-07	1.46E-07	1.46E-07
V <sub>dsat</sub> (V) - HP	0.141	0.160	0.149	0.163	0.140	0.140	0.140
V <sub>dsat</sub> (V) - HD	0.155	0.177	0.168	0.163	0.140	0.140	0.140
I <sub>on</sub> (uA/um) at I <sub>off</sub> =10nA/um - HP when R <sub>sd</sub> =0 [5]	2027	1733	1950	1670	1975	1696	1501
I <sub>on</sub> (uA/um) at I <sub>off</sub> =10nA/um - HP [6]	972	853	911	820	923	826	759
I <sub>on</sub> (uA/device) at I <sub>off</sub> =100nA/um - HP [7]	86	91	157	157	144	132	91
I <sub>on</sub> (uA/um) at I <sub>off</sub> =100pA/um - HD when R <sub>sd</sub> =0 [5]	1447	1157	1271	991	1334	1139	859
I <sub>on</sub> (uA/um) at I <sub>off</sub> =100pA/um - HD [6]	596	483	495	390	520	467	347
I <sub>on</sub> (uA/device) at I <sub>off</sub> =100pA/um - HD [7]	52	52	75	75	81	75	42
C <sub>ch,total</sub> (fF/um <sup>2</sup> ) - HP/HD [8]	31.38	31.38	34.52	34.52	38.35	38.35	38.35
Gate height over fin (nm)	30	25	20	15	15	15	15
Spacer k value	4.5	4.0	3.5	3.5	3.0	3.0	3.0
C <sub>ch</sub> (fF/um) - HP [8]	0.50	0.45	0.44	0.39	0.37	0.37	0.37
C <sub>ch</sub> (fF/um) - HD [8]	0.55	0.50	0.50	0.39	0.37	0.37	0.37
CV/I (ps) - FO3 load, HP [9]	1.16	1.11	1.02	0.92	0.78	0.80	0.80
I/(CV) (1/ps) - FO3 load, HP [10]	---	---	---	---	1.00	1.25	1.25
Energy per switching [CV] (fJ/switch) - FO3 /						0.40	0.33

Solutions Exist

Solutions Being Pursued

No Known Solution

Year 2028 -> 1.5nm Generation

Gate pitch: saturation at 40 nm

L: saturation at 12 nm

Supply Voltage: 0.65 V

S: 68 mV/dec

Parts of current are not longer drift-diffusion but other transport effects, like tunneling, hot electrons, avalanche.

Roadmaps are only a list of required device properties in time to ensure the prosperity of semiconductor industry

-> Roadmaps are based on planar technology and shrinking  
-> for a lot of requirements today no realization possibility (red walls) exist

History (1960- 2000): multi-purpose MOSFETs were designed according the scaling laws

Trend (since ~2000): scaling limits for several structures reached (e.g.: gate oxide  $\rightarrow$  tunneling  $\rightarrow$  no high-k available)  
 $\rightarrow$  depending due to the application different performance is needed

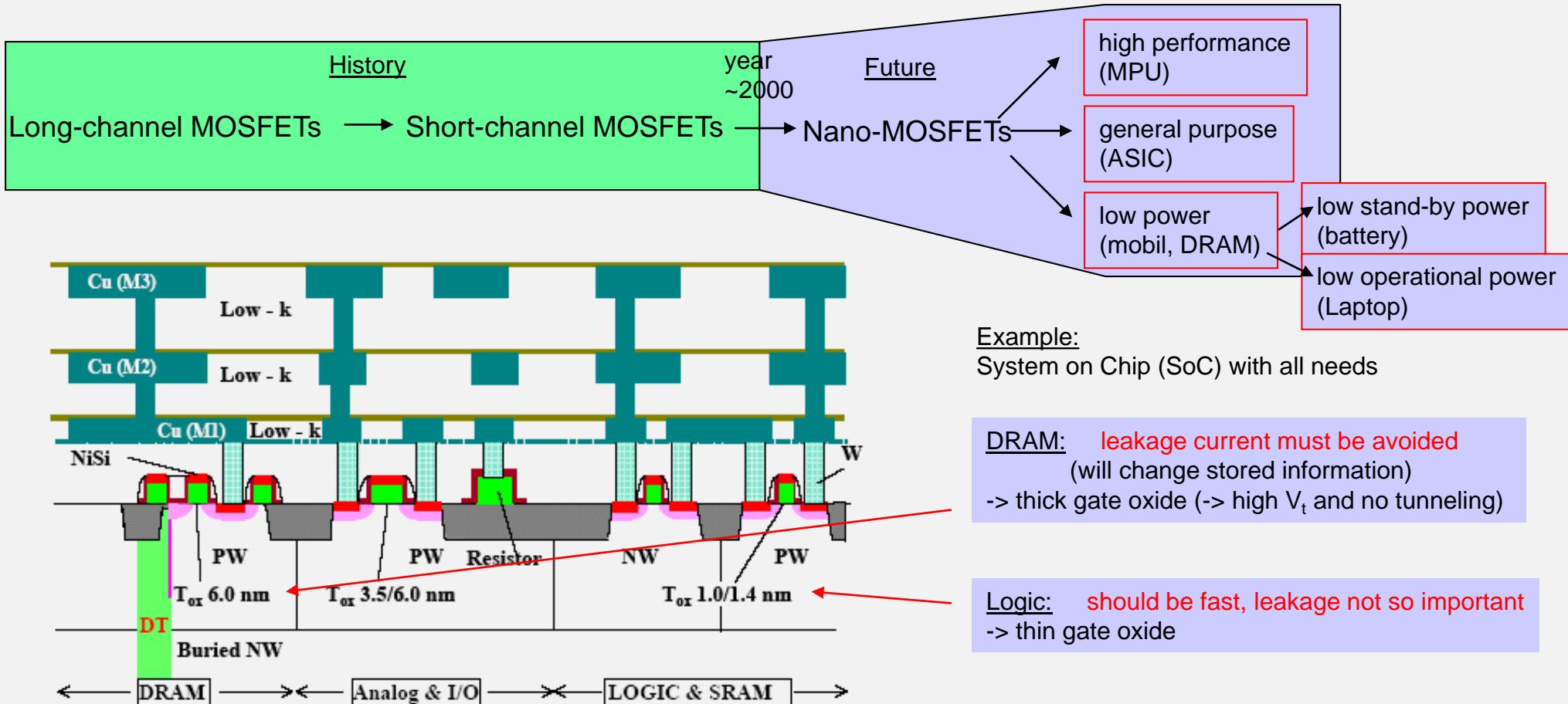


Fig.2 Cross-section of proposed 65nm SoC. Source: Toshiba, IEDM 2002

## 6.1 Overview

Feature size, pitch, node

## 6.2 State of the Art MOSFETs

Electrical and geometrical parameters

## 6.3 Mobility Enhanced MOSFETs

Strain, SiGe, Liners, Directed Channels, Gate-Last

## 6.4 High-k Metal-Gate MOSFETs

Poly-Depletion, FuSi

## 6.5 Vertical MOSFETs

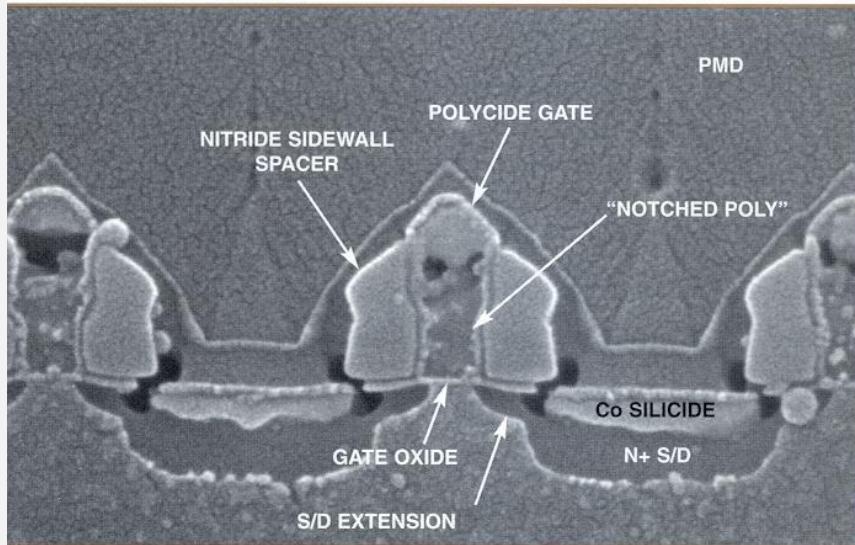
## 6.6 Halo MOSFETs

## 6.7 SOI-MOSFETs

## 6.8 Multi-Gate MOSFETs

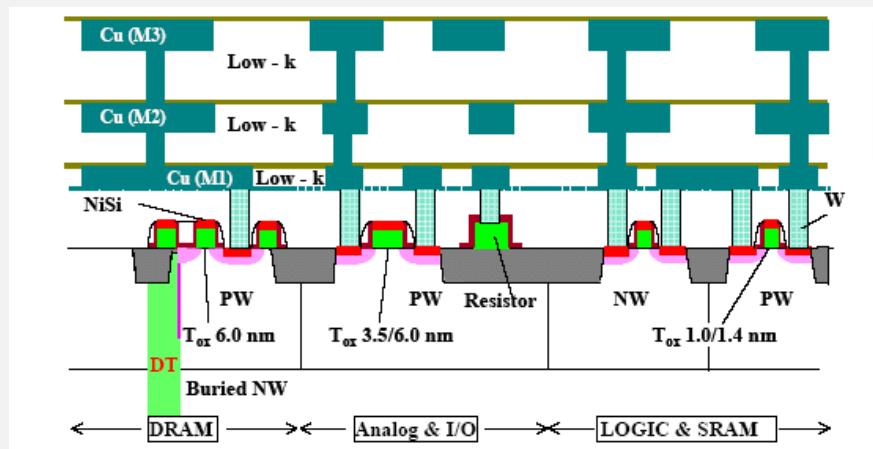
Fully-depleted MOSFETs, FinFETs, Nanowire FETs

## 6.9 Advanced Memory Devices



- $L_G = 70 \text{ nm}$  defined by lithography (printed)  
(KrF 248 nm, ArF 193 nm with phase shift masks)
- SDEs self-aligned to gate
- High-quality gate oxide (1.7-1.5 nm)
- Small parasitic capacitances

Typical values from INTEL, FUJITSU, Toshiba, ... :

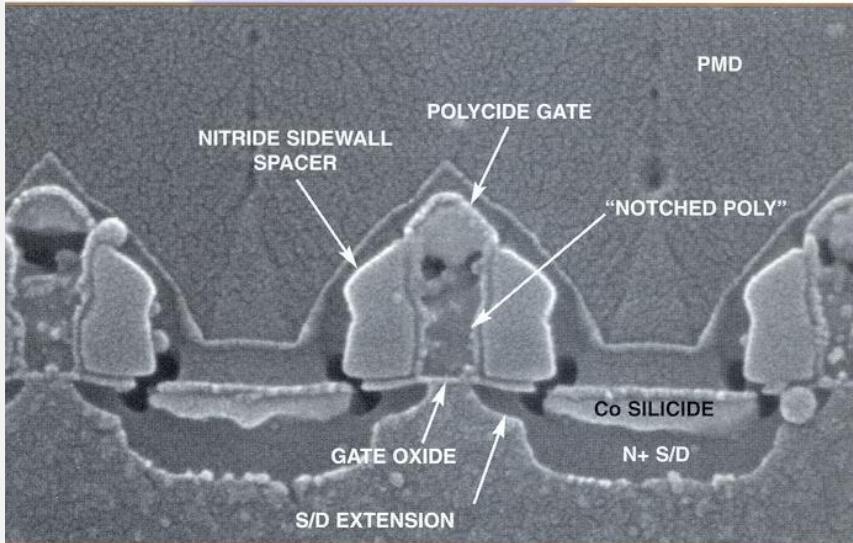


Usually **dual  $V_T$**  (by 2 oxide thicknesses) is used:  
 - high  $V_T$  (thick gate oxide) -> low leakage, slowly -> low power  
 - low  $V_T$  (thin gate oxide)-> high speed, high leakage -> high performance

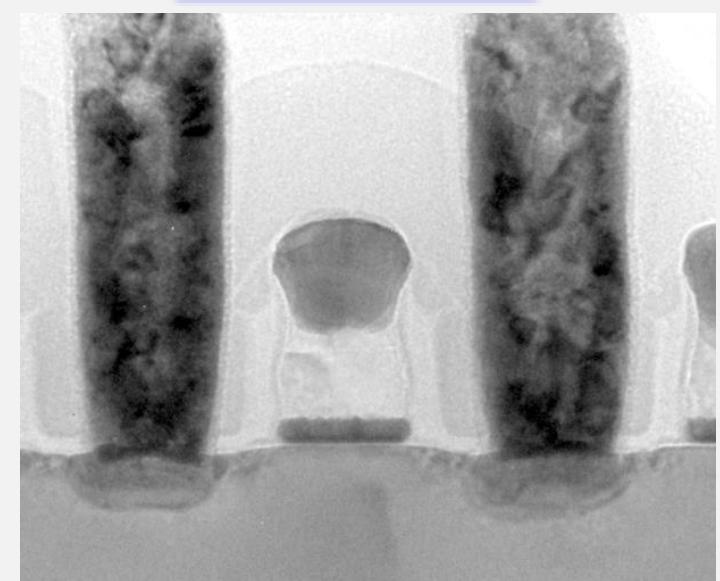
Source: Toshiba, IEDM 2002

Typical performance values from INTEL, IBM, FUJITSU, Motorola, Hitachi:

2000 Node: 130 nm



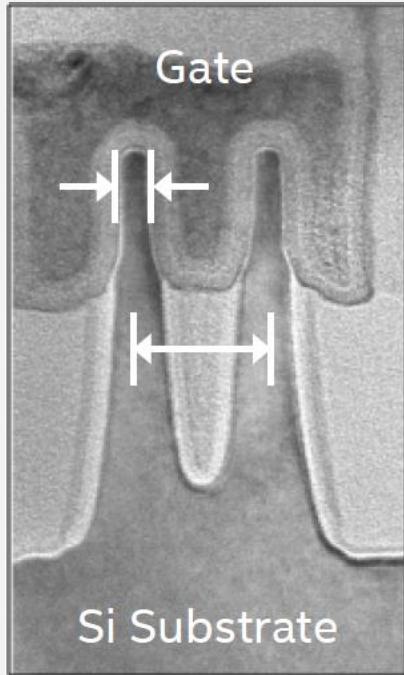
2009 Node: 32 nm



- $L_G = 70$  nm defined by lithography (printed)  
(KrF 248 nm, ArF 193 nm with phase shift masks)
- SDEs self-aligned to gate
- High-quality gate oxide (1.7-1.5 nm)
- Small parasitic capacitances

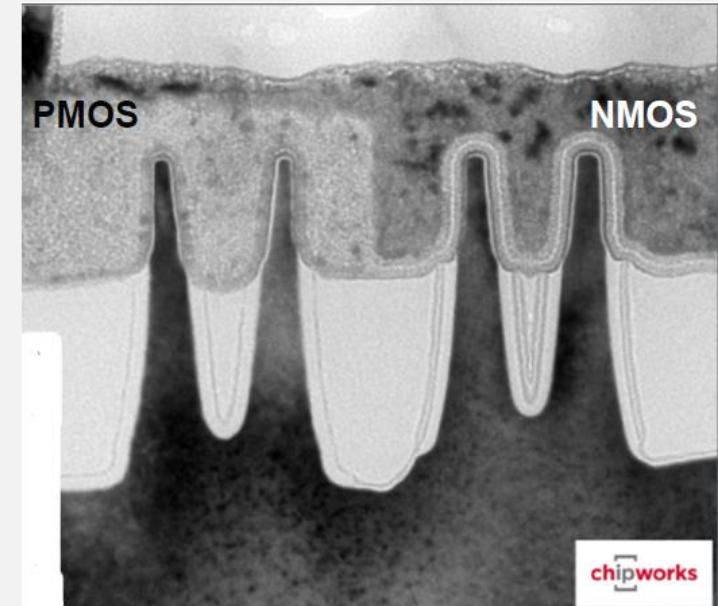
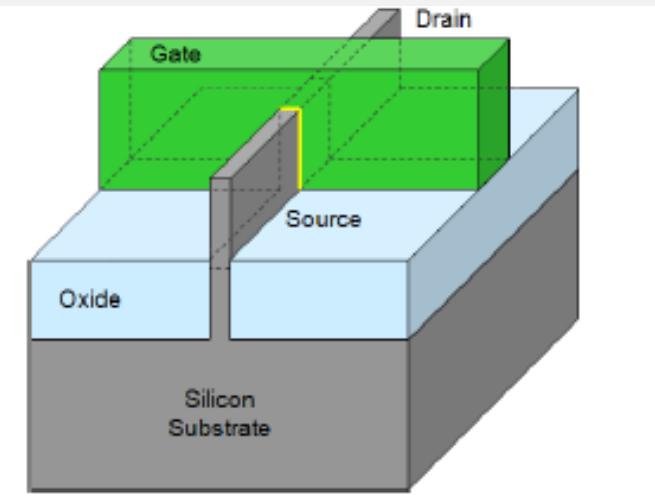
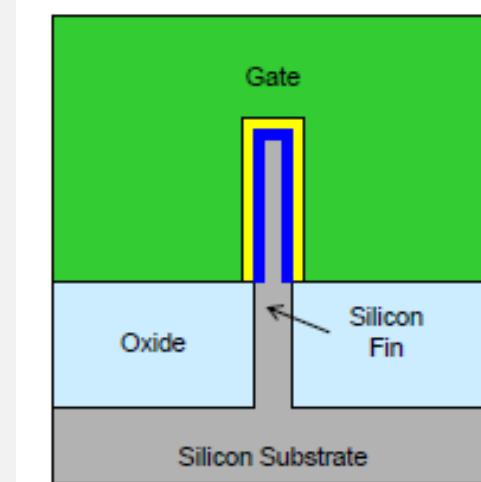
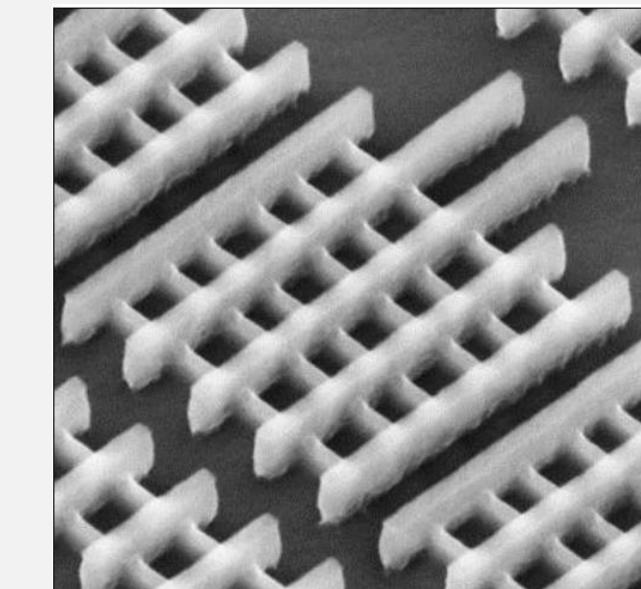
- $L_G = 32$  nm defined by lithography (printed)  
(ArF 193 nm immersion with phase shift masks)
- SDEs self-aligned to gate
- High-k (EOT = 0.9 nm) + metal gate

2014 Node (FinFET): 14 nm



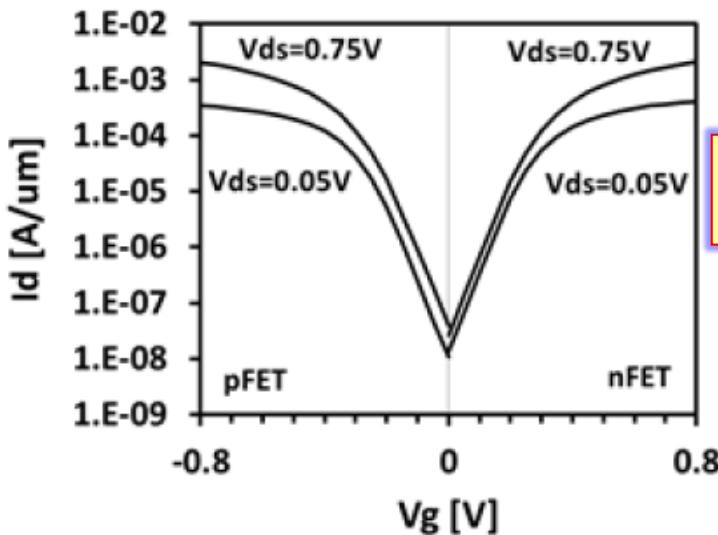
INTEL, IEDM 2014

- Fin width 8 nm
- Fin pitch 42 nm
- Replacement Gate
- High-k + Metal Gate



2016 Node (FinFET): 10 nm

Samsung 2016



$S: \sim 65\text{mV/dec}$   
 $DIBL: \sim 30\text{mV/V}$

- Fin width: ?
- Fin pitch: 42 nm
- Replacement Gate
- High-k + Metal Gate
- Triple patterning

04.01.2017: Snapdragon 835 (Qualcomm) is the first chipset to be made on the 10 nm FinFET process node of Samsung.

SAMSUNG  
NEWSROOM

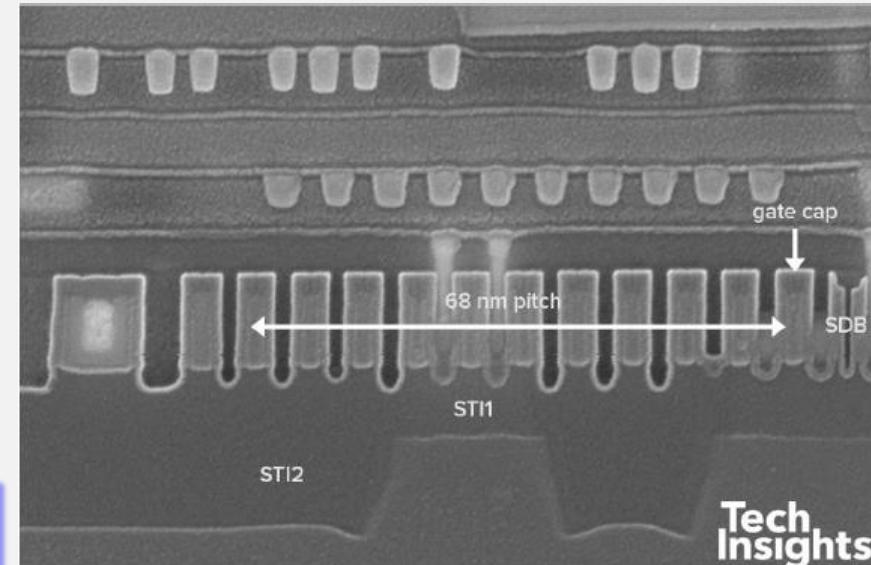
CORPORATE PRODUCTS P

Press Resources > Press Release

## Samsung Starts Industry's First Mass Production of System-on-Chip with 10-Nanometer FinFET Technology

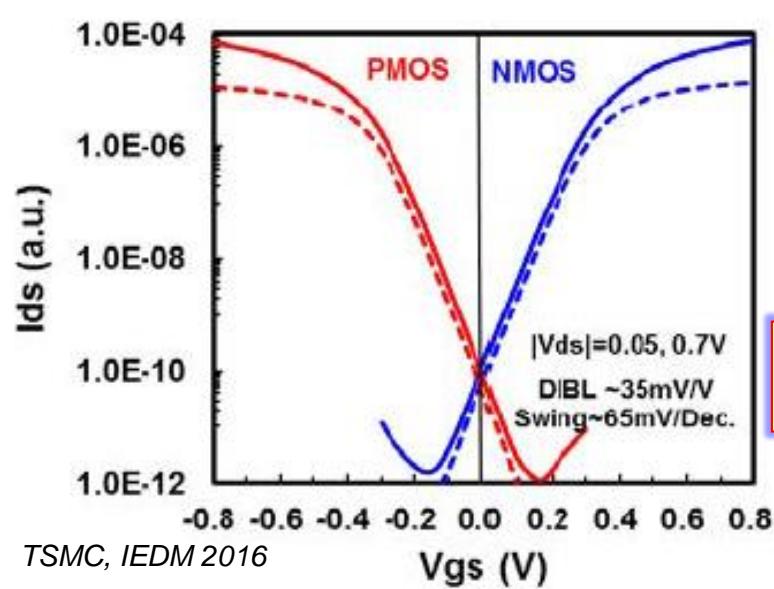
Korea on October 17, 2016

<https://news.samsung.com/global/samsung-starts-industrys-first-mass-production-of-system-on-chip-with-10-nanometer-finfil-technology>

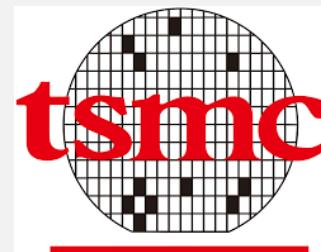


Tech  
Insights

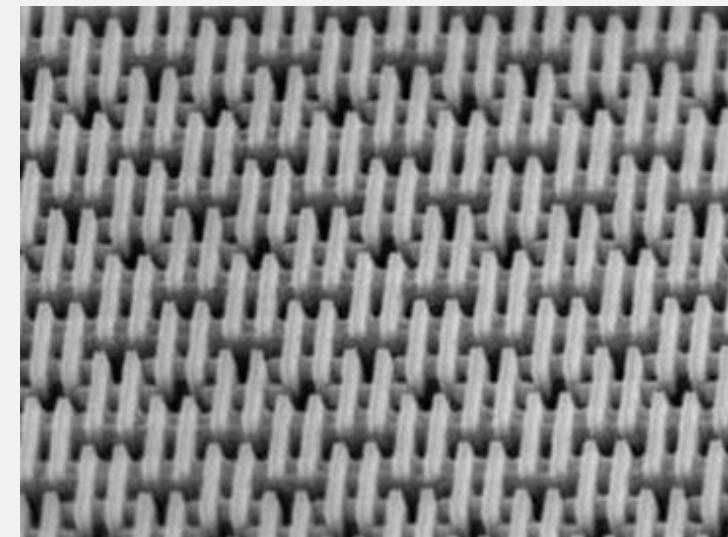
2018 Node (FinFET): 7 nm



TSMC, IEDM 2016



$S: \sim 65mV/dec$   
 $DIBL: \sim 35mV/V$



<https://www.tsmc.com/english/dedicatedFoundry/technology/7nm.htm>

DESIGNLINES | WIRELESS AND NETWORKING DESIGNLINE

## Apple Describes 7nm iPhone SoC

New handsets expand to 512 GB and \$1,099

By Rick Merritt, 09.12.18 □ 6

[https://www.eetimes.com/document.asp?doc\\_id=1333705](https://www.eetimes.com/document.asp?doc_id=1333705)

Share Post Share on Facebook Share on Twitter G+ in

SAN JOSE, Calif. — Apple announced a family of three iPhones powered by a 7nm SoC enabling up to 512 GBytes of memory. The handsets range in starting prices from \$749 to \$1,099, increase battery life by 30 to 90 minutes, and ship within two to six weeks.

The smartphones, and two new models of the Apple Watch, generally packed larger screens and upgraded chips, in some cases in slightly smaller devices. None of the devices support 5G cellular networks, expected to start switching on later this year, but the handsets support Gbit/s data rates, an LTE capability that Qualcomm was early to support.

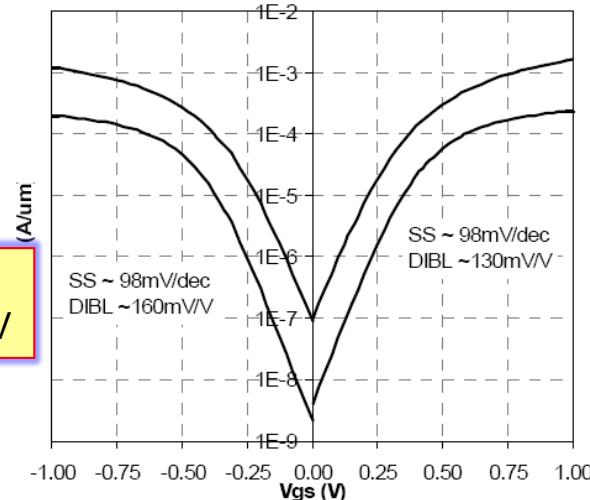
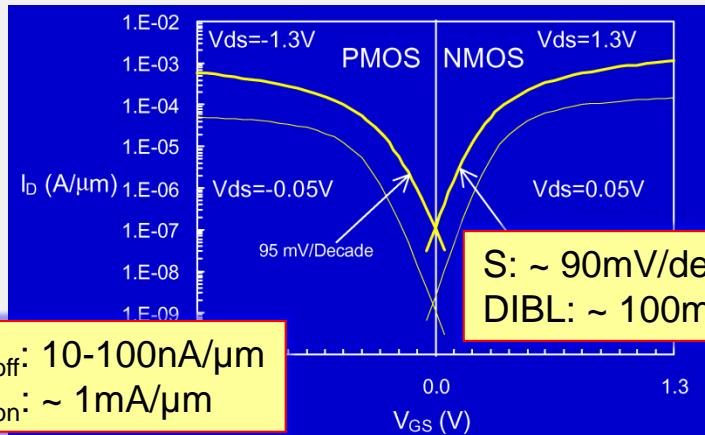
Apple's 7nm A12 Bionic chip packs 6.9 billion transistors and is "the most powerful chip in a smartphone," said chief executive Tim Cook.

Typical performance values from INTEL, FUJITSU, Motorola, Hitachi:

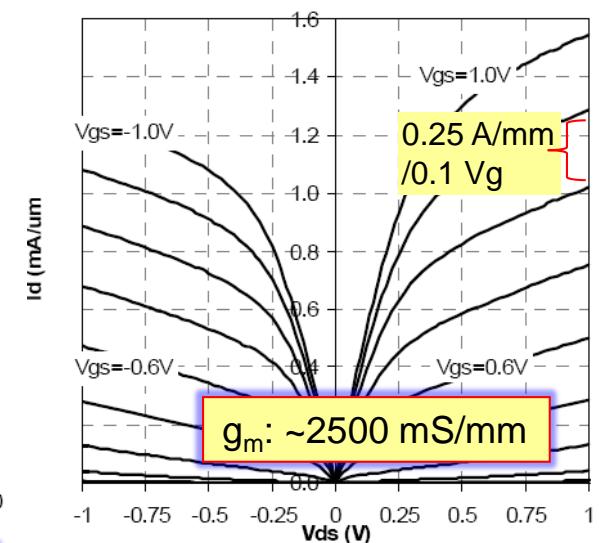
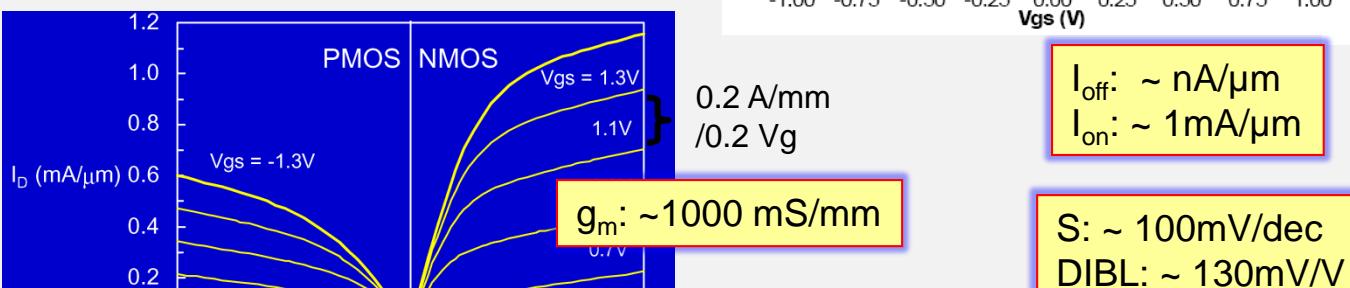
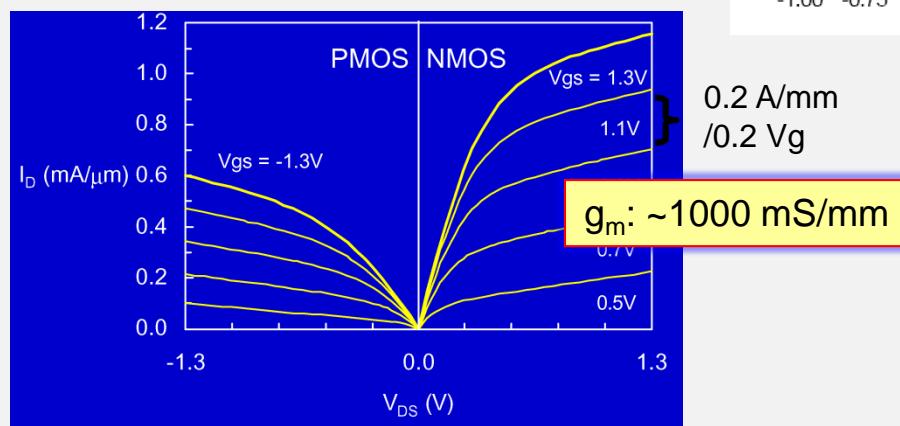
2009 Node: 32 nm

2000 Node: 130 nm

$L_{gate}$ : ~ 70 nm,  $t_{ox}$ : ~ 1.5 nm  
 $V_{DD}$ : ~ 1.2 V,  $V_t$ : ~ 0.5 - 0.4 V



L: ~ 32 nm, EOT: ~ 0.9 nm high-k HfSiON  
 $V_{DD}$ : ~ 1.0 V,  $V_t$ : ~ 0.4 V

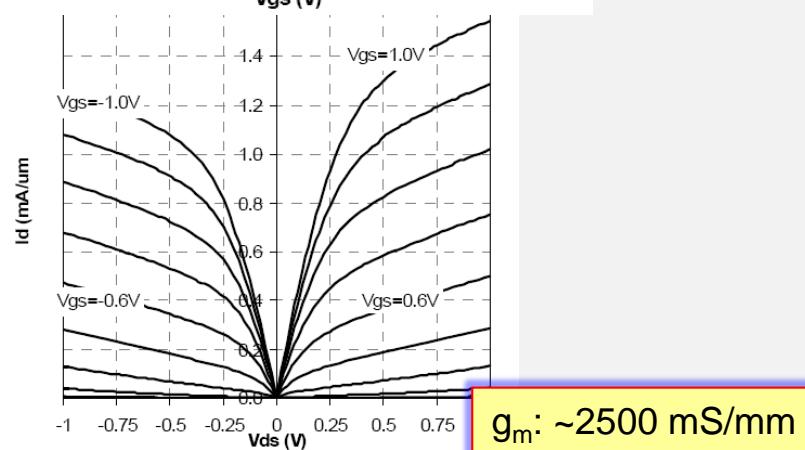
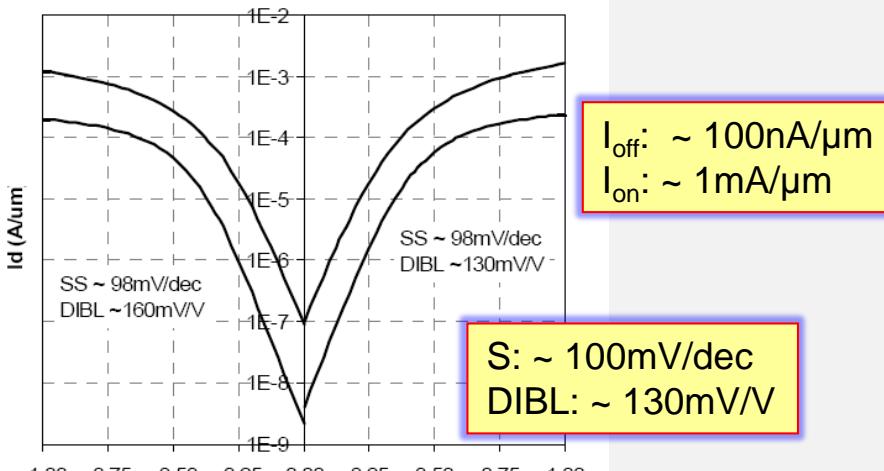


INTEL, IEDM 2008

2009 Node: 32 nm

INTEL, IEDM 2008

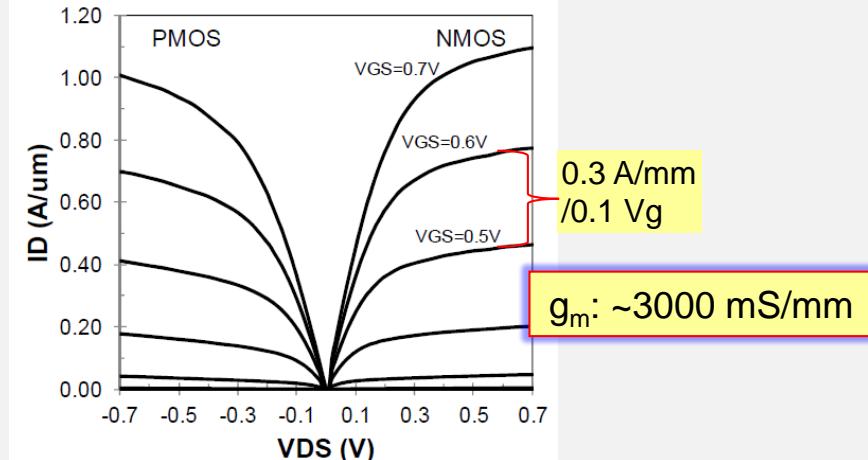
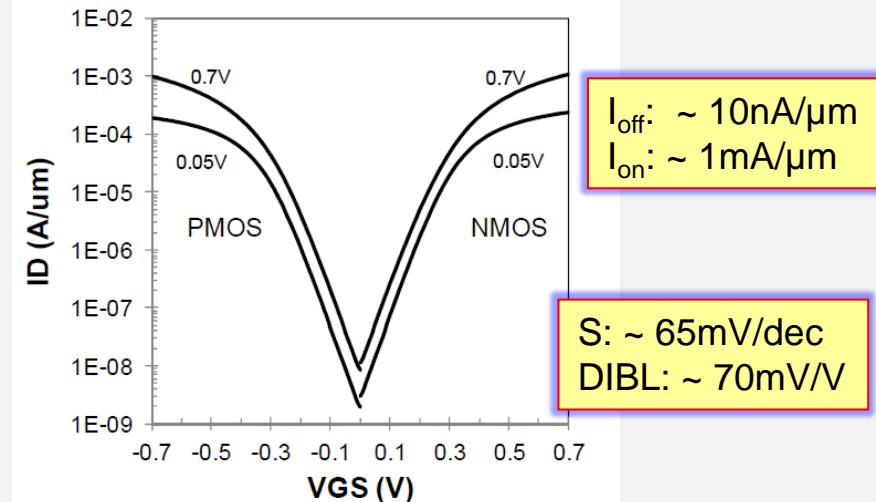
L: ~ 32 nm, EOT: ~ 0.9 nm high-k HfSiON  
V<sub>DD</sub>: ~ 1.0 V, V<sub>t</sub>: ~ 0.4 V



2014 Node (FinFET): 14 nm

INTEL, IEDM 2014

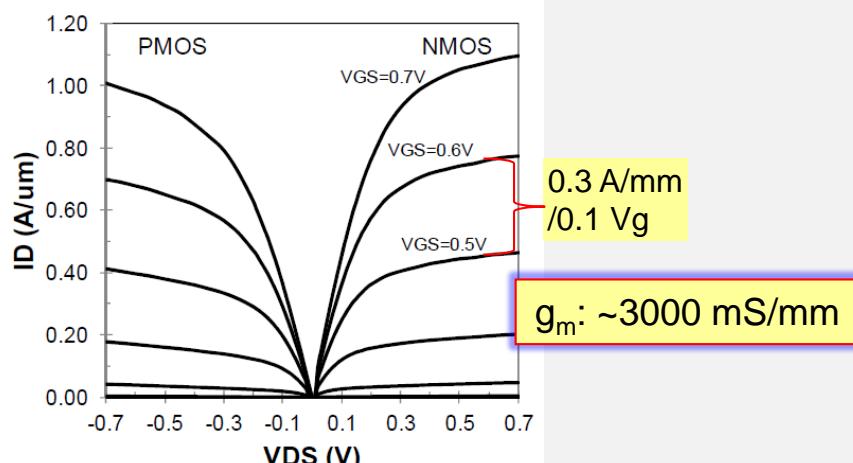
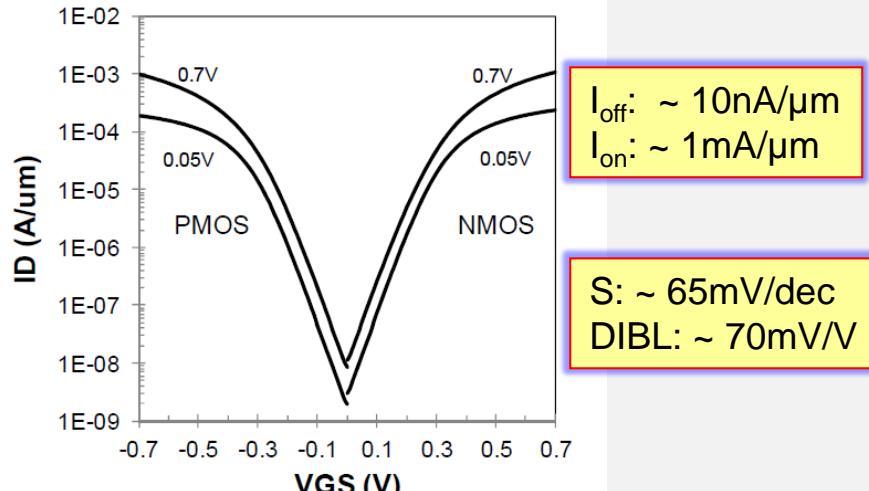
L: ~ 20 nm, high-k HfSiON  
V<sub>DD</sub>: ~ 0.7 V, V<sub>t</sub>: ~ 0.3 V



## 2014 Node (FinFET): 14 nm

INTEL, IEDM 2014

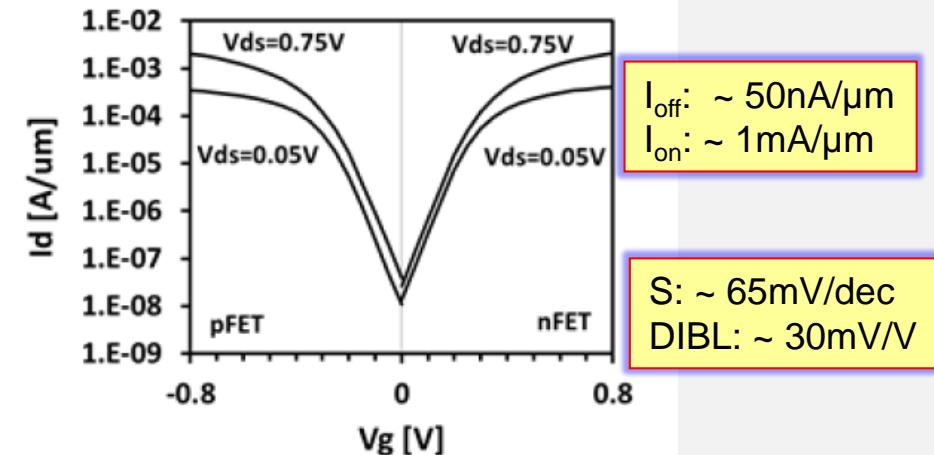
L: ~ 20 nm, high-k HfSiON  
 $V_{DD}$ : ~ 0.7 V,  $V_t$ : ~ 0.3 V



## 2016 Node (FinFET): 10 nm

Samsung 2016

L: ~ nm, high-k HfSiON  
 $V_{DD}$ : ~ 0.75 V,  $V_t$ : ~ 0.3 V



## Minimum Feature Size

	Intel 22 nm	Intel 14 nm	TSMC 16 nm	Samsung 14 nm	Samsung 10 nm	Intel 10 nm
Transistor Fin Pitch	60 nm	42 nm	48 nm	48 nm	42 nm	34 nm
Transistor Gate Pitch	90 nm	70 nm	90 nm	84 nm	64 nm	54 nm
Interconnect Pitch	80 nm	52 nm	64 nm	64 nm	48 nm	36 nm
SRAM Cell Area	.1080 $\mu$ m <sup>2</sup>	.0588 $\mu$ m <sup>2</sup>	.0700 $\mu$ m <sup>2</sup>	.0645 $\mu$ m <sup>2</sup>	0.04 $\mu$ m <sup>2</sup>	0.03 $\mu$ m <sup>2</sup>

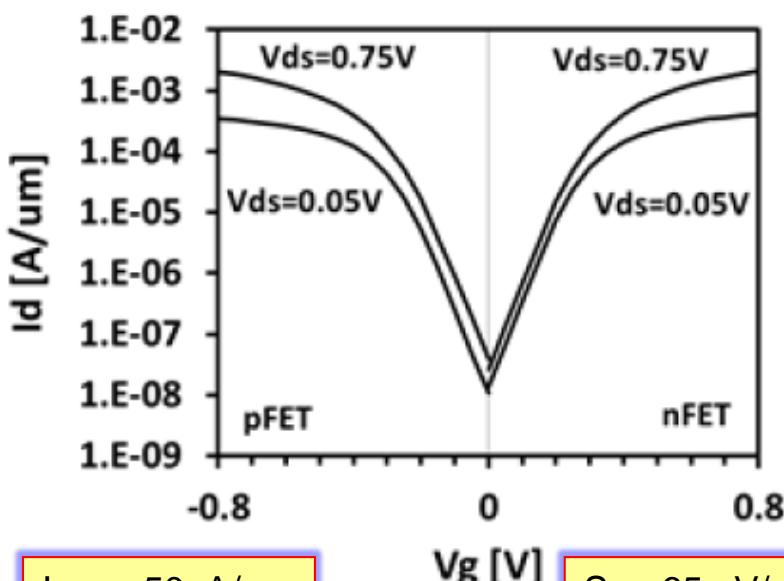
<https://www.extremetech.com/computing/221532-tsmc-will-begin-10nm-production-this-year-claims-5nm-by-2020>

S.-Y. Wu  
2014 IEDM  
T. Song  
2014 ISSCC

2016 Node (FinFET): 10 nm

Samsung 2016

L: ~ 10 nm, high-k HfSiON  
 $V_{DD}$ : ~ 0.75 V,  $V_t$ : ~ 0.3 V



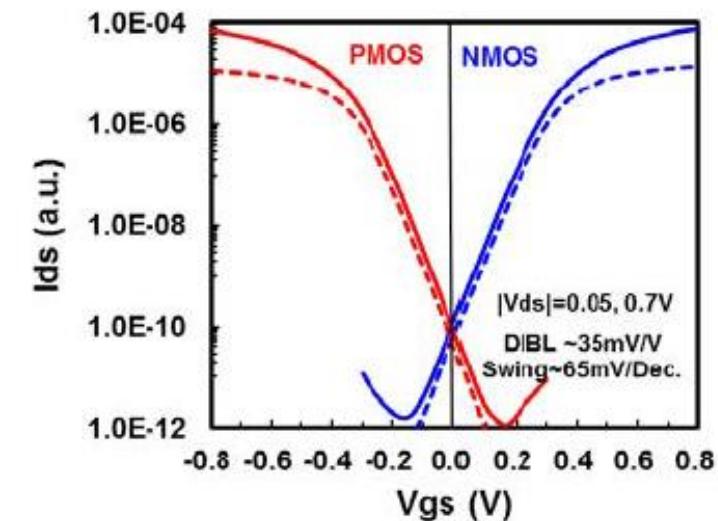
$I_{off}$ : ~ 50nA/μm  
 $I_{on}$ : ~ 1mA/μm

S: ~ 65mV/dec  
DIBL: ~ 30mV/V

2018 Node (FinFET): 7 nm

TSMC, IEDM 2016

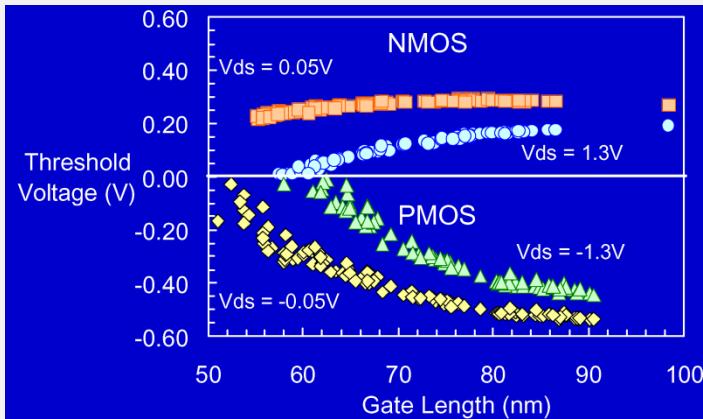
L: ~ 15 nm, high-k HfSiON  
 $V_{DD}$ : ~ 0.7 V,  $V_t$ : ~ 0.3 V



$I_{off}$ : ~ ?  
 $I_{on}$ : ~ ?

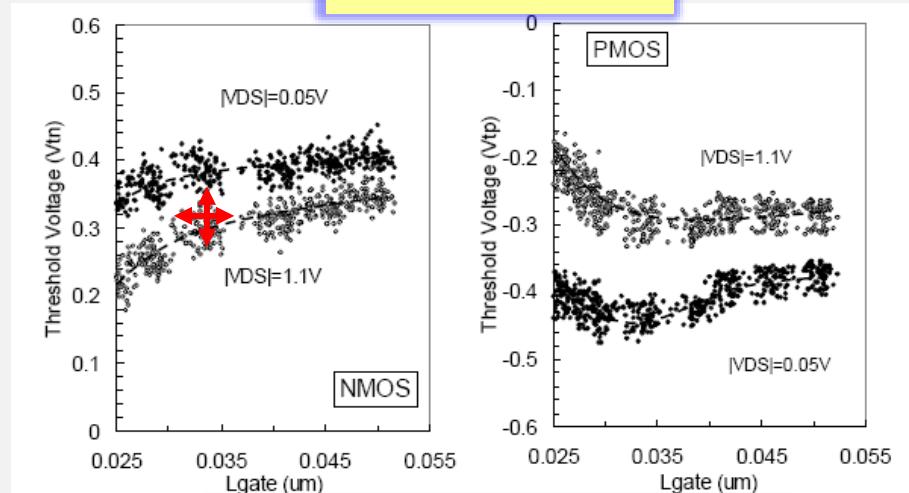
S: ~ 65mV/dec  
DIBL: ~ 35mV/V

2000 Node: 130 nm

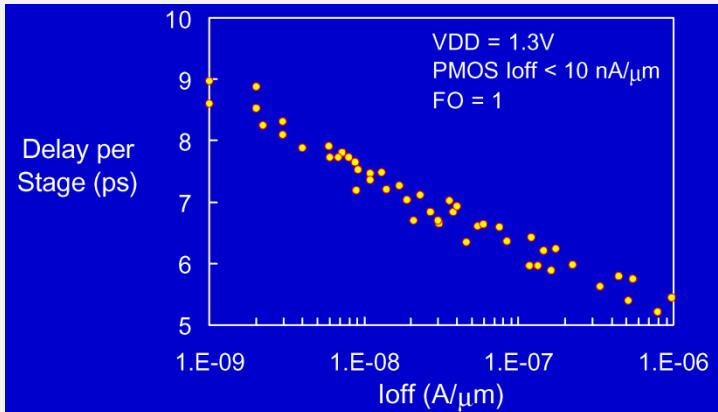


$V_t$  roll-off: good above  $L > 70$  nm

2008 Node: 45 nm



$V_t$  roll-off: good above  $L > 35$  nm



stage delay: ~7psec

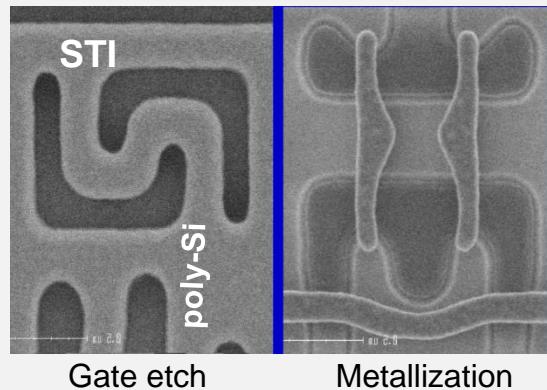


high fluctuations  $\rightarrow$  challenge for circuit design

SRAM:

For demonstration of design improvement **6T-SRAM cells** are used (usually embedded in MPU):  
 - regular structure  
 - easy to evaluate **density** and **yield** (like DRAM, DRAM-pitch)

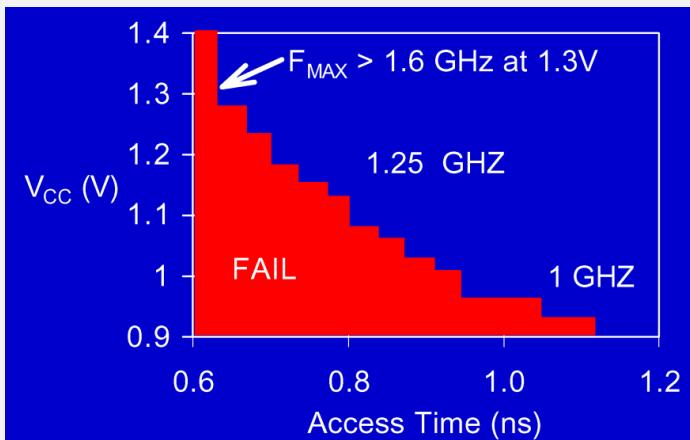
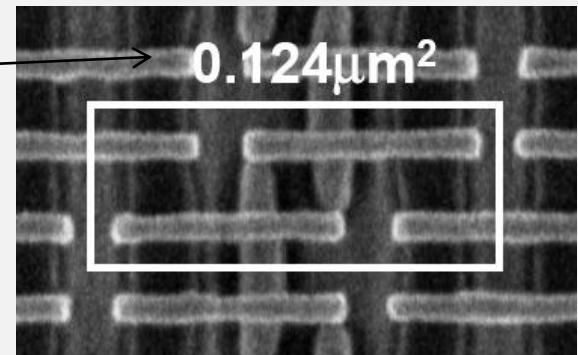
2000 Node: 130 nm



Typical cell size:  
 $(1.3 \times 1.7) \mu\text{m}^2 = 2.2 \mu\text{m}^2$

! Make a test:  
 2000  $\rightarrow$  2009 = 3 generations  
 $\rightarrow$  does scaling work ?

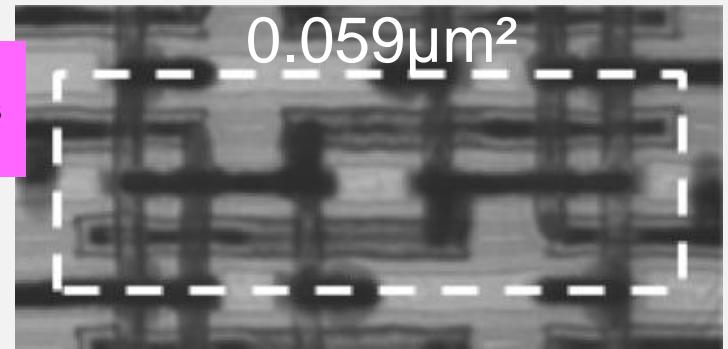
2009 Node: 32 nm



Schmoo-plot ( $f_{max} = f(V_{DD})$  of SRAM performance)

! Make a test:  
 2009  $\rightarrow$  2014 = 2 generations  
 $\rightarrow$  does scaling work ?

2014 Node (FinFET): 14 nm



2014 Node (FinFET): 14 nm

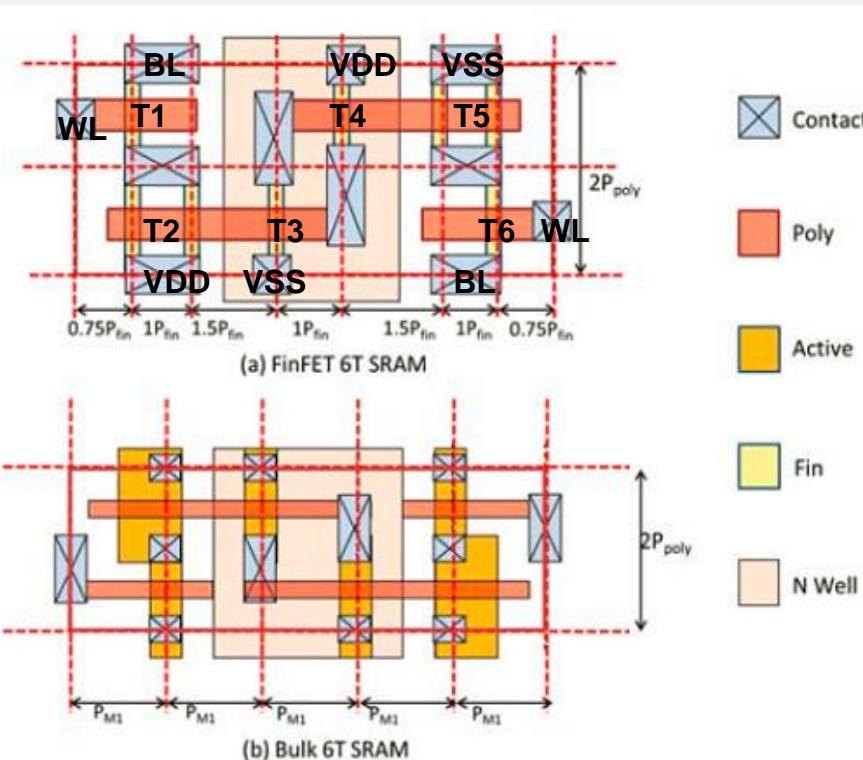
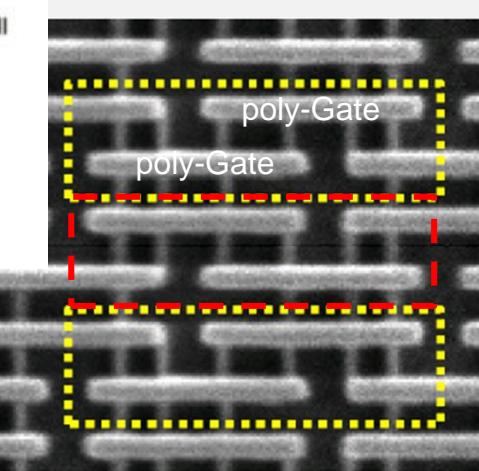
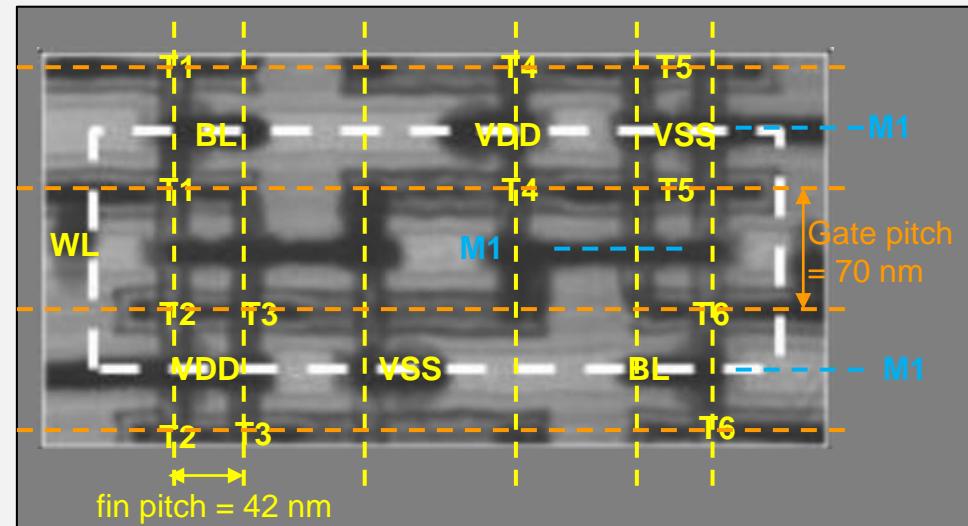
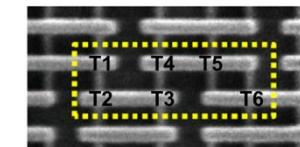
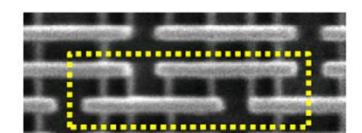
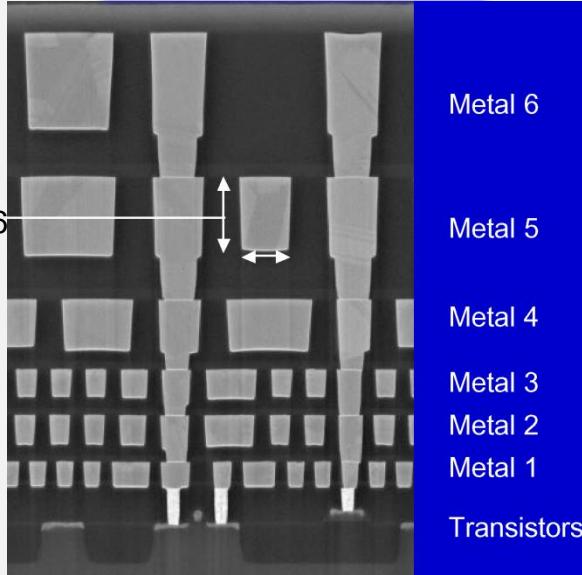


Figure 5: Layout of 6T SRAMs for (a) FinFET and (b) bulk.

courtesy: A.B.Kahng; "The ITRS Design ... ",  
Design Automation Conference, 2013

Intel's 6T SRAM options in their SoC technology, including high density / low leakage (HDC), low voltage (LVC), and high performance (HPC)

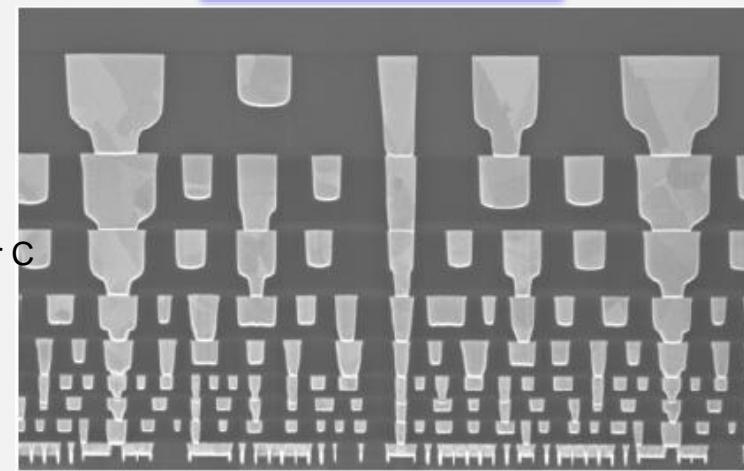
6T SRAM  
HDC6T SRAM  
LVC6T SRAM  
HPC



Interconnects limit the RC delay of faster devices

-> more metallization layers -> lower R  
-> higher aspect ratios/ low-k ILD -> lower C

Typically:  
# 6-8 layers of interconnects  
# Cu + ILD  
# every layer separately optimized



**Interconnect Dimensions**

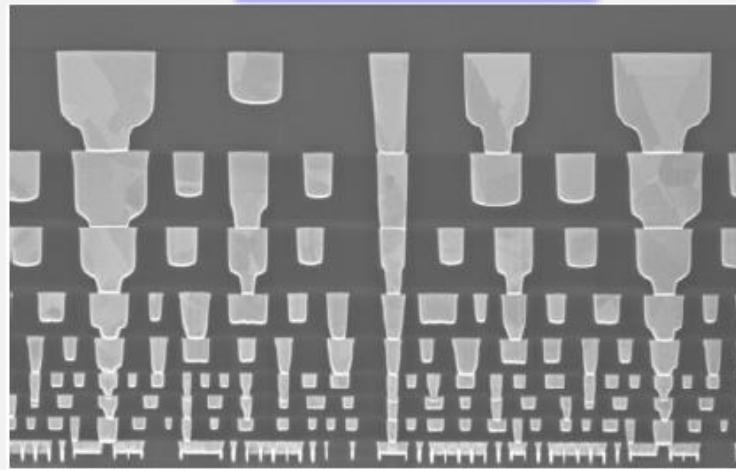
Layer	Pitch	Thick	A.R.
Metal 1	350	280	1.6
Metal 2	448	360	1.6
Metal 3	448	360	1.6
Metal 4	756	570	1.5
Metal 5	1120	900	1.6
Metal 6	1204	1200	2.0

nm nm

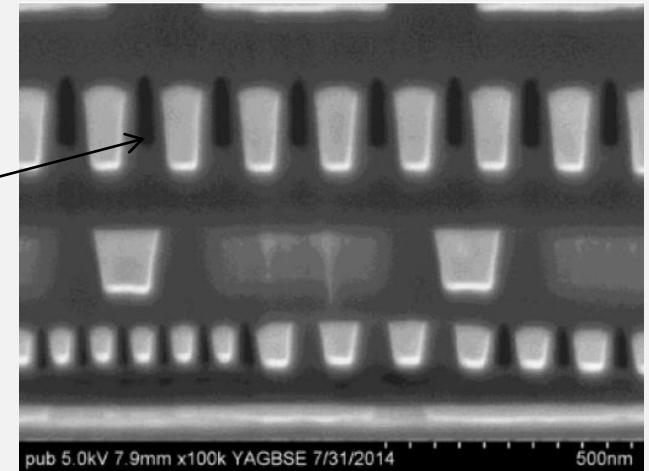
Minimum pitch on device layer  
(isolation, gate, contact, metal1)

Layer	Pitch (nm)	Thick (nm)	Aspect Ratio
Isolation	140.0	200	-
Contacted Gate Pitch	112.5	35	-
Metal 1	112.5	95	1.7
Metal 2	112.5	95	1.7
Metal 3	112.5	95	1.7
Metal 4	168.8	151	1.8
Metal 5	225.0	204	1.8
Metal 6	337.6	303	1.8
Metal 7	450.1	388	1.7
Metal 8	566.5	504	1.8
Metal 9	19.4 $\mu$ m	8 $\mu$ m	1.5
Bump	145.9 $\mu$ m	25.5 $\mu$ m	-

2009 Node: 32 nm



2014 Node (FinFET): 14 nm



Interconnects limit the RC delay of faster devices

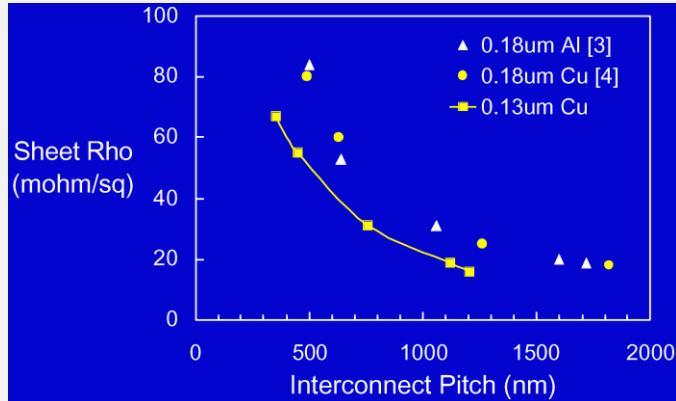
- > more metallization layers -> lower R
- > higher aspect ratios/ low-k ILD -> lower C

Layer	Pitch (nm)	Thick (nm)	Aspect Ratio
Isolation	140.0	200	-
Contacted Gate Pitch	112.5	35	-
Metal 1	112.5	95	1.7
Metal 2	112.5	95	1.7
Metal 3	112.5	95	1.7
Metal 4	168.8	151	1.8
Metal 5	225.0	204	1.8
Metal 6	337.6	303	1.8
Metal 7	450.1	388	1.7
Metal 8	566.5	504	1.8
Metal 9	19.4 $\mu$ m	8 $\mu$ m	1.5
Bump	145.9 $\mu$ m	25.5 $\mu$ m	-

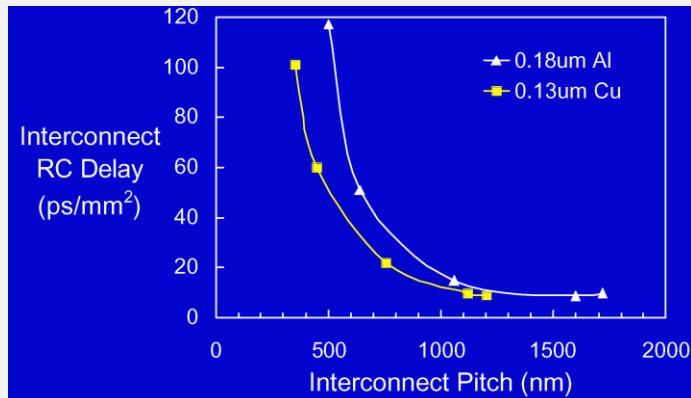
Minimum pitch on device layer  
(isolation, gate, contact, metal1)

Layer	Pitch (nm)	Scale Factor to [1]
Fin	42	0.70
Contacted Gate Pitch	70	0.78
Metal 0	56	N/A
Metal 1	70	0.78
Metal 2	52	0.65

2000 Node: 130 nm

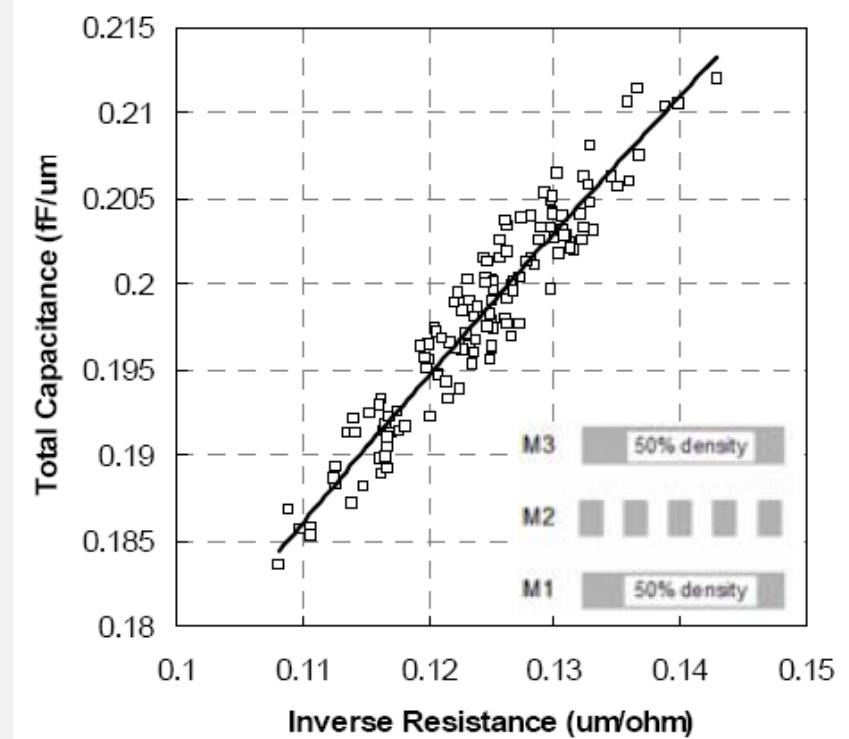


pitch-dependent sheet resistance



RC delay for 1mm wire length, line capacity: ~ 250fF/mm for M1-M5

2009 Node: 32 nm



- **Willamette (2000)**

180nm, channel: 90nm, 217mm<sup>2</sup>, < 2 GHz, L2 =256kB, 42 Mill. T

- **Northwood (2002)**

130nm, channel: 65nm, 146mm<sup>2</sup>, < 3GHz, L2=512kB, 55 Mill. T

- **Prescott (2004)**

90nm, low-k, metal: 7 levels Cu, channel: 45nm, 112mm<sup>2</sup>, < 3.4GHz,  
L2=1024kB, 125 Mill. T

- **Cedar Mill (2006) [tick]**

65nm, strain channels with SiGe S/D &Si<sub>3</sub>N<sub>4</sub>, 2 cores, 90mm<sup>2</sup>, 184 Mill. T

- **Merom (2007) [tock]**

65nm, strain channels with SiGe S/D &Si<sub>3</sub>N<sub>4</sub>, 2 cores, 143mm<sup>2</sup>, 291 Mill. T

- **Penryn (2008) [tick]**

45nm, high-k + metal gate, 6 cores, 263mm<sup>2</sup> up to 731 Mill. T.

- **Nehalem (2009) [tock]**

45nm, high-k + metal gate, 8 cores, 684mm<sup>2</sup> up to 2.3 Bill. T.

- **Westmere (2010) [tick]**

32nm, high-k + metal gate, 10 cores, 512mm<sup>2</sup> up to 2.6 Bill. T.

- **Sandy Bridge (2011) [tock]**

32nm, high-k + metal gate, 10 cores, 434mm<sup>2</sup> up to 2.3 Bill. T.

- **Ivy Bridge (2012) [tick]**

22nm, FinFET, 15 cores, 541mm<sup>2</sup> up to 4.3 Bill. T.

- **Haswell (2013) [tock]**

22nm, FinFET, 18 cores, 661mm<sup>2</sup> up to 5.5 Bill. T.

INTEL Prescott: die view with marked memory areas

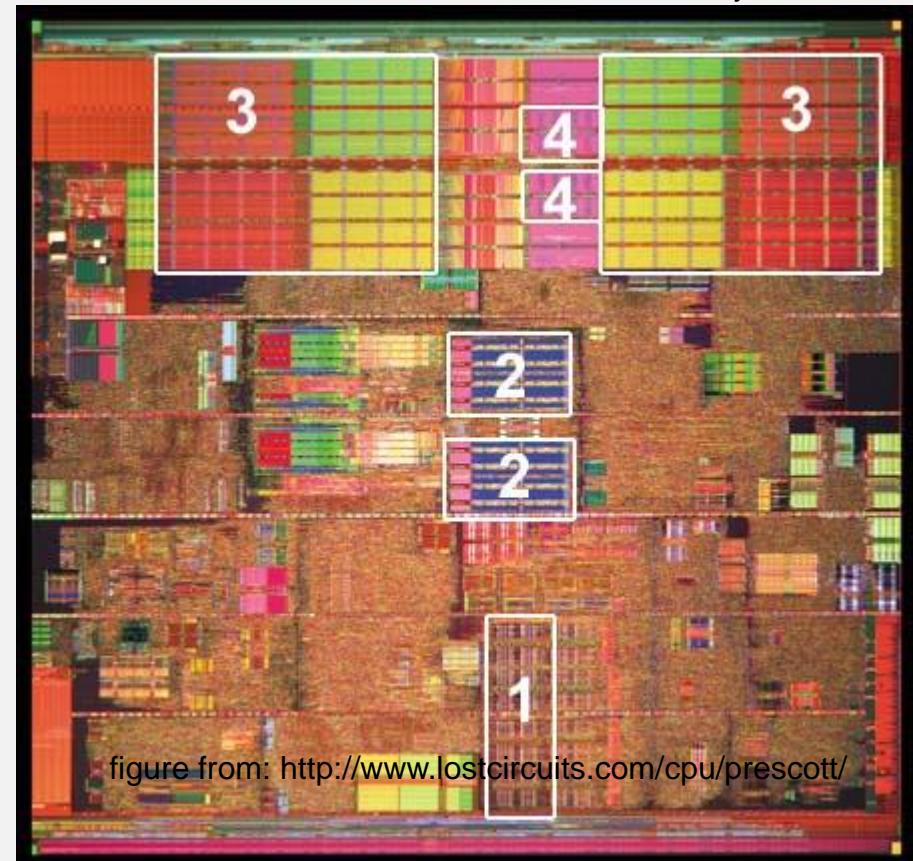


figure from: <http://www.lostcircuits.com/cpu/prescott/>

1: Execution Trace Cache; 2: L 1 Data Cache with muxer (pink) = 20kB;  
3: L 2 cache blocks = 1024kB; 4: L 2 tag cache

- **Broadwell (2014) [process]**

14nm, FinFET, 22 cores, 456mm<sup>2</sup> up to 7.2 Bill. T

- **Skylake (2015) [architecture]**

14nm, FinFET, 28 cores, 485 mm<sup>2</sup> up to 8.3 Bill. T

- **Kaby Lake (2016) [optimization]**

14nm, FinFET, 4 cores, 126 mm<sup>2</sup> up to 2.2 Bill. T

- **Coffee Lake (2017) [2<sup>nd</sup> optimization]**

14nm, FinFET, 8 cores, 174 mm<sup>2</sup> up to ? Bill. T

- **Cannon Lake (2018) [process]**

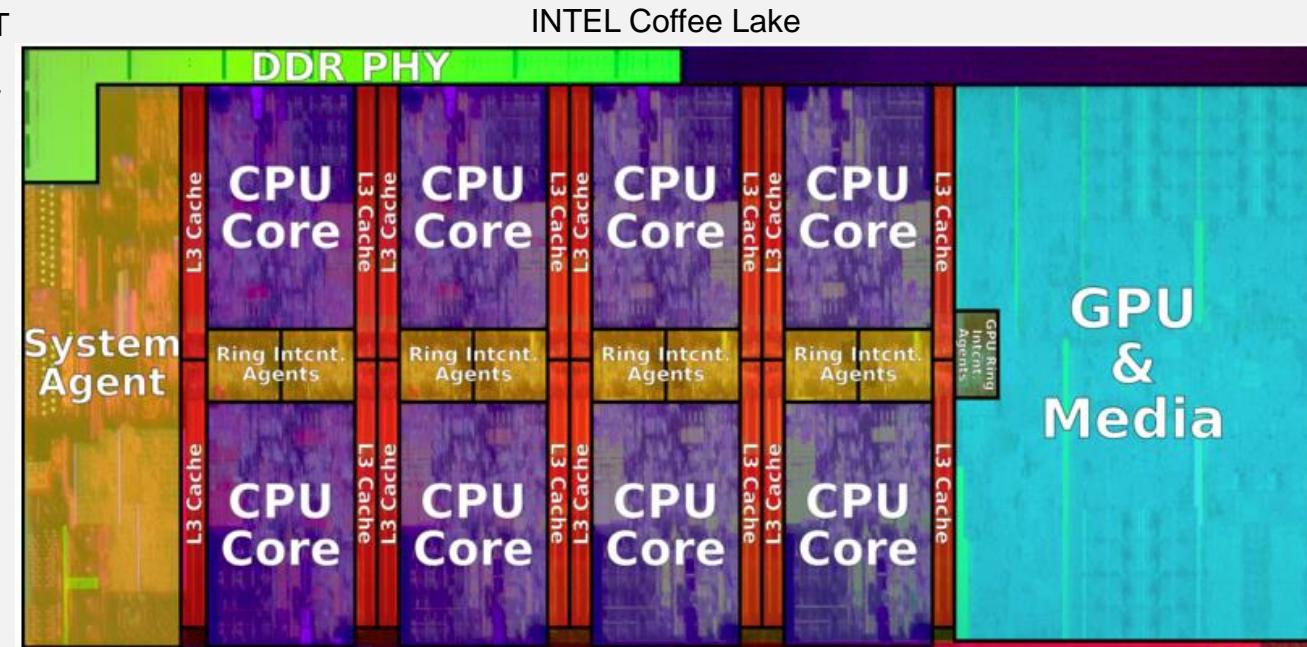
10nm, FinFET, ?? cores, ? mm<sup>2</sup> up to ? Bill. T

- **Ice Lake (2020) [architecture]**

10nm, FinFET, ?? cores, ? mm<sup>2</sup> up to ? Bill. T

- **Tiger Lake (2021?) [optimization]**

10nm, FinFET, ?? cores, ? mm<sup>2</sup> up to ? Bill. T



### Cannon Lake stumbles into the market: The IdeaPad 330-15ICN is the first laptop with a 10-nm-CPU

*Intel Cannon Lake once was a very highly anticipated CPU-release. After massive delays, the first Intel CPU generation that is manufactured in the advanced 10 nanometer process has finally arrived – and it enters the market silently, in a low-budget device.*

by Benjamin Herzig, 2018/05/13

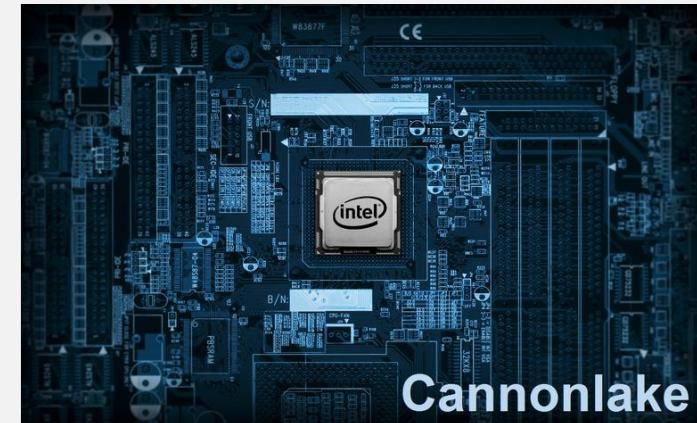
Cannon Lake Ultrabook Laptop Coffee Lake AMD

The first 10 nm laptop has been announced: When Lenovo introduced a couple of new IdeaPad laptops [a few days ago](#), this release also included the IdeaPad 330-15ICN – "ICN" stands for "Intel Cannon Lake". For those not in the know, Cannon Lake was once upon a time announced as the next big step in Intels release schedule, the first CPU generation to be manufactured in the 10 nm process. Originally, its release was expected in late 2016, almost two years ago.

<https://www.notebookcheck.net/Cannon-Lake-stumbles-into-the-market-The-IdeaPad-330-15ICN-is-the-first-laptop-with-a-10-nm-CPU.303330.0.html>



*Cannon Lake stumbles into the market: The IdeaPad 330-15ICN is the first laptop with a 10-nm-CPU*



<https://www.tweaktown.com/news/59209/intel-delays-10nm-cpu-tech-third-time-late-2018/index.html>



Because of yield problems, mass production of the new Cannon Lake processors will start 2019



It is expected that Cannon Lake is a test drive for 10 nm production, will be replaced by Ice Lake in 2020

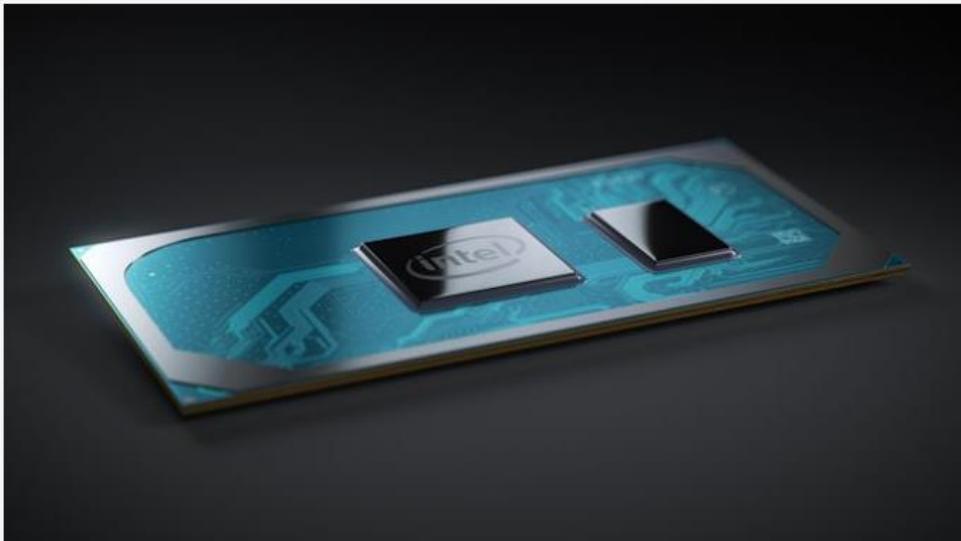
### Intel Begins Commercial Shipments of 10nm Ice Lake CPUs to OEMs

by Anton Shilov on July 26, 2019 5:30 PM EST

Posted in [CPUs](#) [Intel](#) [10nm](#) [Ice Lake](#)

115  
Comments

+ Add A Comment



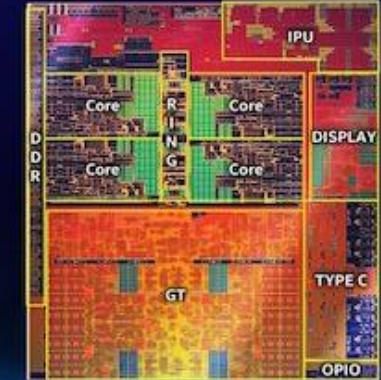
Intel has begun shipments of its 10th generation Core "Ice Lake" processors as of the second quarter, according to the company in an earnings call this week. Made using Intel's 10nm process technology, these laptop CPUs were qualified by OEMs earlier in 2019 and are on track to reach the market inside mobile PCs by the holiday season.

<https://www.anandtech.com/show/14679/intel-begins-commercial-shipments-of-10-nm-ice-lake-cpus-to-oems>

Mass production of Ice Lake started in late 2019

7nm node is planned for mass production in 2021

10NM ICE LAKE CLIENT  
Shipping in June



INTEL, 2019

INVESTOR MEETING

### OUR GAME PLAN... INVESTING IN PROCESS LEADERSHIP

#### EXTEND 14NM

Build Capacity  
to Support  
Customer Growth

vs. TSMC 10NM

#### RAMP 10NM

Client Systems on  
Shelf for 2019  
Holiday Season  
Server in 1H'20

vs. TSMC 7NM

#### ACCELERATE TO 7NM

Production and  
Launch in 2021

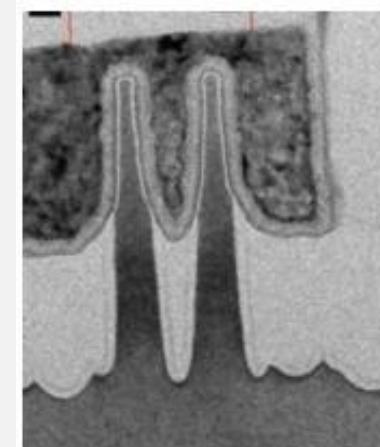
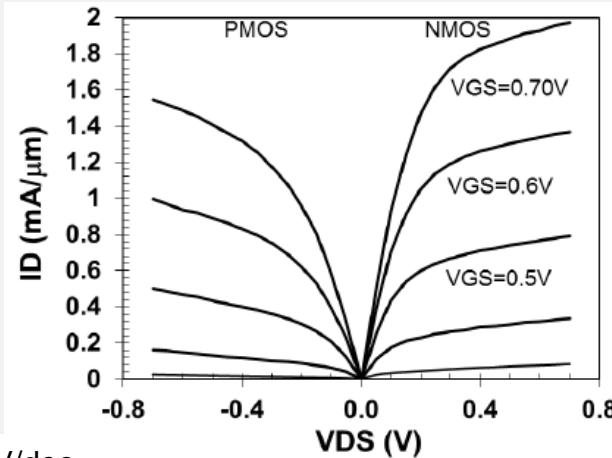
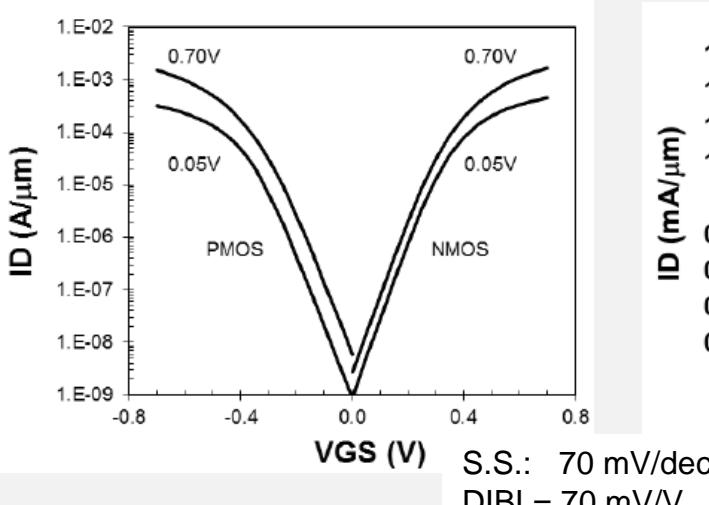
vs. TSMC 5NM

WORLD CLASS PACKAGING TECHNOLOGY COMPLEMENTS PROCESS LEADERSHIP

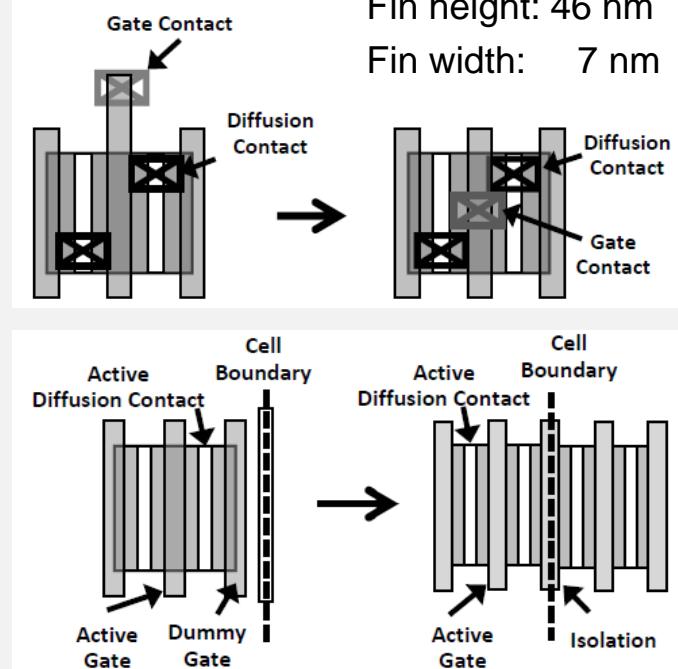
intel INVESTOR MEETING

**A 10nm High Performance and Low-Power CMOS Technology Featuring 3<sup>rd</sup> Generation FinFET Transistors, Self-Aligned Quad Patterning, Contact over Active Gate and Cobalt Local Interconnects**

C. Auth, A. Aliyarunkunju, M. Asoro, D. Bergstrom, V. Bhagwat, J. Birdsall, N. Bisnik, M. Buehler, V. Chikarmane, G. Ding, Q. Fu, H. Gomez, W. Han, D. Hanken, M. Haran, M. Hattendorf, R. Heussner, H. Hiramatsu, B. Ho, S. Jaloviar, I. Jin, S. Joshi, S. Kirby, S. Kosaraju, H. Kothari, G. Leatherman\*, K. Lee, J. Leib, A. Madhavan, K. Marla, H. Meyer, T. Mule, C. Parker, S. Parthasarathy, C. Pelto, L. Pipes, I. Post, M. Prince, A. Rahman\*, S. Rajamani, A. Saha, J. Dacuna Santos\*, M. Sharma, V. Sharma, J. Shin, P. Sinha, P. Smith, M. Sprinkle, A. St. Amour, C. Staus, R. Suri, D. Towner, A. Tripathi, A. Tura, C. Ward, A. Yeoh  
 Logic Technology Development, \*Quality and Reliability Engineering, Intel Corporation IEEE 2017



Fin pitch: 34 nm  
Fin height: 46 nm  
Fin width: 7 nm



► Working 10nm FinFET, Si or SiGe?

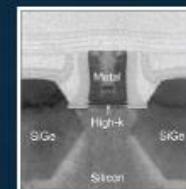
► Cobalt is used for three local interconnects

► Volume production began in 2019

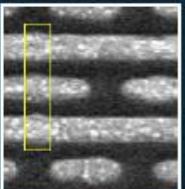
## INTEL INNOVATION LEADERSHIP



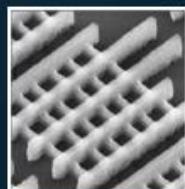
Strained  
Silicon



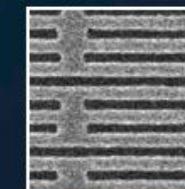
High-k  
Metal Gate



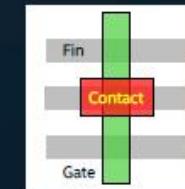
Self  
Align Via



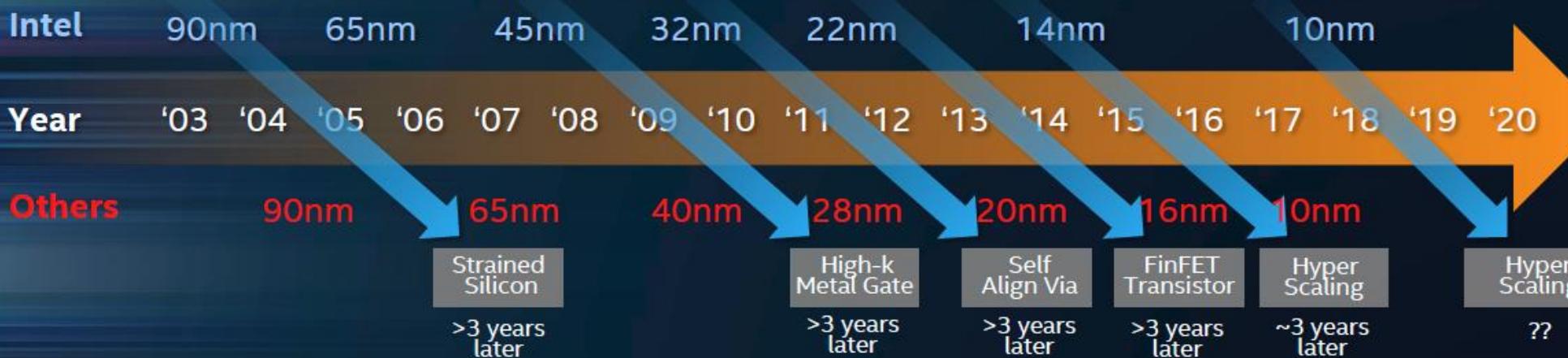
FinFET  
Transistor



Hyper  
Scaling



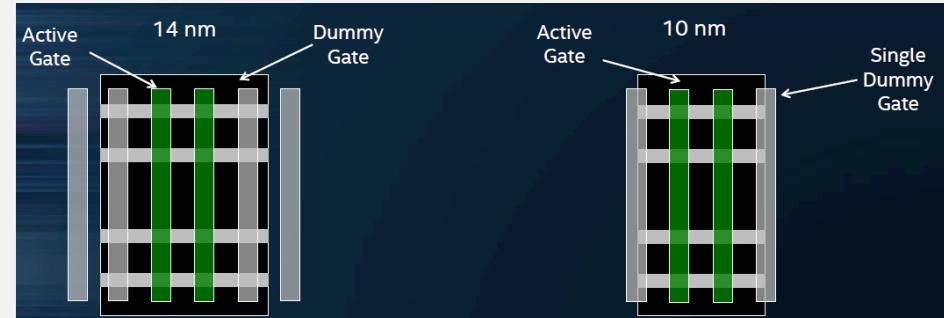
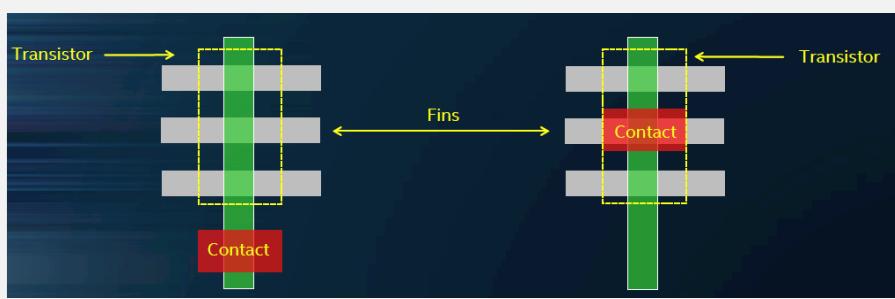
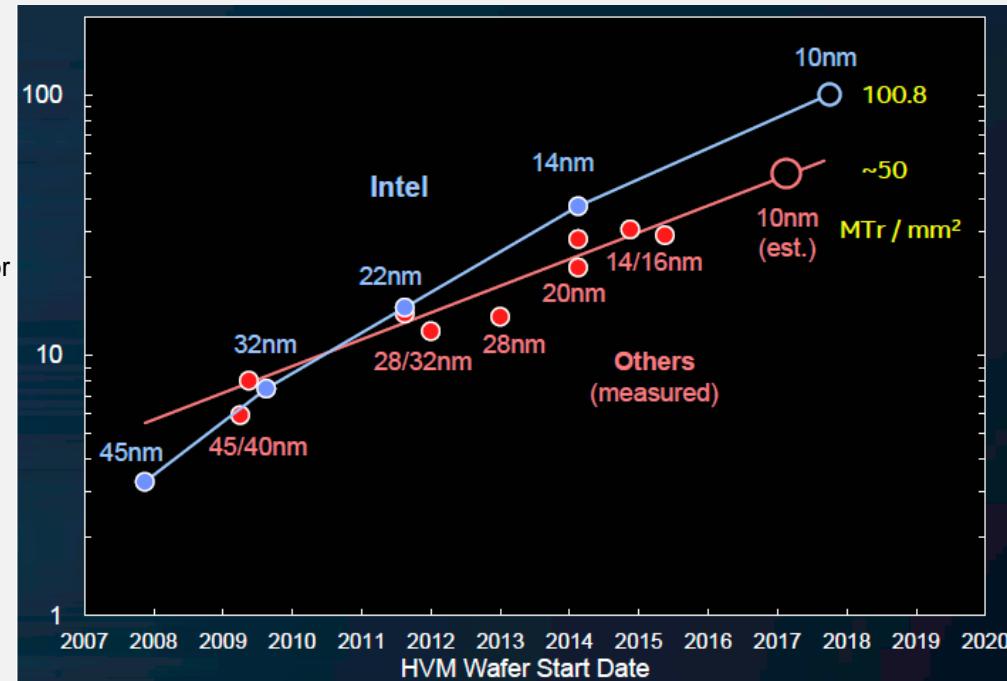
Hyper  
Scaling



INTEL, 2017



In history INTEL wanted to be the most innovative semiconductor manufacturer



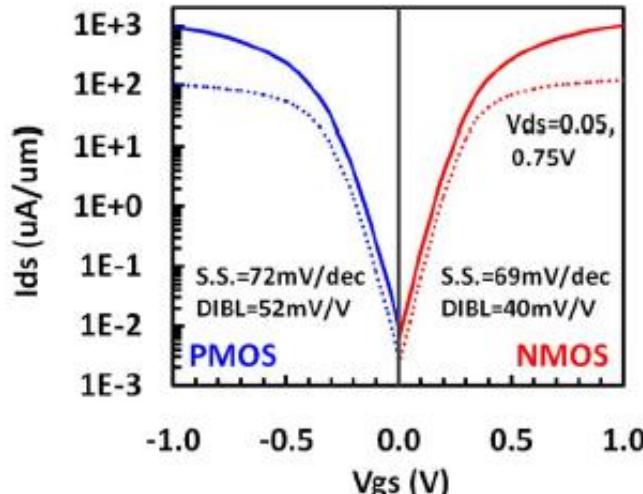
Through clever technological design (hyper scaling) INTEL seems to achieve a higher transistor density

# A 7nm CMOS Technology Platform for Mobile and High Performance Compute Application

IEEE 2017

S. Narasimha, B. Jagannathan, A. Ogino, D. Jaeger, B. Greene, C. Sheraw, K. Zhao, B. Haran, U. Kwon, A.K.M. Mahalingam, B. Kannan, B. Morganfeld, J. Dechene, C. Radens, A. Tessier, A. Hassan, H. Narisetty, I. Ahsan, M. Aminpur, C. Ari, M. Aquilino, A. Arya, R. Augur, N. Baliga, R. Bhelkar, G. Biery, A. Blauberg, N. Borjemscaia, A. Bryant, L. Cao, V. Chauhan, M. Chen, L. Cheng, J. Choo, C. Chua, T. Chu, B. Colgan, R. Coleman, D. Conklin, S. Crown, A. da Silva, D. Dechene, G. Derderian, S. Deshpande, G. Dillaway, K. Donegan, M. Eller, Y. Fan, Q. Fang, A. Gasse, C. Gauthier, C. Ghosh, C. Gifford, T. Gordon, M. Gribelyuk, G. Han, J.H. Han, K. Han, M. Hasan, J. Higman, J. Holt, L. Hu, L. Huang, T. Hung, Y. Jin, J. Jimenez, S. Johnson, V. Joshi, P. Justison, S. Kalaga, T. Kim, W. Kim, R. Krishnan, B. Krishnan, Anil K., M. Kumar, J. Lee, R. Lee, J. Lemke, S.L. Lie, P. Liao, J. Lipnugari, J. Lipnick, P. Liu, J. Liu, S. Lucarini, W. Ma, E. Maciejewski, S. Madisetti, A. Malinowski, J. Mata, C. Menendez, J. Montoya, T. Naresh, T. Ngan, C. North, K. Onishi, C. Ordonio, M. Ozbek, R. Pal, S. Parihar, O. Patterson, E. Ramanathan, I. Ramirez, R. Ranjan, J. Sarad, V. Sardesai, S. Saudari, C. Schiller, B. Senapati, C. Serrau, N. Shah, T. Shen, H. Sheng, J. Shepard, Y. Shi, M.C. Silvestre, D. Singh, Z. Song, J. Sporre, P. Srinivasan, Z. Sun, A. Sutton, R. Sweeney, K. Tabakman, M. Tan, X. Wang, E. Woodard, G. Xu, D. Xu, T. Xuan, Y. Yan, J. Yang, K.B. Yeap, M. Yu, A. Zainuddin, J. Zeng, K. Zhang, M. Zhao, Y. Zhong, R. Carter, C.-H. Lin, S. Grunow, C. Child, M. Lagus, R. Fox, E. Kaste, G. Gomba, S. Samavedam, P. Agnello, and DK Sohn

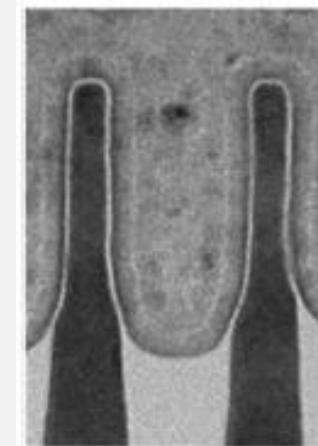
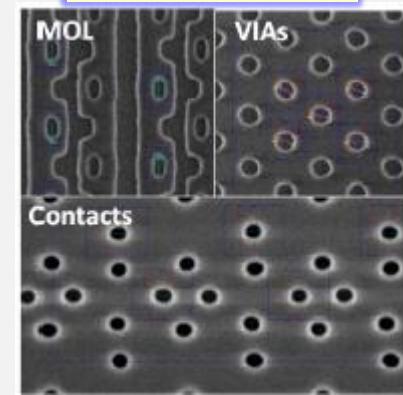
GLOBALFOUNDRIES, 400 Stonebreak Road Extension, Malta, NY 12020



- SRAM cell of  $0.027 \mu\text{m}^2$
- High-k/Metal-Gate
- Replacement-Gate Process
- 12-level Cu/low-k
- Cobalt as contact metallization
- Si or SiGe?



EUV lithography



Fin pitch: 30 nm

Working 7nm FinFET, Si or SiGe?

Immersion lithography is used

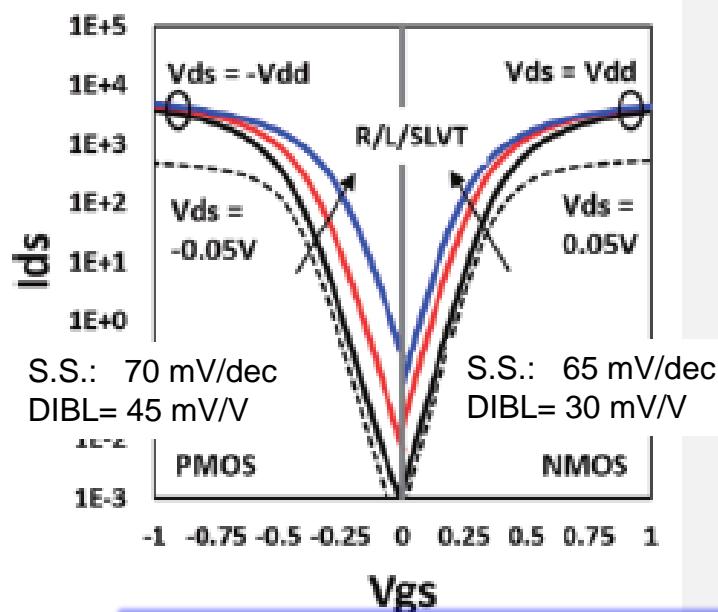
Volume production is put on hold, because the costs are too high

## Highly Manufacturable 7nm FinFET Technology featuring EUV lithography for Low Power and High Performance Applications

IEEE 2017

Daewon Ha, C. Yang, J. Lee, S. Lee, S. H. Lee, K.-I. Seo, H.S. Oh, E.C. Hwang, S.W. Do, S.C. Park, M.-C. Sun, D.H. Kim, J.H. Lee, M.I. Kang, S.-S. Ha, D.Y. Choi, H. Jun, H.J. Shin, Y.J. Kim, J. Lee, C.W. Moon, Y.W. Cho, S.H. Park, Y. Son, J.Y. Park †, B.C. Leet, C. Kim†, Y.M. Oh†, J.S. Park†, S.S. Kim†, M.C. Kim†, K.H. Hwang†, S.W. Nam†, S. Maeda, D.-W. Kim, J.-H. Lee, M.S. Liang, ES Jung

Logic TD & †Process Development Team, Semiconductor R&D Center, Samsung Electronics Co. Ltd., San#16, Banweol-Dong, Hwasung-City, Gyeonggi-Do, 445-701, Republic of Korea, email: daewon.ha@samsung.com



- High-k/Metal-Gate
- Replacement-Gate Process
- Si or SiGe?

Working 7nm FinFET, Si or SiGe?

EUV lithography is used for some critical process steps

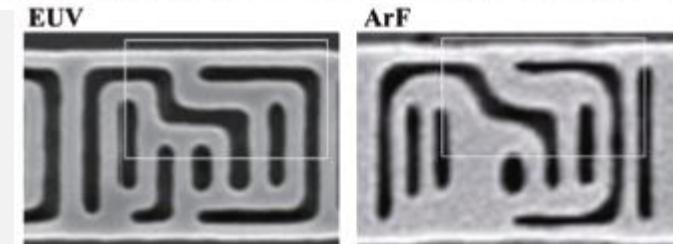
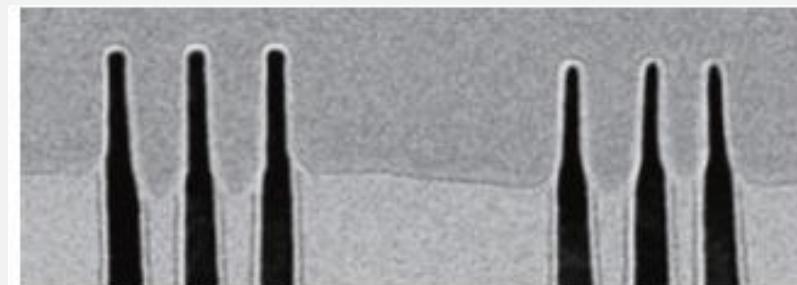
Volume production started in end of 2018

Samsung Electronics Starts Production of EUV-based 7nm LPP Process

Korea on October 18, 2018

AUDIO

SHARE



	Intel 22 nm	Intel 14 nm	Samsung 10 nm	TSMC 10 nm	Intel 10 nm	Globalfoundries 7 nm	Samsung 7 nm	TSMC 7 nm
Fin Pitch [nm]	60	42	42	36	34	30	?	?
Gate Pitch [nm]	90	70	64/68	66	54	56	57	57
Min. Metal Pitch [nm]	80	52	48	42	36	40	36	40
SRAM Cell Area [ $\mu\text{m}^2$ ]	0.108	0.059	0.04	?	0.031	0.027	0.026	0.027
MTx/mm <sup>2</sup>	x	x	x	x	106 	98	95	96
Fin Width [nm]	?	8	?	?	7	~7	?	?
N-channel DIBL mV/V	46	65	30	?	70	40	30	35
N-channel S.S. [mV/dec]	69	70	65	?	70	69	65	65



Intel 10nm node is more close to the 7nm node of the foundries



Foundries do a hybrid scaling and not a full node scaling

## 6.1 Overview

Feature size, pitch, node

## 6.2 State of the Art MOSFETs

Electrical and geometrical parameters

## 6.3 Mobility Enhanced MOSFETs

Strain, SiGe, Liners, Directed Channels, Gate-Last

## 6.4 High-k Metal-Gate MOSFETs

Poly-Depletion, FuSi

## 6.5 Vertical MOSFETs

## 6.6 Halo MOSFETs

## 6.7 SOI-MOSFETs

## 6.8 Multi-Gate MOSFETs

Fully-depleted MOSFETs, FinFETs, Nanowire FETs

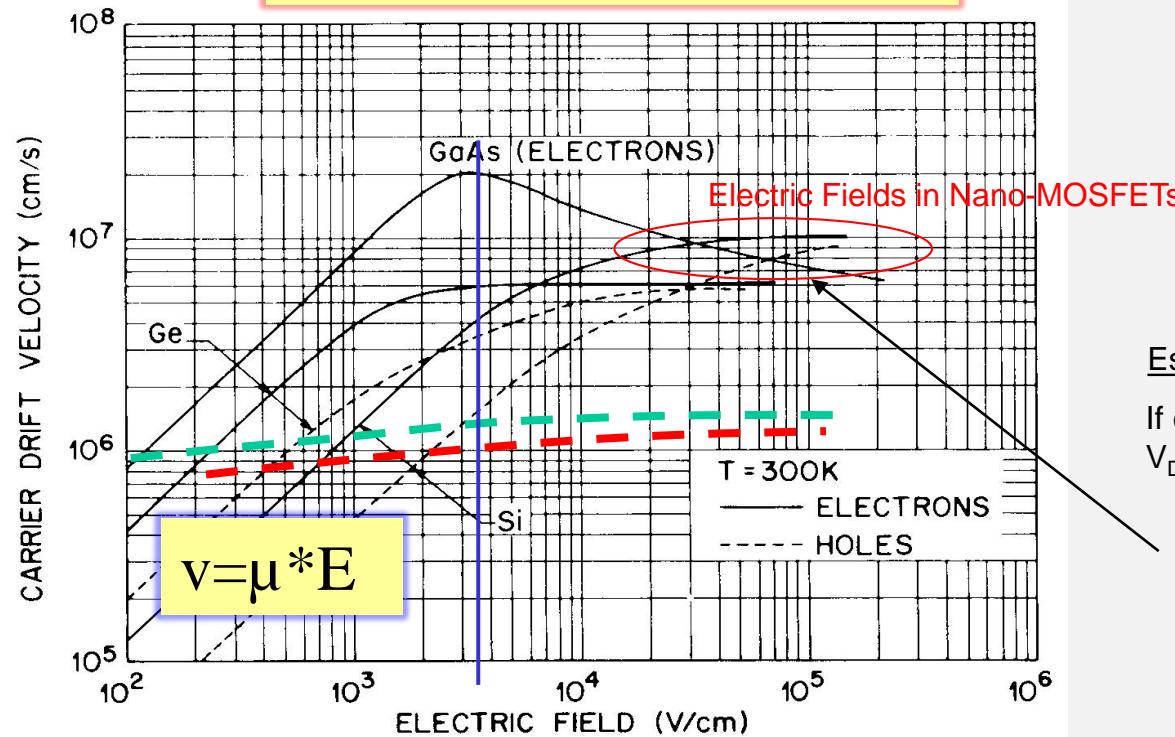
## 6.9 Advanced Memory Devices

The  $I_{on}$  current of a MOSFET is given by:

$$I_D = \mu \cdot C_{ox} \cdot \frac{w}{L} \cdot \left[ (V_{GS} - V_T) \cdot V_{DS} - \frac{V_{DS}^2}{2} \right]$$

**mobility** increases/decreases drive current and therefore High-Frequency performance

The problem: Mobility decreases in high Electric Fields



Sze, Irvin: RESISTIVITY, MOBILITY AND IMPURITY LEVELS IN GaAs, Ge, AND Si AT 300°K, Solid-State Electronics, 11(1968)599

Challenge for sub-100nm CMOS: increase mobility in high electric fields  
→ introduction of strain techniques for 90nm node

Estimation:

If channel length: 100nm  
 $V_{Drain} = 1V$

(Linear) Electric Field:  
 $1V/100nm = 10^5$  V/cm

bulk mobility is reduced

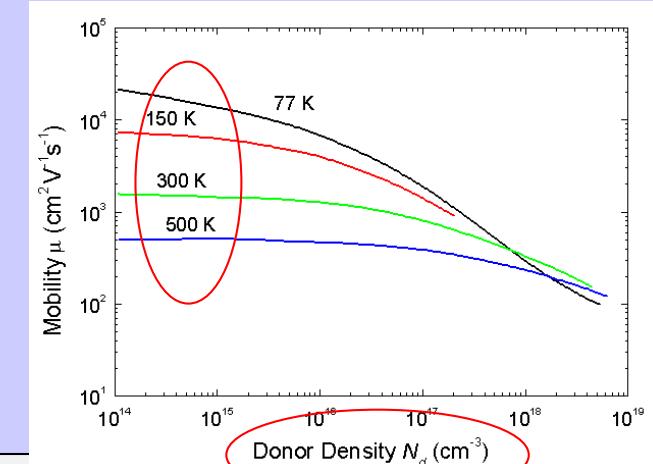
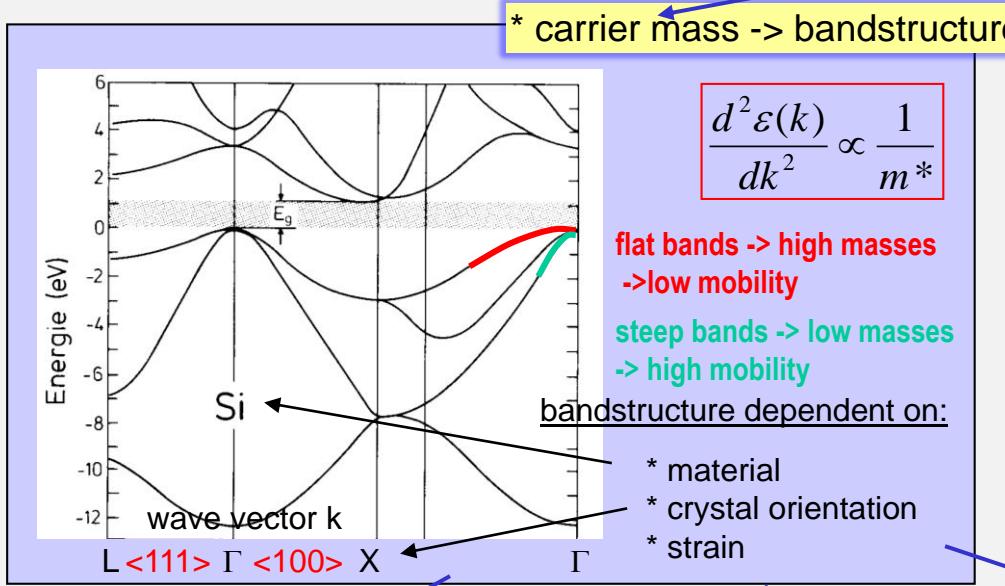


Field-effect mobility is basically dependent on:

$$\vec{v}_{drift} = \mu \cdot \vec{E} = -\frac{e \cdot \tau_\lambda}{2m^*} \cdot \vec{E}$$

\* carrier scattering events

- > non-periodic atoms (doping, surface roughness)
- > lattice vibrations (temperature)
- > transport (ballistics, tunneling)



## Material

with higher bulk mobilities

Ge, SiGe

inside the channel

-> material  
(SiGe, ...)

from surroundings

-> liner  
-> S/D

## Crystal Orientation

surface

(100), (110), ...

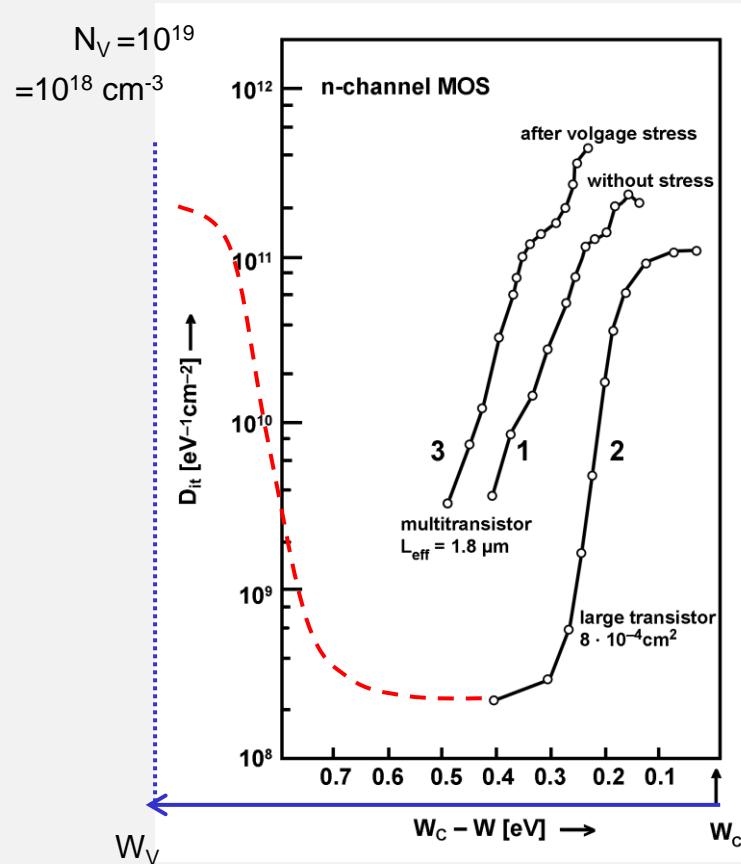
direction

<110>, <100>, ...

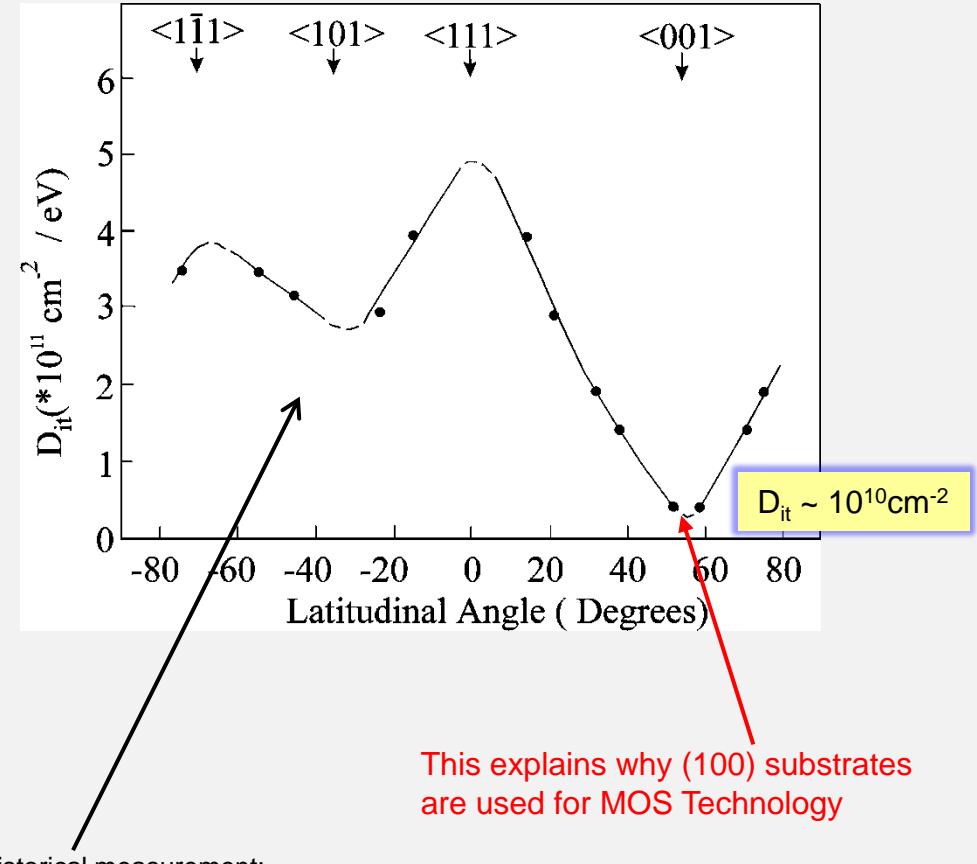
and combinations

Due to interface states  $D_{it}$  from 1960 (first working MOSFET) until ~ 2007 the only possibility to create a proper working MOSFET was using Si(100) and thermal oxidation for the Gate oxide

Energetic Distribution in the Band Gap



Dependence on Surface Orientation



historical measurement:

Arnold et al., CRYSTALLOGRAPHIC SYMMETRY OF SURFACE STATE DENSITY IN THERMALLY OXIDIZED SILICON, Appl.Phys.Lett 13(1968)413

Why we need an interface state density  $D_{it} \sim 10^{10} \text{ cm}^{-2}$  or lower ?

Let's make an estimation:

### 1) How many charges are in a switched-on channel ?

Estimation 1) a conducting channel acts as a metallic layer

-> for silicon this is the case, if the Fermi level is within the bands -> this needs dopings above  $N_D > 10^{19} \text{ at/cm}^3 = n_e$  because the thickness of the channel is only a few nanometer, the (volume) charge carrier density  $n_e$  can be expressed as area charge density  $n_e''$  (carrier/cm<sup>2</sup>).

-> to calculate the area charge density  $n_e''$  we can multiply the volume charge density  $n_e$  by the thickness of the channel (~5nm)

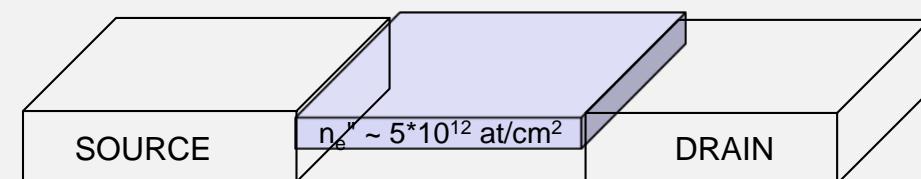
$$\rightarrow 10^{19} \text{ at/cm}^3 * 5 \text{ nm} = 10^{19} \text{ at/cm}^3 * 5 * 10^{-7} \text{ cm} = 5 * 10^{12} \text{ e/cm}^2$$

Estimation 2)

The Gate capacity is about:  $C'' = \epsilon_0 \epsilon_{ox} / t_{ox}$ ,  $\rightarrow C'' \sim 10^{-13} * 4 / 4 \text{ nm} = 10^{-6} \text{ F/cm}^2$ .

This results in a charge of:  $Q = C'' * U = 10^{-6} \text{ F/cm}^2 * 1 \text{ V} = 10^{-6} \text{ Cb/cm}^2$

$$\Rightarrow \text{number of electrons: } n_e'' = 10^{-6} \text{ Cb/cm}^2 / 1.6 * 10^{-19} \text{ Cb/e} \rightarrow 6 * 10^{12} \text{ e/cm}^2$$

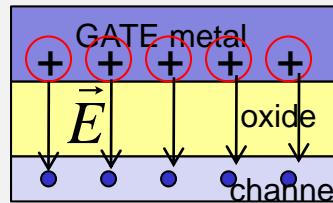


In a switched-on MOSFET-channel around  $10^{12} - 10^{13}$  electrons/cm<sup>2</sup> are existing



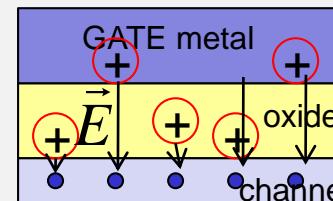
How many in a 32nm MOSFET ?

### 2) How are the channel charges created ?



Ideal case:

- perfect (charge-free dielectrics)
- > each metal gate charge creates a channel charge



Non-ideal case:

- oxide charges may create channel charges without external Gate voltage or are screening the external voltage
- > no channel charge control, if oxide charges ( $D_{it}$ ) reach same amount as desired channel charges



for good channel control the  $D_{it}$  should be less 1% of the channel charge  
 $\rightarrow D_{it} \leq 10^{10} \text{ cm}^{-2}$

Due to interface states  $D_{it}$  from 1960 (first working MOSFET) until ~ 2007 the only possibility to create a proper working MOSFET was using Si(100) and thermal oxidation for the Gate oxide

### 1) Using **Germanium** in old days (1948-1954)

- > the oxidation of Ge delivers a very hygroscopic thermal Gate oxide
- > no long-term stable MOSFET
- > standard technology with Ge is bipolar

### 2) Using **Compound Semiconductors** (GaAs, InP, ...)

- > delivers inhomogeneous thermal Gate oxides  
(because one element is oxidizing faster than the other)
- > inhomogeneous oxide thickness along channel  
with different materials (work functions are changing)
- > no reproducible  $V_t$  for devices
- > standard technology with III-V is bipolar

Oxidfilme auf A<sup>III</sup>B<sup>V</sup>-Halbleitern

K. Löschke<sup>a</sup>, G. Kühn<sup>b</sup>, H.-J. Bilz Und<sup>c</sup> and G. Leonhardt<sup>a</sup>

#### Abstract

Thermal oxidation of GaP, InP, GaAs, GaSb and InSb was carried out at room temperature and at 300 °C in air. The optical constants and film thicknesses were measured ellipsometrically and the same samples, including oxidized GaSb, were investigated by photoelectron spectroscopy. The concentration profiles of the elements in the films and their binding state were determined. It is shown that the elements in oxide films of GaP are oxidized perfectly, whereas films of the other compounds exhibit a very inhomogeneous structure. The thicknesses of the oxide films increase in the order  $GaP \approx InP < GaAs < GaSb, InSb$ . The interpretation of the ellipsometric measurements is complicated by the presence of a non-oxidized B<sup>V</sup> component, especially in the case of arsenides and antimonides.

### 3) For **deposited Gate dielectrics** the interface state density is for all materials close to $10^{12} /cm^2$

**GaAs metal-oxide-semiconductor devices with a complex gate oxide composed of SiO<sub>2</sub> and GaAs oxide grown using a photoelectrochemical oxidation method**  
Hsin-Ying Lee et al 2010 *Semicond. Sci. Technol.* 25 015005 (4pp)

**Abstract.** In this study, a SiO<sub>2</sub>/GaAs oxide bi-layer layer was used as the gate oxide in GaAs-based metal-oxide-semiconductor (MOS) devices.

The GaAs oxide layer of the bi-layer layer was directly formed on the GaAs surface by using the photoelectrochemical (PEC) oxidation method.

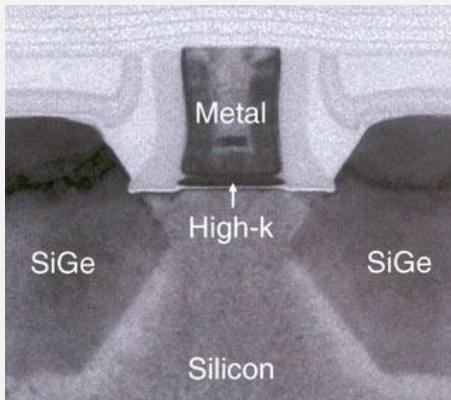
Some samples were thermally treated at 200 °C and 300 °C in O<sub>2</sub> ambience for 30 min.

The surface state density of the oxide/GaAs interface with and without GaAs oxide thermal treatment was  $7.2 \times 10^{11}$  and  $7.9 \times 10^{11} \text{ cm}^{-2} \text{ eV}^{-1}$ , respectively.

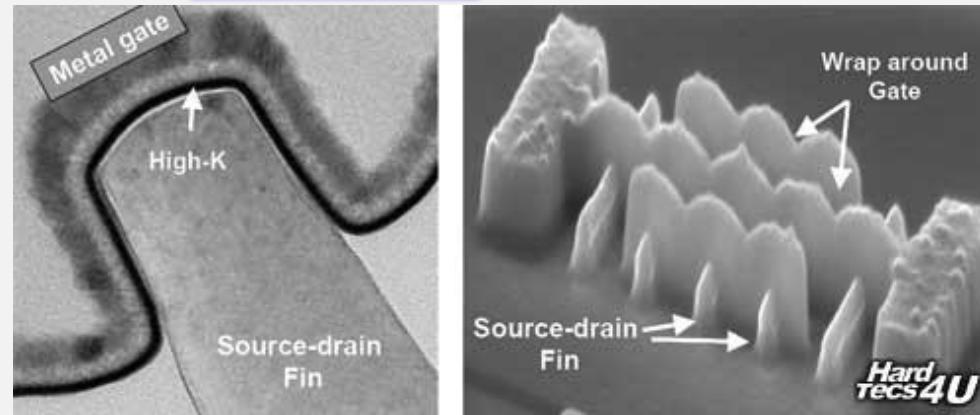
The GaAs MOS field effect transistors (MOSFETs) with the PEC-deposited GaAs oxide thermally treated showed an output current of  $152 \text{ mA mm}^{-1}$  at  $V_{DS} = 2.4 \text{ V}$  and  $V_{GS} = 0 \text{ V}$  and an extrinsic transconductance of  $89 \text{ mS mm}^{-1}$ .

But why we use deposited high-k today ?

INTEL, IEDM 2007

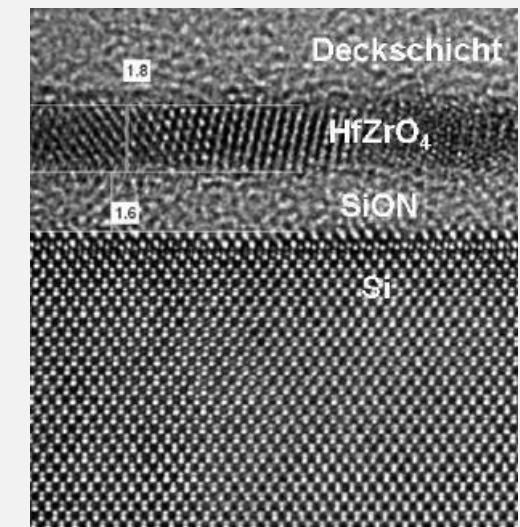
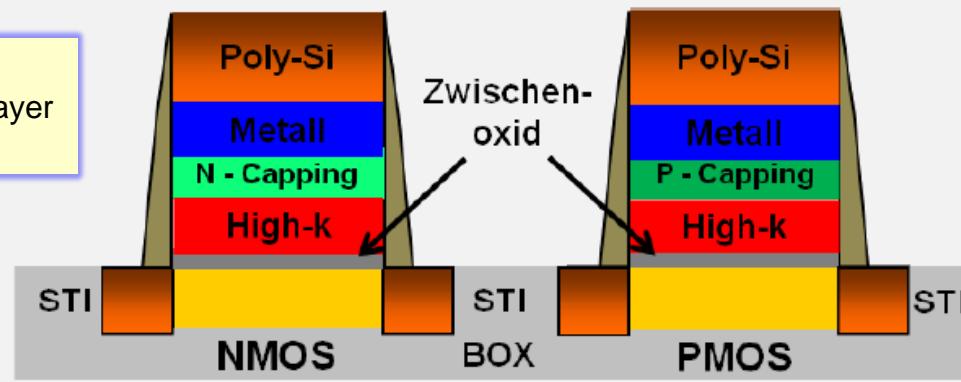


INTEL, 2012



Solution:

We use silicon substrate and create a thin interface layer of  $\text{SiO}_x$  to achieve good  $D_{it}$

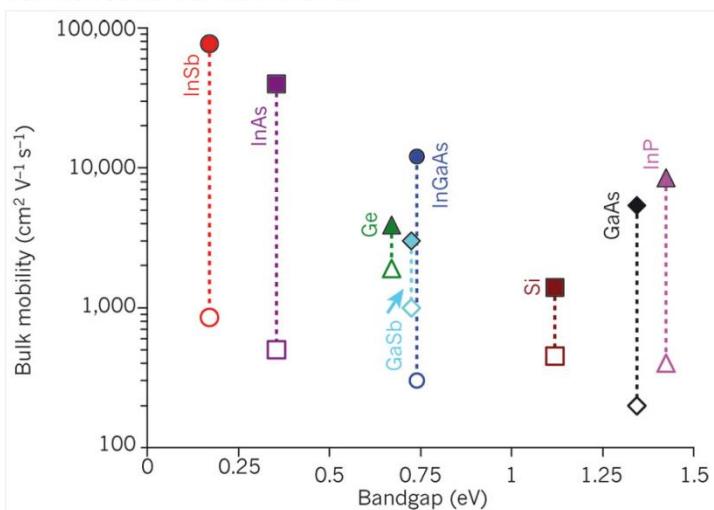


courtesy: AMD, Ph.D. thesis Kelwing 2014

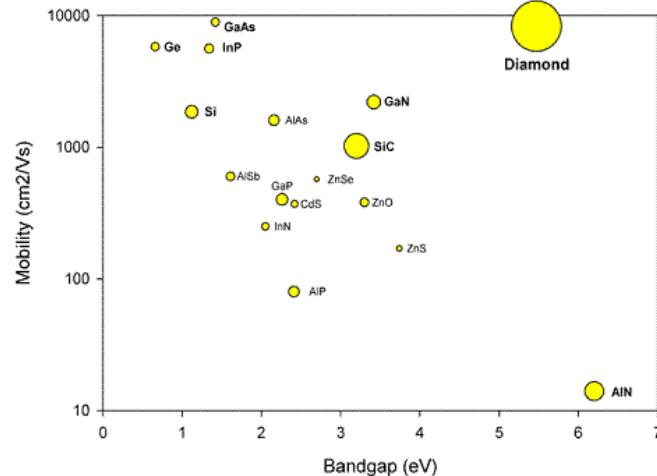
Figure 1: The mobility landscape of semiconductors.

From  
Academic and industry research progress in germanium nanodevices  
Ravi Pillarisetti

Nature 479, 324–328 (17 November 2011) | doi:10.1038/nature10678



The bulk mobility is plotted against the bandgap for silicon, germanium and a variety of group III-V materials. □ indicate holes. Germanium offers the highest hole mobility of any known semiconductor material.



Comparison of common semiconductor materials:  
x-axis: Bandgap; y-axis: Combined electron + hole mobility;  
Area of circles proportional to thermal conductivity.

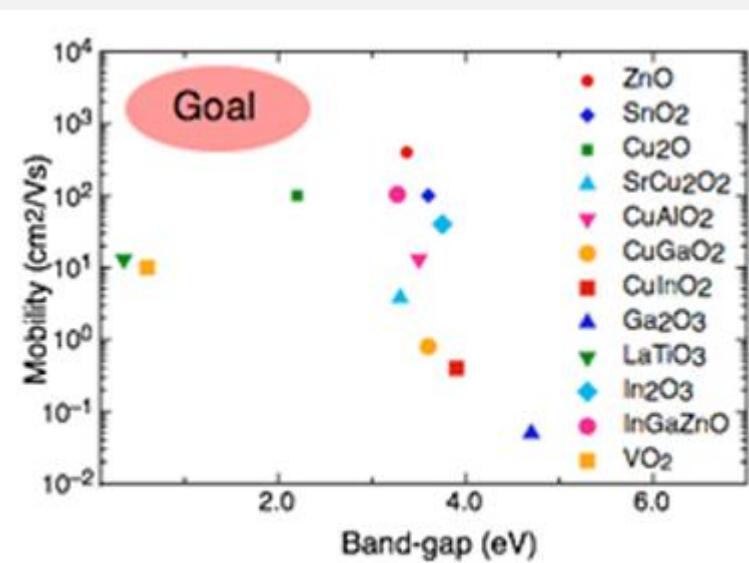


Figure 1 illustrates the reported carrier mobility versus the band gap for a whole range of oxide semiconductors.

YEAR OF PRODUCTION	2018	2020	2022	2025	2028	2031	2034
Logic industry "Node Range" Labeling (nm)	G54M36	G48M30	G45M24	G42M21	G40M16	G40M16T2	G40M16T4
IDM-Foundry node labeling	"7"	"5"	"3"	"2.1"	"1.5"	"1.0 eq"	"0.7 eq"
Logic device structure options	i10-f7	i7-f5	i5-f3	i3-f2.1	i2.1-f1.5	i1.5e-f1.0e	i1.0e-f0.7e
Mainstream device for logic	FinFET	finFET	finFET	LGAA	LGAA	LGAA-3D	LGAA-3D
LOGIC TECHNOLOGY ANCHORS							
Patterning technology inflection for Mx interconnect	193i, EUV	193i, EUV DP	193i, EUV DP	193i, High-NA EUV	193i, High-NA EUV	193i, High-NA EUV	193i, High-NA EUV
Beyond CMOS as complimentary to mainstream CMOS	-	-	-	2D Device, FeFET	2D Device, FeFET	2D Device, FeFET	2D Device, FeFET
Channel material technology inflection	Si	SiGe25%	SiGe50%	Ge, 2D Mat	Ge, 2D Mat	Ge, 2D Mat	Ge, 2D Mat
Process technology inflection	Conformal deposition	Conformal Doping, Contact	Channel, RMG	Stacked-device Non-Cu Mx	Stacked-device Non-Cu Mx	3DVLSI	3DVLSI
Stacking generation inflection	2D	2D	3D-stacking: W2W D2W	3D-stacking: W2W D2W	Fine-pitch + High-BW 3D stacking, P-over-N (CFET), VGAA use	3DVLSI: Mem-on-Logic, VGAA use	3DVLSI: Logic-on-Logic

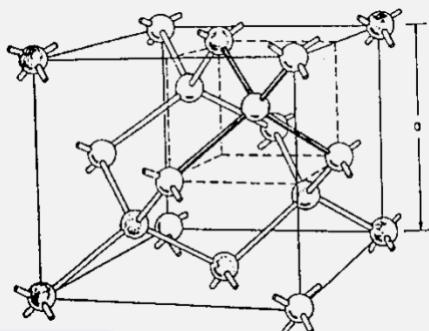
Roadmap suggests to use materials like Germanium and IIIV in the channel from 5nm node

For the moment it looks like problems must be solved

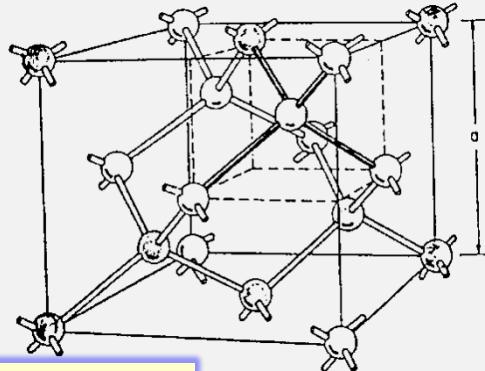
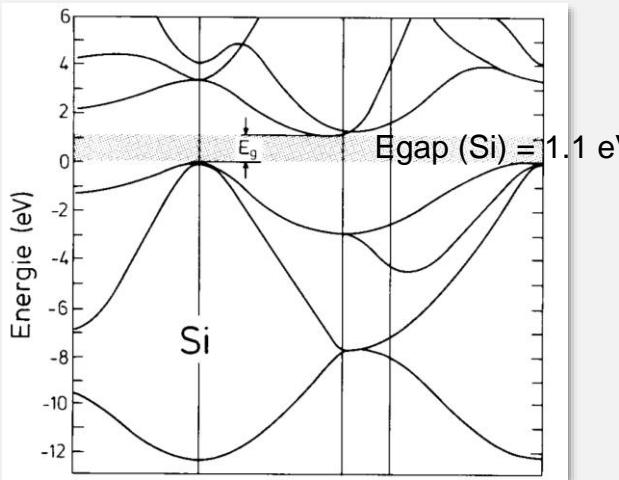


INTERNATIONAL ROADMAP FOR DEVICES AND SYSTEMS™  
INTERNATIONAL ROADMAP FOR DEVICES AND SYSTEMS™

2018 UPDATE

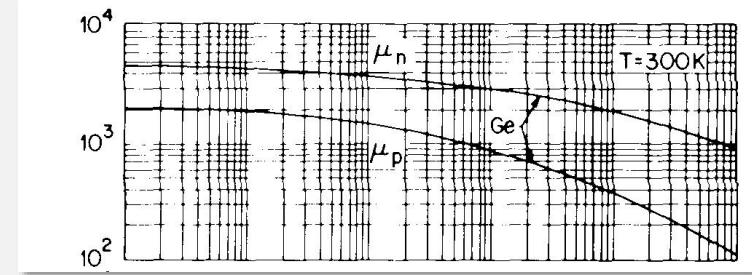
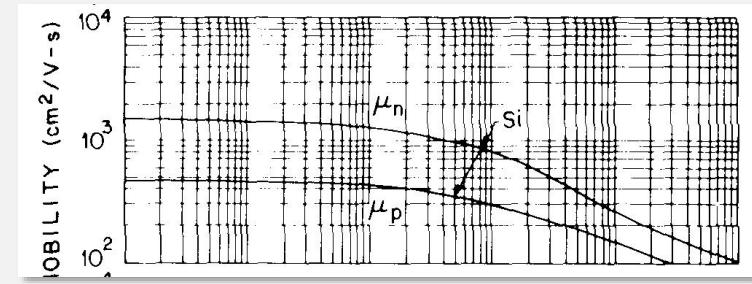
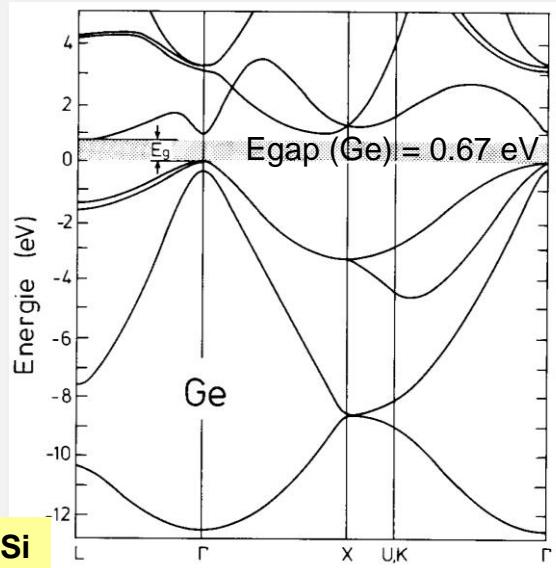


**Silicon**      cubic fcc lattice  
lattice constant:  $a_{\text{Si}} = 0.54 \text{ nm}$



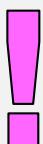
**Germanium**      cubic fcc lattice  
lattice constant:  $a_{\text{Ge}} = 0.56 \text{ nm}$

Ge lattice is about 4.2% larger than Si



**Mobility of electrons and holes 2 times higher in Ge**

So let's try to use Ge in the channel



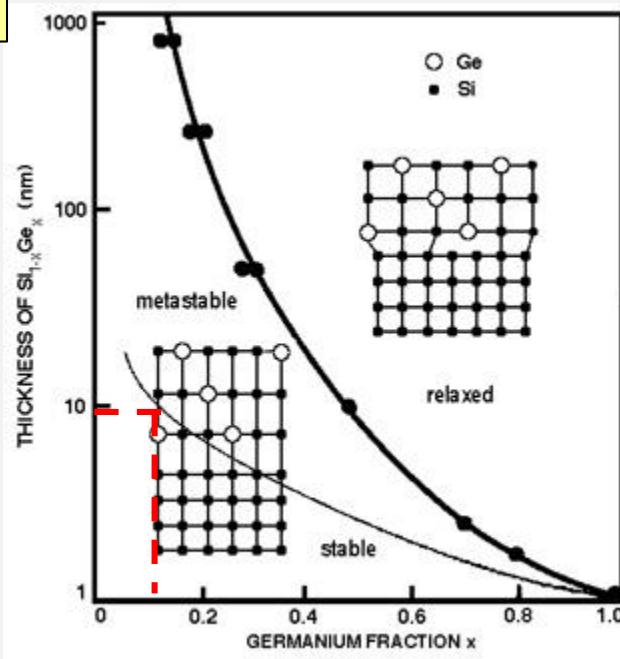
For defectfree devices the Ge must be grown epitaxially on Si-substrate.

-> No problem, since both elements have the same cubic lattice,  
but Ge lattice is 4.2 % larger

-> because Ge lattice is larger, with epitaxial growth strain is induced

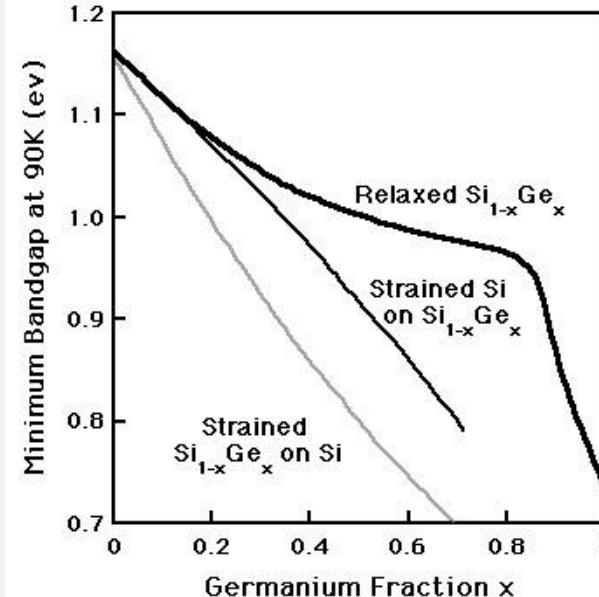
1

Ge or mixtures of Si-Ge  
can be grown epitaxially on Si



2

Mixing Si ( $E_{gap} = 1.1$  eV) and Ge ( $E_{gap} = 0.67$  eV) results in an intermediate bandgap of the mixture (Vegard's rule)



Dependent on the Ge-fraction a **critical thickness** limits the **pseudomorphic growth** of SiGe on Si

- > for a 10nm thick channel only 10 % Ge can be used
- > only 10-20 % improvement of mobility



using Ge in the channel an additional Si-cap is necessary to form MOS-oxide

In a Si bipolar transistor the current gain is determined by the thickness and doping of the base layer.

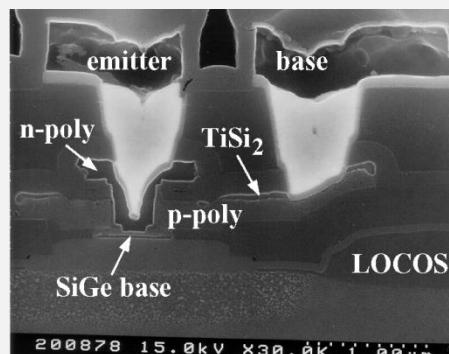
A trade-off has to be found for the thickness and doping to optimize the current gain:

- as thinner the base as higher the current gain, but as higher the resistance (decreasing current gain)
- as higher the doping as lower the resistance, but as higher the leakage current (tunneling)

In the **SiGe Hetero-Bipolar Transistor (HBT)**, Ge is selectively introduced into the base region of a Si-bipolar transistor

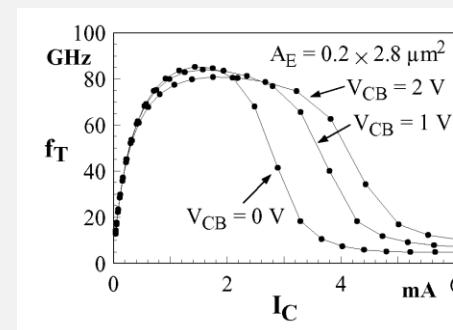
The bandgap of the SiGe-base arranges linearly between the bandgap of pure Si (1.12eV) and pure Ge (0.67eV) (called Vegard's law). The band-offset is almost in the valence band.

SiGe (HBT) technology:



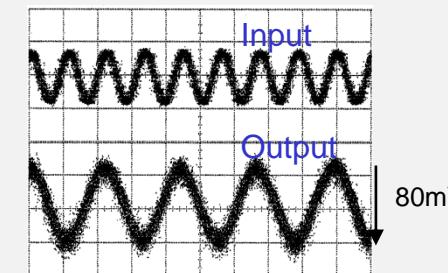
Source: Infineon, IEDM2000

Transistor HF-performance:



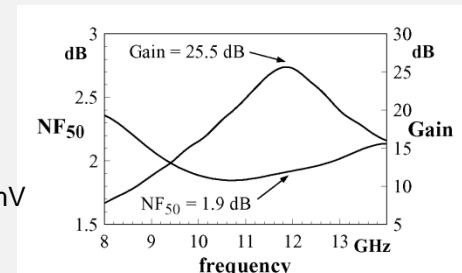
fmax: 128 GHz,  
FT : 85 GHz at 2mA

Frequency divider:  
(digital application)



88 GHz input,  
consumes 100mA at 6V

Low Noise Amplifier (LNA):  
(analogue application)



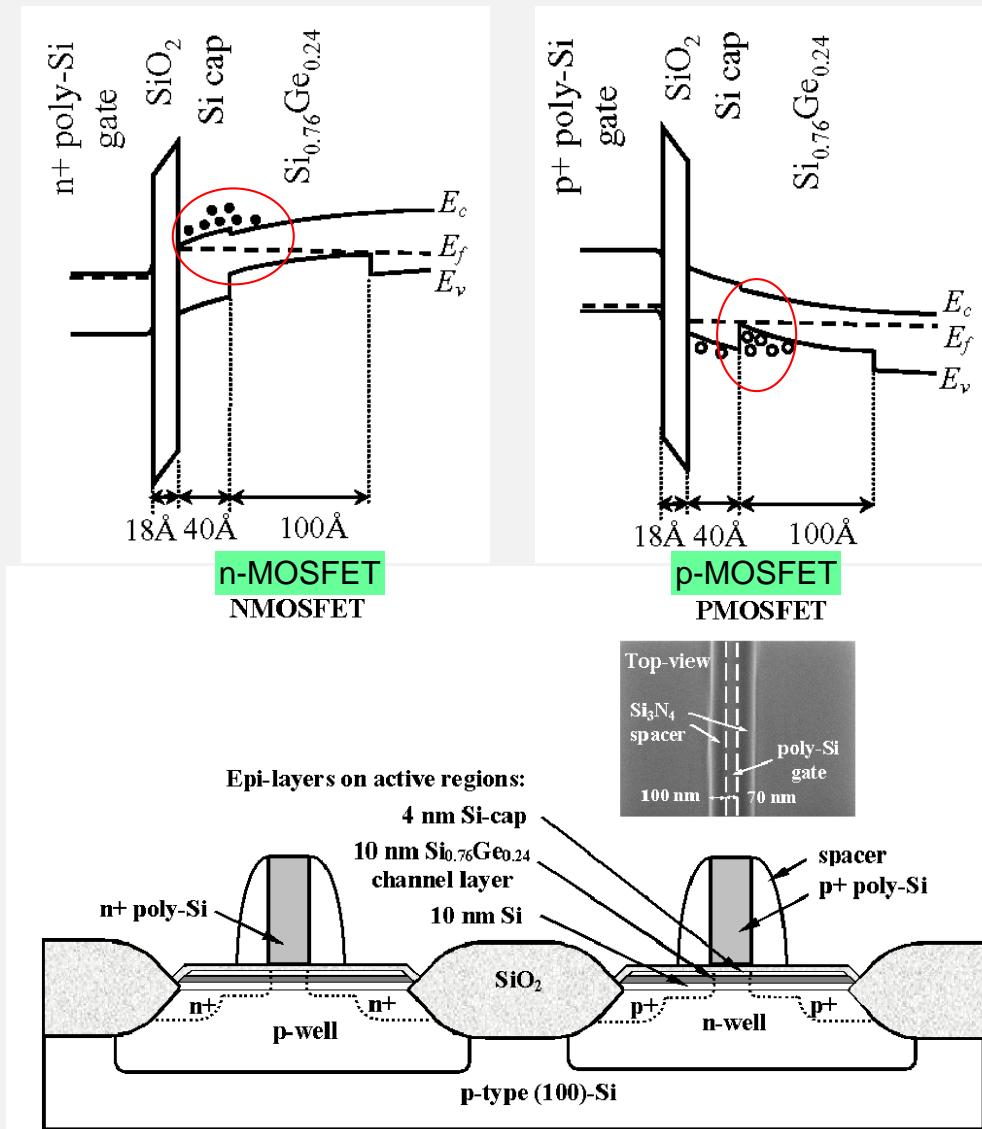
at 12 GHz:  
noise: 1.9dB, gain: 25dB

The smaller base bandgap enhances electron injection, producing a higher current gain compared to Si for the same base doping level.

In addition with varying vertical Ge-content in the base an intrinsic drift-field can be incorporated to improve the transistor frequency



SiGe is standard high performance Bipolar technology



The band offset in SiGe is used to improve mobility

but Si-cap needed for MOS

Band offset for SiGe is almost in the valence band  
-> for p-MOSFET most holes in the channel are confined to SiGe  
-> for n-MOSFET a fraction of channel electrons move in SiGe

Band offset and strain properties improve performance

CMOS with 100nm gate length (University research)

Source:UCLA, IEDM2000

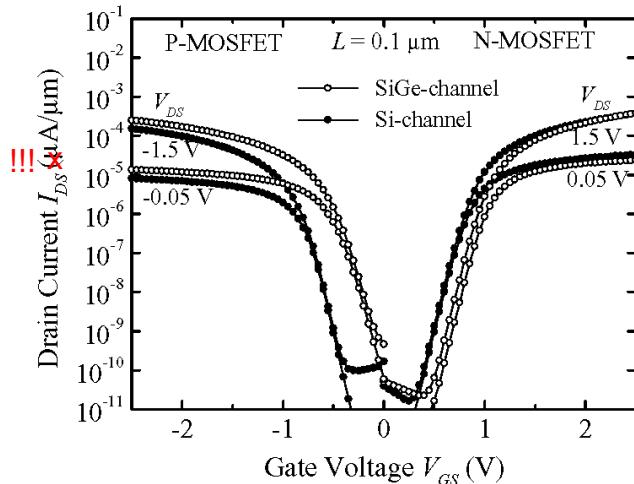
### Selective epitaxy with UHV-CVD

for Si<sub>0.76</sub>Ge<sub>0.24</sub> the critical thickness is 10nm  
S/D recrystallization: 950°C, 30sec

Epitaxy in CVD-reactors is standard CMOS technology.

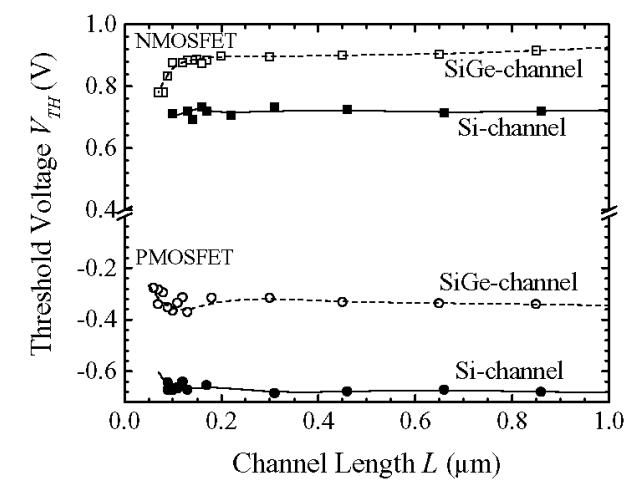
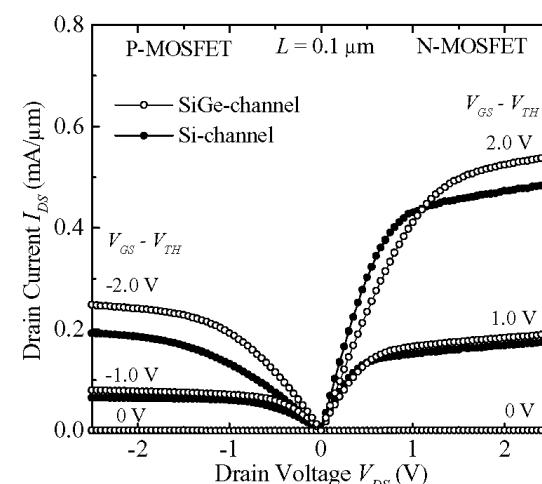
Introducing Ge in Si-CMOS improves performance  
for nearly no cost rise

Electrical properties of SiGe-MOSFETs (experimental state): Source:UCLA, IEDM2000

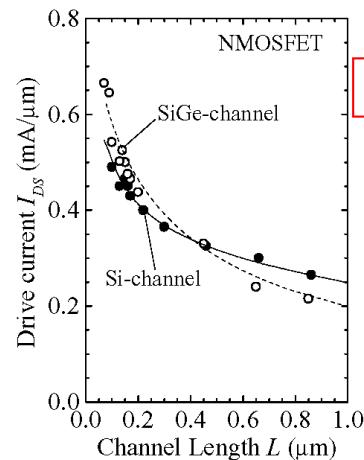
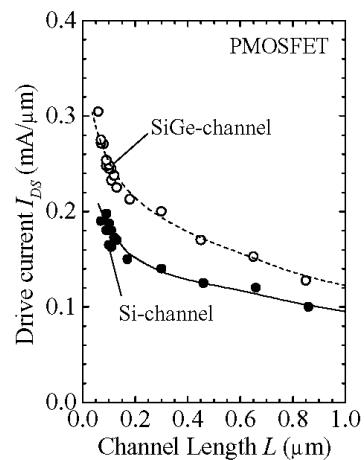


$I_{off} < 0.1 \text{nA}/\mu\text{m}$

S:  $\sim 100 \text{ mV}$   
DIBL:  $< 10 \text{ mV/V}$



Threshold voltage is shifted to positive values for SiGe



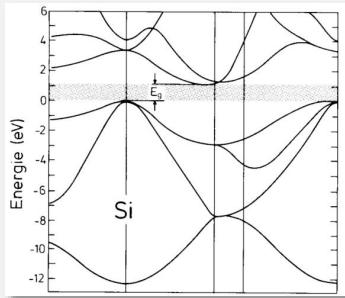
$I_{on}$  increased by 25% in p-MOSFETs with SiGe channel

-> Fabricating defectfree SiGe with additional Si-cap is not easy

-> Performance increase in mobility larger as expected just from Ge content



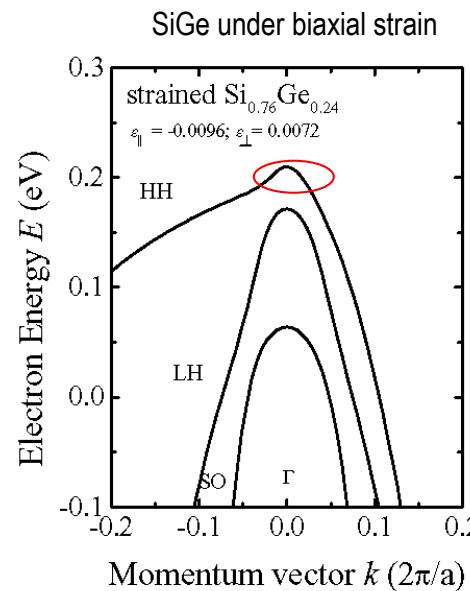
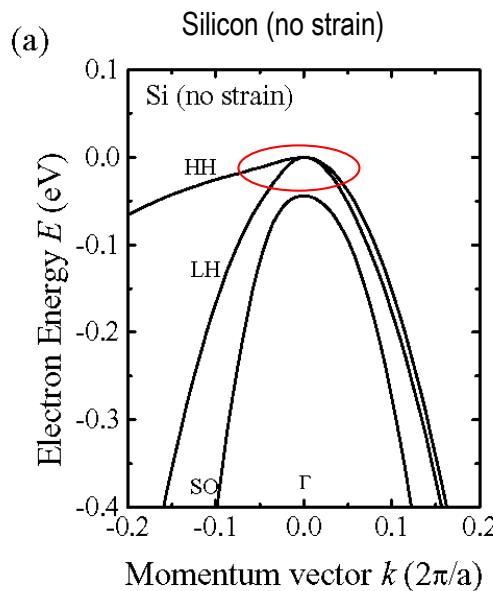
An additional increase in mobility is induced by strain



How can strain increases mobility ?

Not easy to answer, not completely understood

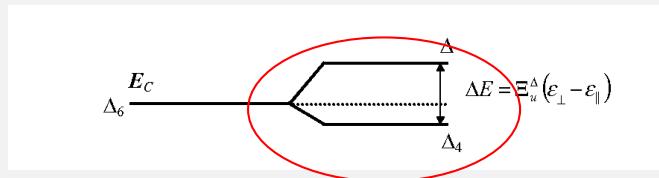
### Holes



The biaxial stress (lateral and vertical) breaks down the existing cubic lattice symmetry.

For holes in strained Si the top-most heavy hole mass is reduced  
-> higher mobility

### Electrons



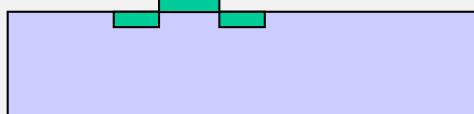
The six-fold degeneracy for electron masses is decoupled in 4 in-plane lower energy valleys and 2 perpendicular transport, higher states

-> higher mobility

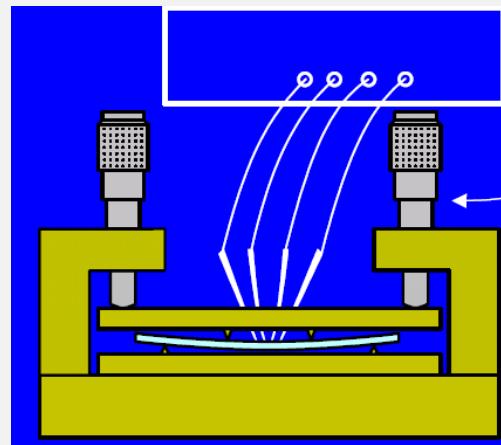
To create strain different approaches exist



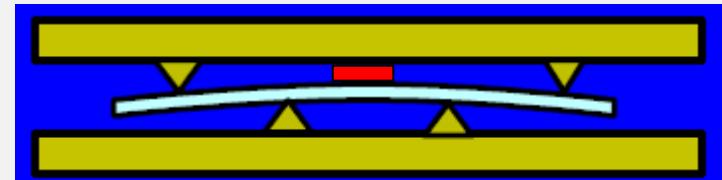
## Creation/Measurement of strain



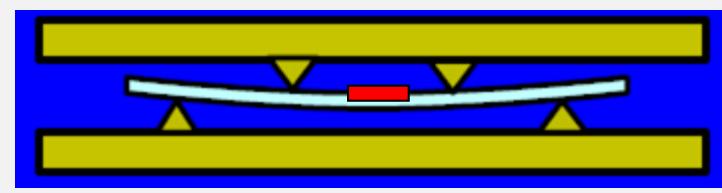
wafer with devices on top



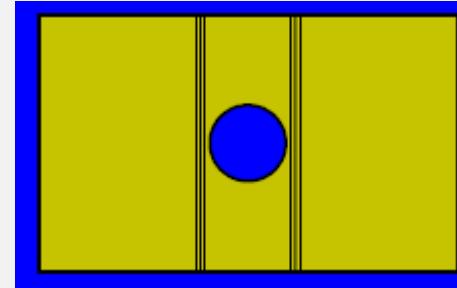
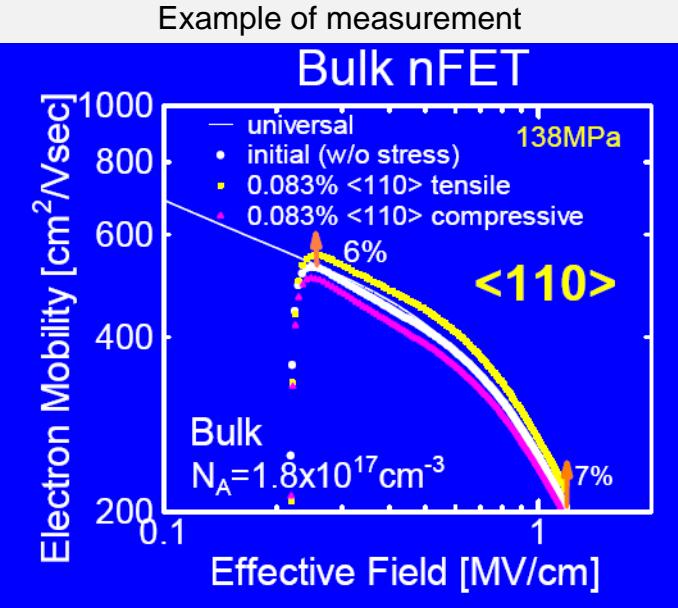
Depending on wafer and pit orientation tensile or compressive stress, and also uniaxial and biaxial stress can be created



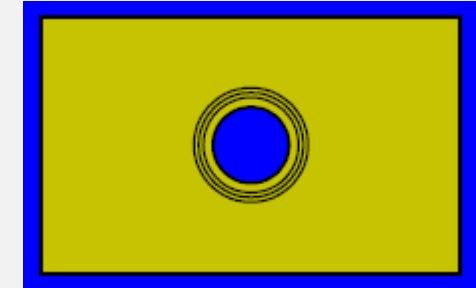
tensile stress



compressive stress



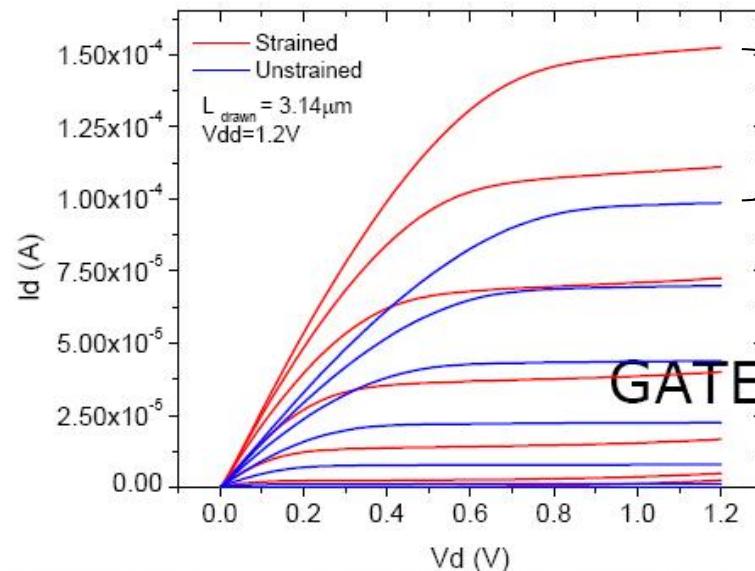
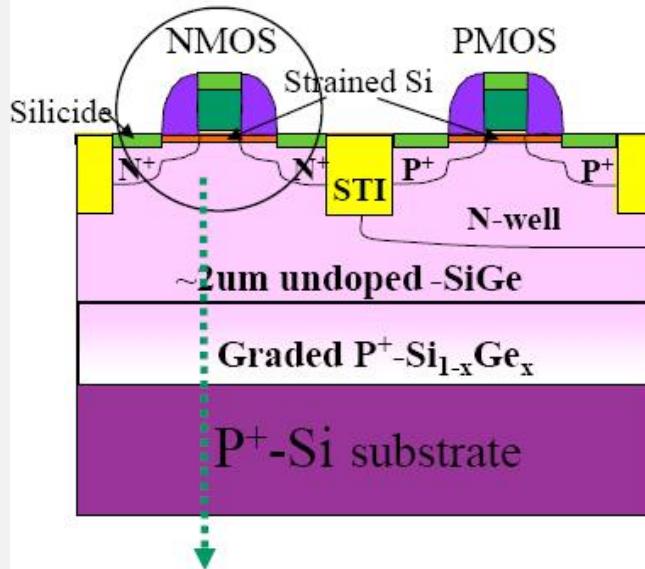
uniaxial stress



biaxial stress



Strain is induced on special fabricated Si substrates with relaxed SiGe-buffer



Advantage:

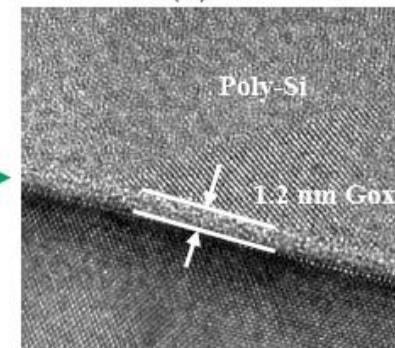
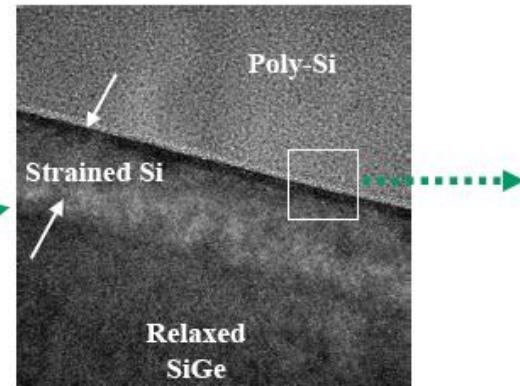
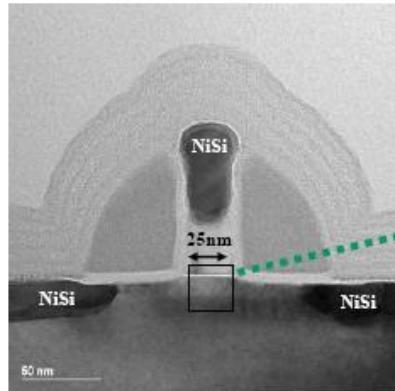
50 % current gain !

Disadvantage:

high sophisticated process for buffer growth

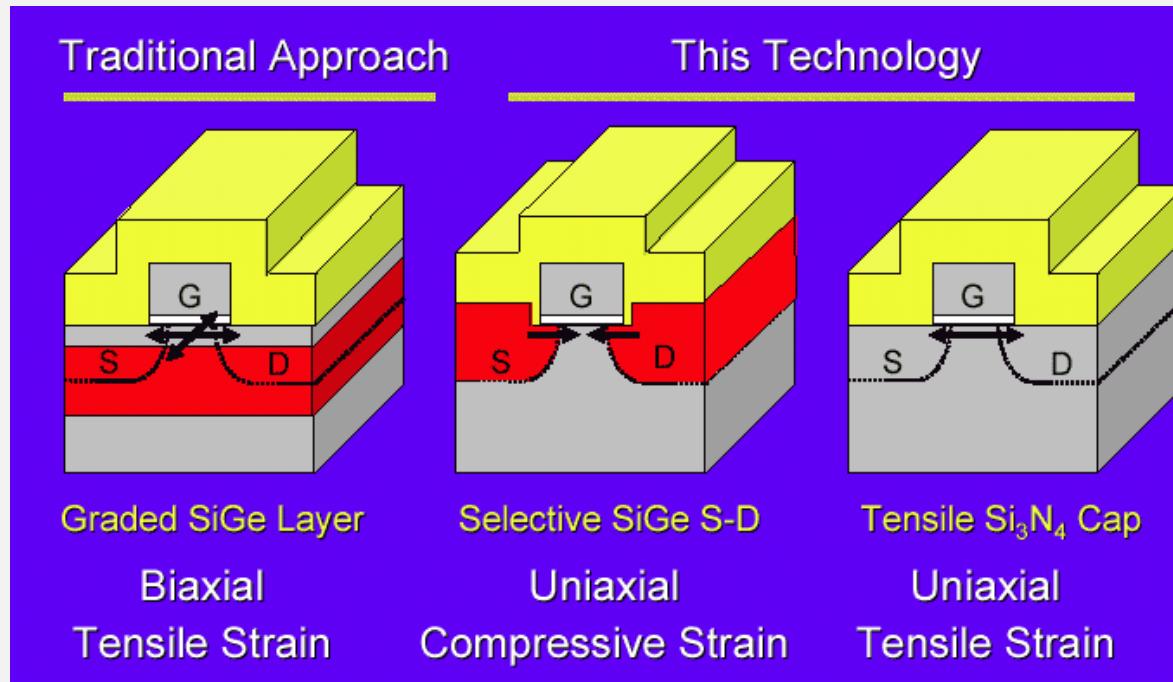
takes a long time > 3h

very expensive



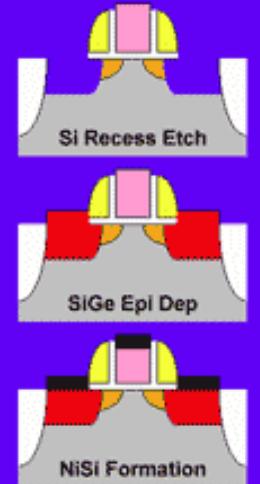
Strained substrate: 20% Ge; 175Å strained Si

Courtesy: AMD  
presented at Analyst Meeting 2002



### Strained PMOS Process Flow

- SiGe introduced late in the process flow → source-drain
- Si Recess Etch + SiGe Epi deposition inserted post spacer formation to standard non-strained process
- Ease of implementation



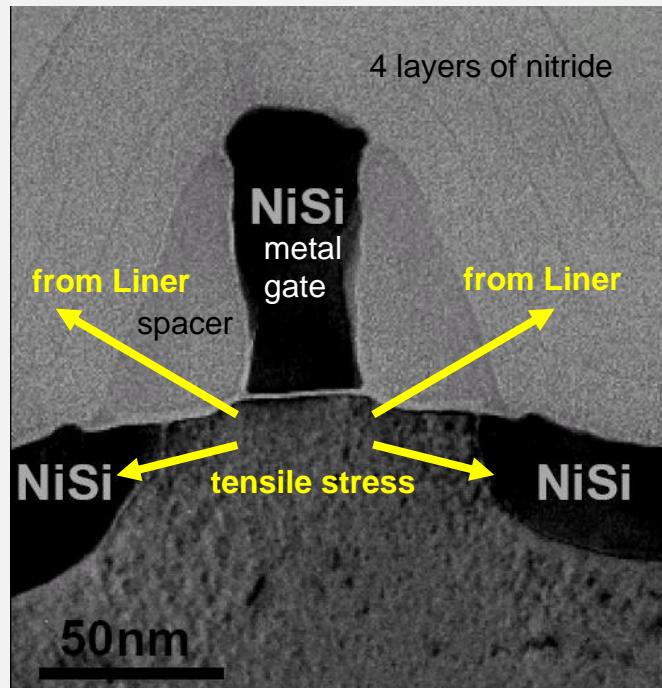
IEDM 2003

5

Other concepts: Induce strain from surroundings, not from substrate

### n-MOSFET

- \* introduce tensile strain in the channel

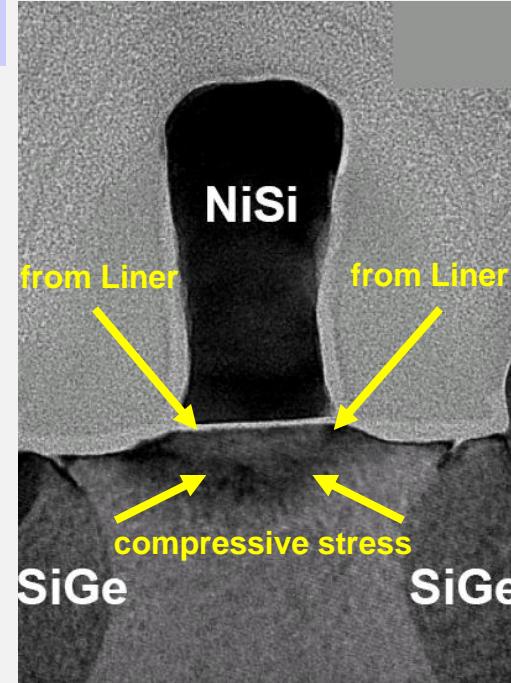


### INTEL CMOS:

65 nm node  
L = 35 nm  
tox (SiON) = 1.2 nm  
FUSI: NiSi

### p-MOSFET

- \* introduce compressive strain in the channel



tensile strain introduced by:

- \* 4 layers of nitride + NiSi<sub>2</sub> - S/D
- nitride: strain can be adjusted by process
- NiSi: volume reduction during formation

compressive strain introduced by:

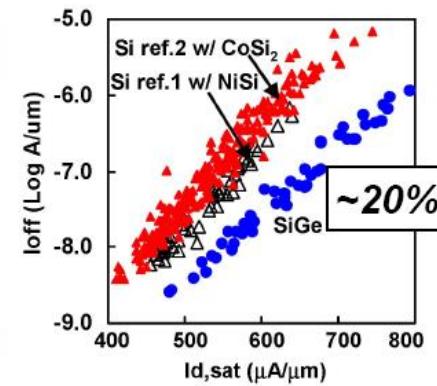
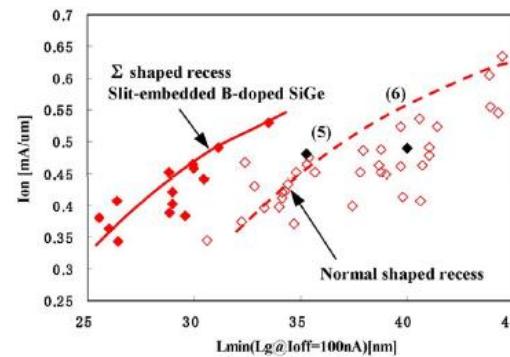
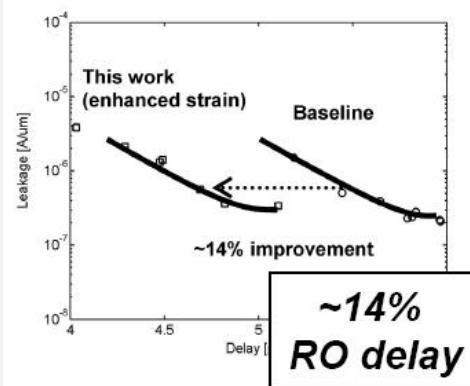
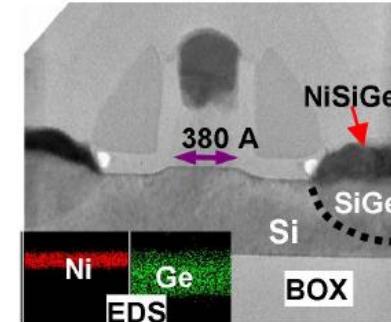
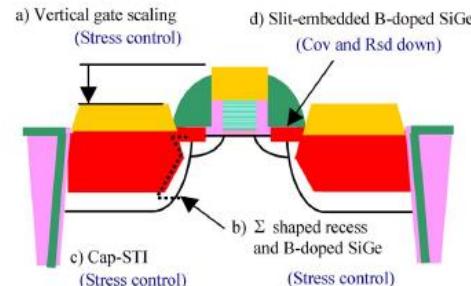
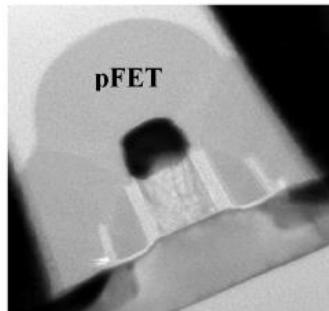
- \* SiGe
- Ge: larger atom than Si + advantage of higher doping

INTEL: IEDM 2005

<http://www.intel.com/technology/itj/2008/v12i2/1-transistors/6-stressenhancement.htm>

der Bundeswehr

# Embedded SiGe (PMOS)



Lee - IBM  
IEDM 2005 [38]  
SOI and e-SiGe

Ohta – Fujitsu  
IEDM 2005 [39]  
Profile engr.

Zhang – Freescale  
VLSI 2005 [40]  
Thin body SOI

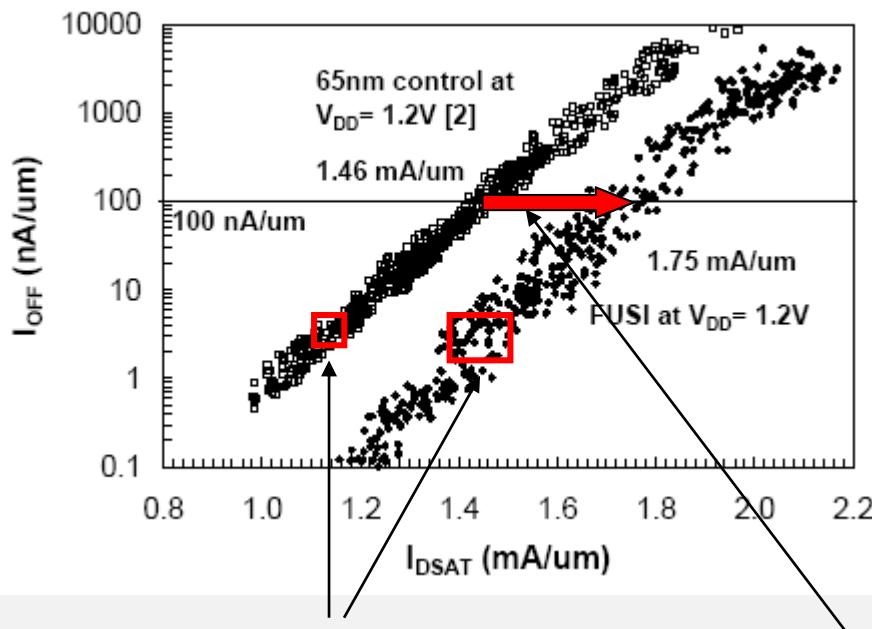


Kelin Kuhn / IWCE / Beijing / 2009

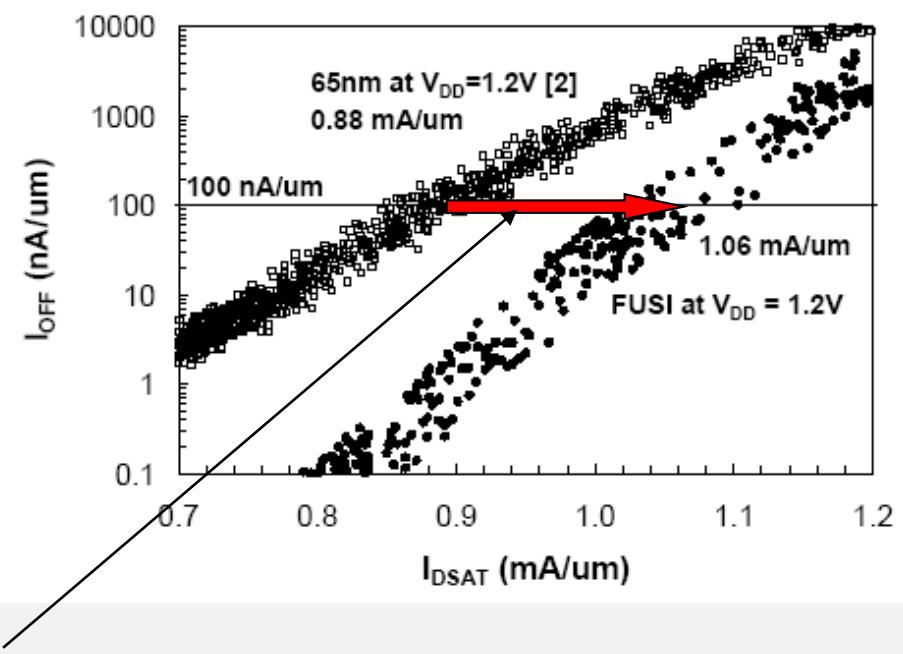
46

INTEL CMOS:  
65 nm node  
 $L = 35 \text{ nm}$   
 $\text{tox (SiON)} = 1.2 \text{ nm}$   
FUSI: NiSi

n-MOSFET



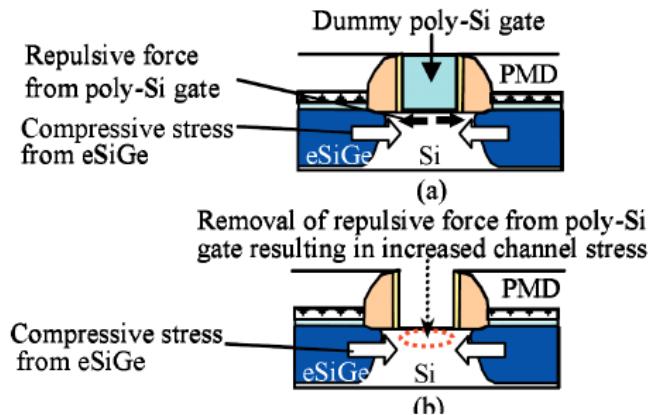
p-MOSFET



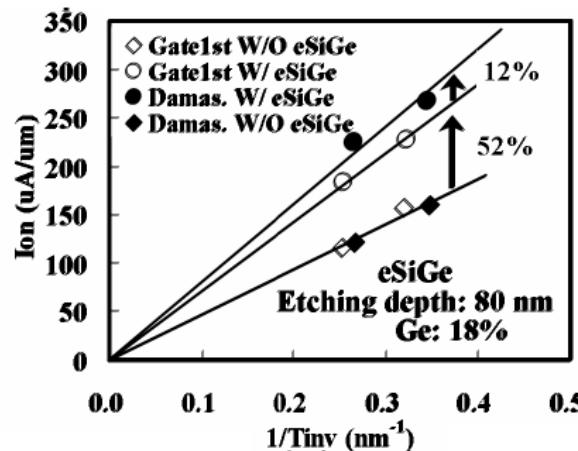
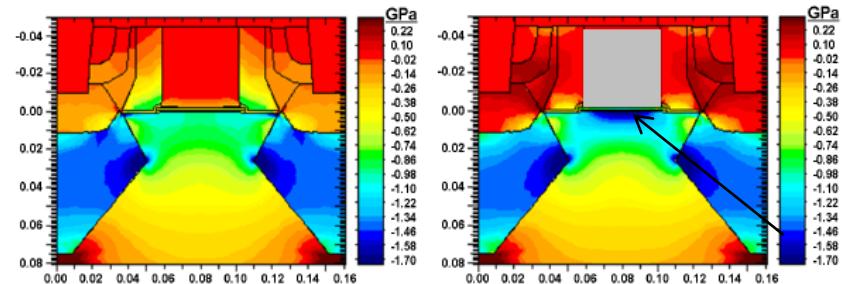
highest reported drive current  $I_{on}$  in industrial devices !

(but with the help of metal gate) -> see later

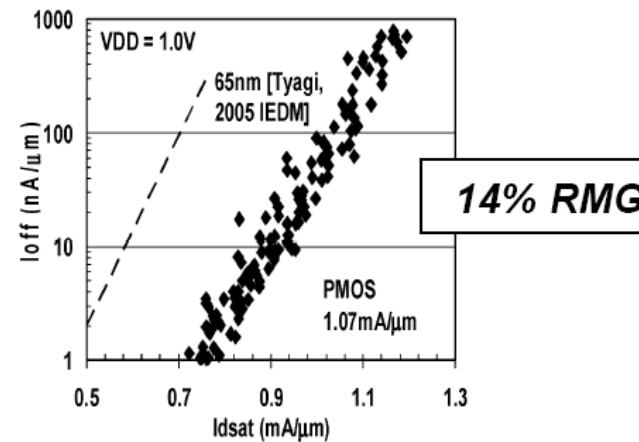
## Enhanced PMOS strain: Gate last HiK-MG



Before gate removal      After gate removal

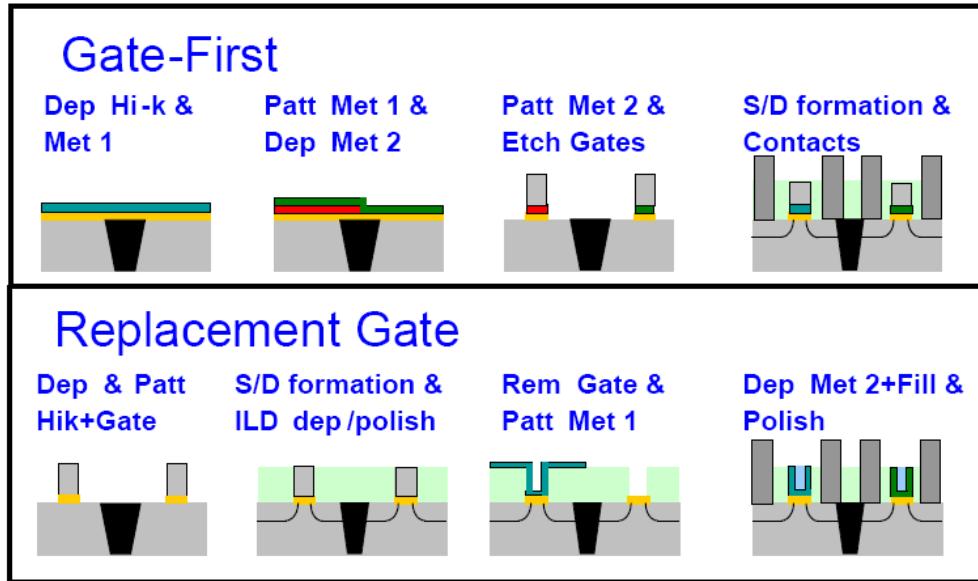


Wang – Sony  
VLSI 2007 [41]



Auth – Intel  
VLSI 2008 [23]

# Gate First vs Replacement Gate



## Advantages of replacement gate flow

- **High Thermal budget available for Midsection**
  - Better Activation of S/D Implants
- **Low thermal budget for Metal Gate**
  - Large range of Gate Materials available
- **Significant enhancement of strain**
  - Both NMOS and PMOS benefit

Auth - Intel –VLSI 2008 [23]

- STI, Wells, and VT Implants
- ALD deposition of high-k gate dielectric
- Polysilicon deposition and gate patterning
- S/D extensions, spacer, Si Recess & SiGe deposition
- S/D Formation, Ni Salicidation, ILD0 Deposition
- Poly Opening Polish, Poly Removal
- PMOS workfunction metal deposition
- Metal gate patterning, NMOS WF metal deposition
- Metal gate fill and polish, ESL deposition

Fig.5 Transistor process flow highlighting differences from [6,7]

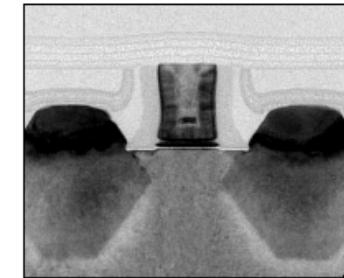


Fig.6 TEM micrograph of high-k + metal gate PMOS transistor.

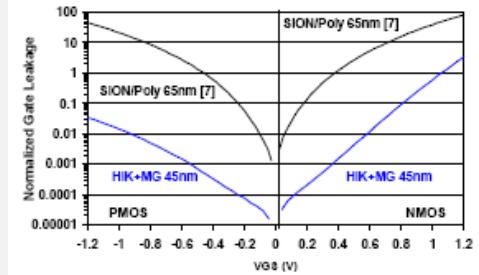


Fig.7 HiK+MG enables 25-1000X gate leakage reduction.

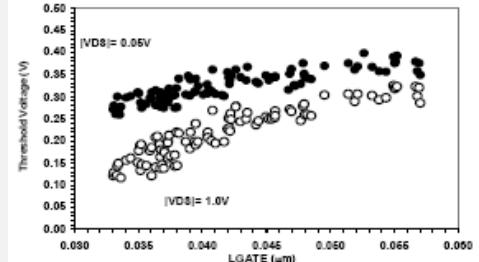


Fig.8 NMOS VT vs. LGATE showing good VT rolloff & DIBL

Jan 2010 Gate-first

## Samsung Foundry: 32/28nm LP High-K Metal Gate Logic Process and Design Ecosystem

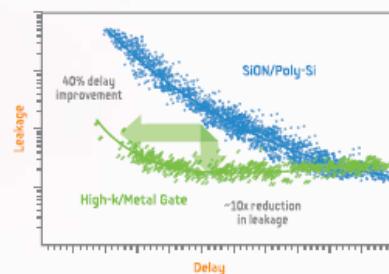


### Advanced Manufacturing for Energy-Efficient, High-Performance Mobile and IT Infrastructure ICs

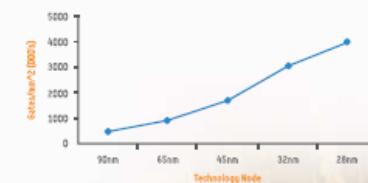
Samsung Foundry's 32/28nm LP gate first High-k Metal Gate (HKMG) process node offers considerable power/performance advantages to a growing spectrum of mobile and IT infrastructure computing applications.

Virtually all mobile applications relying on batteries and extremely tight power budgets demand 32/28nm LP solutions. But, low power solutions are no longer limited to mobile applications. Reducing power is now a concern across a wide set of applications such as communications, networking, servers, and data centers.

#### HKMG ENABLES PERFORMANCE BOOST



#### GATE DENSITY



#### Advances in Material Science

Keeping Moore's Law alive, HKMG reduces transistor level leakage current down to extremely low levels, and permits greater technology scalability down to more advanced process node generations. With leakage becoming an issue, 45nm LP hit the proverbial wall, and as a result, ushered in the need for new material science advancements.

The 32/28nm LP HKMG process node displays an array of benefits and improvements over 45nm LP, including:

- 2x gate density increase (only possible with gate first HKMG)
- >x100 lower gate leakage
- >40% delay improvement at fixed leakage
- X10 leakage reduction at fixed speed

#### Chipfertigung

##### Gerüchte um Probleme mit High-K/Metal-Gate-Prozess

Für Wirbel sorgte ein Bericht der Barclays Bank, in welchem ein Analyst von angeblichen Problemen bei der Chipfertigung im Gate-First-Ansatz berichtet. Dieser wurde im Gegensatz zu Intel und TSMC von IBMs Fab-Club gewählt, betroffen wären also auch die beiden Auftragsfertiger Globalfoundries und Samsung.

Anzeige Sep 2010: problems with Gate-first

Unbestreitbar ist, dass der Gate-First-Ansatz dem Gate-Last-Ansatz bei 28-nm-HKMG-Lösungen sowohl bei der Skalierbarkeit als auch bei der Herstellbarkeit überlegen ist, da er erheblich geringere Chipgrößen und Kosten und ebenso eine Kompatibilität mit bewährten Konstruktionselementen und Prozessabläufen vorheriger Technologieknoten bietet. Im Rahmen der Siliziumplanartechnologie wird das Gate standardmäßig als erstes platziert (Gate-First). Anhand der Gate-Position können Source und Drain ohne weiteres korrekt platziert werden.

Das Problem besteht darin, dass die Dotierung der Source/Drain-Gebiete bei hohen Temperaturen erfolgt, die zwar dem bisher genutzten Polysilizium-Gate nichts ausmachen, einem Metal-Gate jedoch Probleme bereiten können.

Laut dem Bericht der Barclays Bank sollen die Mitglieder von IBMs Fab-Club und der Fishkill-Allianz (IBM, Infineon, NEC, Globalfoundries, Samsung, ST Microelectronics und Toshiba) auf Probleme mit dem Gate-First-Ansatz gestossen sein: Neben thermischen Instabilitäten soll es auch Verschiebungen der Transistor-Schwellenspannungen gegeben haben. Die Schlussfolgerung des Berichtes besteht darin, dass TSMC, das wie Intel auf dem Gate-Last-Ansatz setzt, bei dem das Gate erst nach den Hochtemperaturschritten eingesetzt wird, mit seinem 28-nm-Prozess schneller am Markt sein als die Mitbewerber.

Sowohl Globalfoundries als auch Samsung haben den Bericht zurückgewiesen. Gregg Bartlett, Vice President bei Globalfoundries, fuhrte aus, dass man in Fab 1 bereits Testchips fertigt und Bestellungen für 28-nm-Designs annehme. Zahlreiche Designs seien bereits erfolgreich validiert worden.

##### Verliert TSMC Umsatz?

In dem Bericht der Barlays Bank wird auch auf das Geschäft der weltgrößten Foundry TSMC eingegangen. Hier erwartet man eine Auslastungsreduktion auf 75 Prozent, nachdem sie in Q2/Q3 2010 über einhundert Prozent gelegen hat. Bereits im 4. Quartal 2010 soll sie um bis zu 5 Prozent sinken, im ersten Quartal 2011 dann noch mal um 15 bis 20 Prozent. Entgegen TSMCs eigener Prognose für 2011, die ein Wachstum von 10 Prozent beinhaltet, nimmt die Barlays Bank 2011 einen Umsatzrückgang von 5 Prozent gegenüber 2010 an. Hier kann TSMC gegenüber dem Krisenjahr 2009 um 40 Prozent zulegen.

Die Investitionen für 2011 werden nach dem Bericht voraussichtlich auch auf 5 Mrd. Dollar zurückgefahren (geplant waren 5,9 Mrd. Dollar), allerdings erwartet man auf Grund des harten Wettbewerbs mit Globalfoundries und Samsung im Bereich der 28-nm-Fertigung keine Budgetreduktion.

Jan 20, 2011 13:16 GMT · By Sorin Nita · Share: 8+

**Yesterday, Common Platform members announced that the coalition will move to gate-last technology when it makes the transition to the 20nm manufacturing node, plans for 32nm and 28nm chip fabrication remaining unchanged at this time.**

The Common Platform Initiative is made out of IBM, GlobalFoundries, and Samsung Electronics, some of the world's largest semiconductor manufacturers, in an effort to introduce common process technology across all of their production facilities.

This helps in reducing research and development costs as the three companies can share those expenses between them.

Up until now, the companies were some of the most vehement supporters of the gate-first approach, GlobalFoundries reaffirming its support for the technology more than just a few times.

However, its position has now changed as the company decided to move to gate-last with their 20nm node.

The difference between the two technologies lies in the timing when the metal electrode is deposited, before or after the high temperature activation anneal(s) of the flow, gate-first, at least in theory, allowing customers to transition to a lower manufacturing node without having to redesign their chips.

According to Jon Carvill, VP of GlobalFoundries Communications, the change "is due to the design rules and desired scaling our customers want at 20nm. Gate Last was the best option to hit it and Gate First was the better option at 32/28nm."

"We have multiple customers engaged on Gate First 32nm/28nm as does Samsung (who made this disclosure with us yesterday)," concluded the company's rep.

Until Common Platform announced the switch to gate-last, Intel and TSMC were the major proponents of this technology.

# SolidState TECHNOLOGY®

Insights for Electronics Manufacturing

HOME SEMICONDUCTORS PACKAGING MATERIALS MEMS LEDS DISPLAYS SEMICON DESIGN

## Chipworks

0 0 0

## Common Platform Goes Gate-Last – at Last!

At the IBM/GLOBALFOUNDRIES/Samsung Common Platform Technology Forum on Tuesday, Gary Patton of IBM announced that the Platform would be moving to a gate-last high-k / metal-gate (HKMG) technology at the 20-nm node.

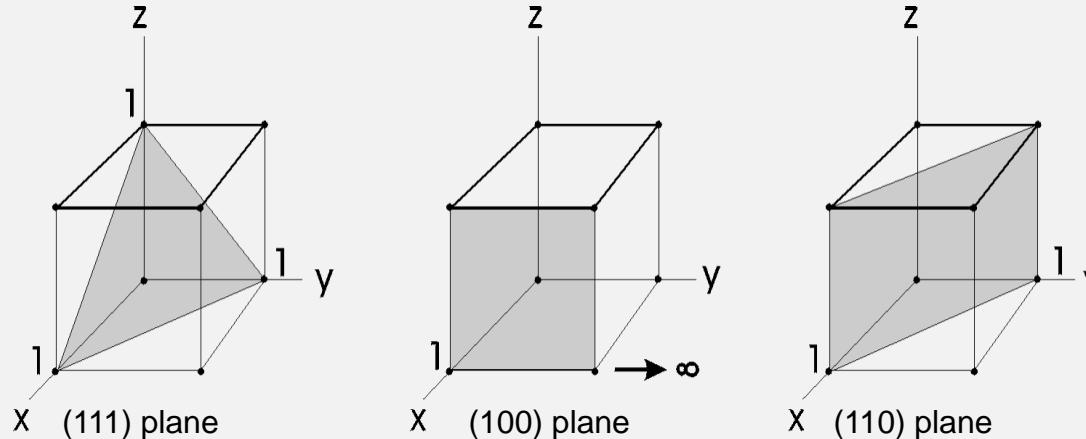
At the 45- and 32-nm nodes there has been a dichotomy between gate-last as embodied by Intel, TSMC, and UMC, and gate-first, promoted by the Common Platform and others such as Panasonic. (Though, to be realistic, Intel's is the *only* HKMG we've seen so far, and the only 32-nm product.)



Since the 20 nm technology node all companies are using Gate-Last/Replacement-Gate

Silicon exhibits a cubic lattice

Orientation of crystal planes and directions can be defined by intersection points with the unit cell axis



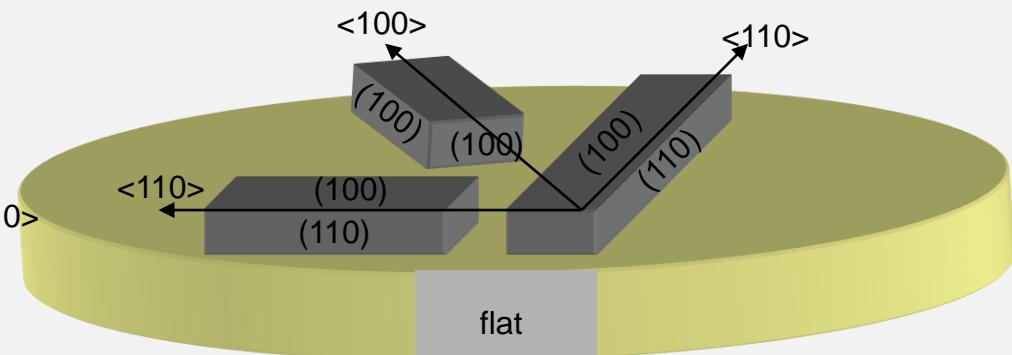
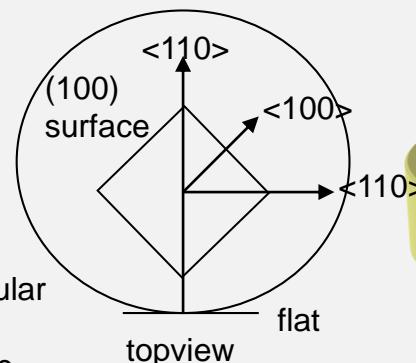
Historically, for CMOS-Technology all wafers are fabricated with:

- (100) surface

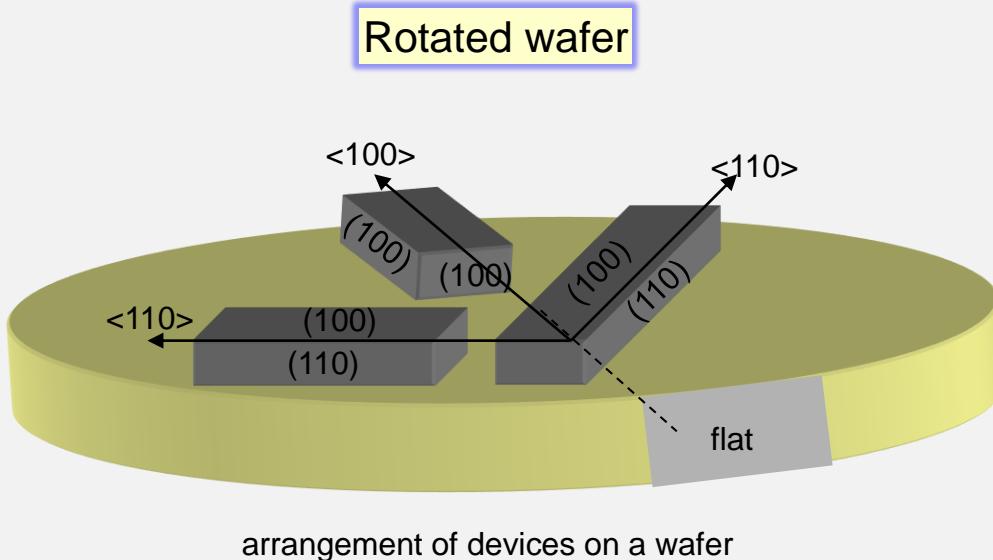
(best interface Si/SiO<sub>2</sub> with lowest interface state density)

- <110> channel direction

- closest package of devices rectangular
- best lithography rectangular
- best orientation for electron and hole mobility



arrangement of devices on a wafer

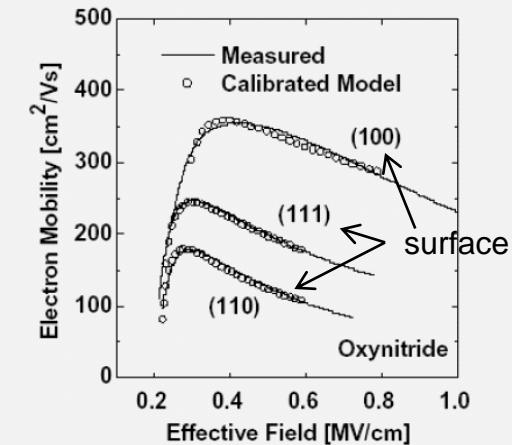


Since ~2004 various companies (INTEL, AMD, IBM,...), which are fabricating high-performance MPUs, are using (100)-wafers, but with notch in <100> direction (= rotated wafers) to use existing equipment. Arranging device channels in <100> direction may increase hole mobility of about 15% without decreasing electron mobility.

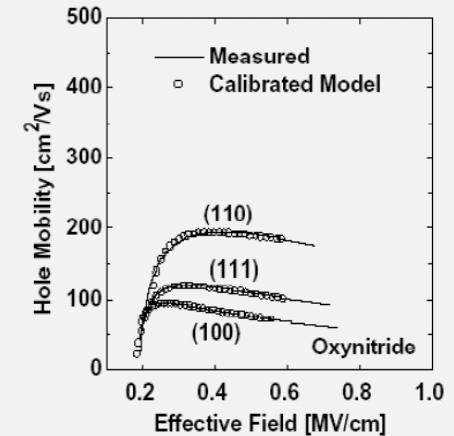
Combined technology  
+ strain

We need to have detailed knowledge of bandstructure

**New surface orientation**



**Hole Mobility**  
in <110> direction

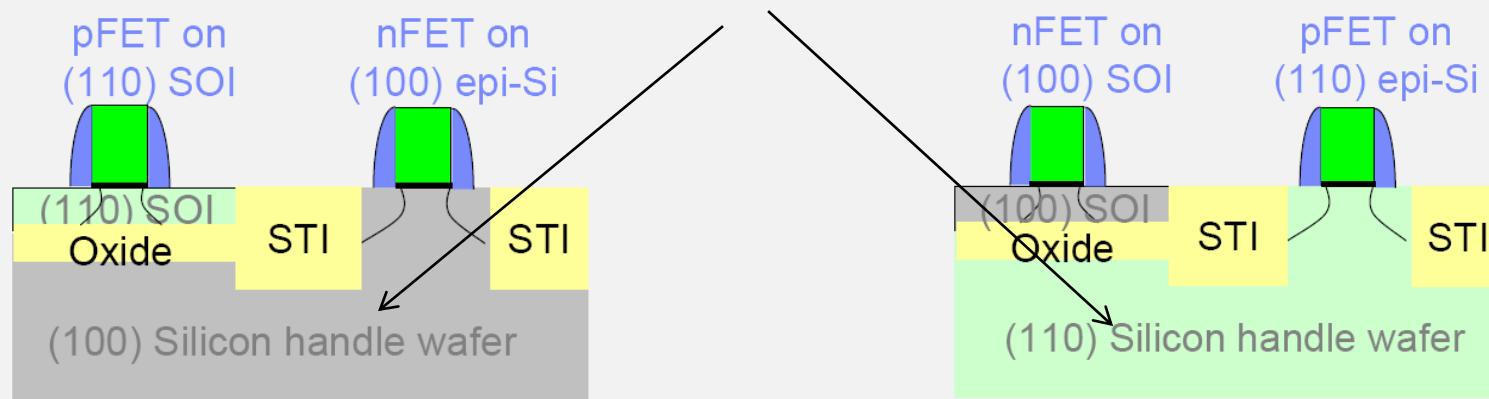


### Hybrid Orientation Technology (HOT)

Due to higher mobilities of holes (p-MOSFET) on (110)-surface, but higher mobility of electrons (n-MOSFETs) on (100)-surface, each type of MOSFET should have its own surface orientation on the same substrate (handling wafer).

*first proposed by IBM at IEDM 2003*

Both types of handling wafers are possible, but today best developed wafers (with 300mm) are (100)-wafers



Advantage:

- + well established (100)-wafer in 300mm (cheap)
- + low temperature wafer fabrication
- + well established (100) selective epitaxy

Disadvantage:

- non-established (110)-wafer bonding
- n-Fet in substrate requires additional technology steps (halo) to avoid SCE

Advantage:

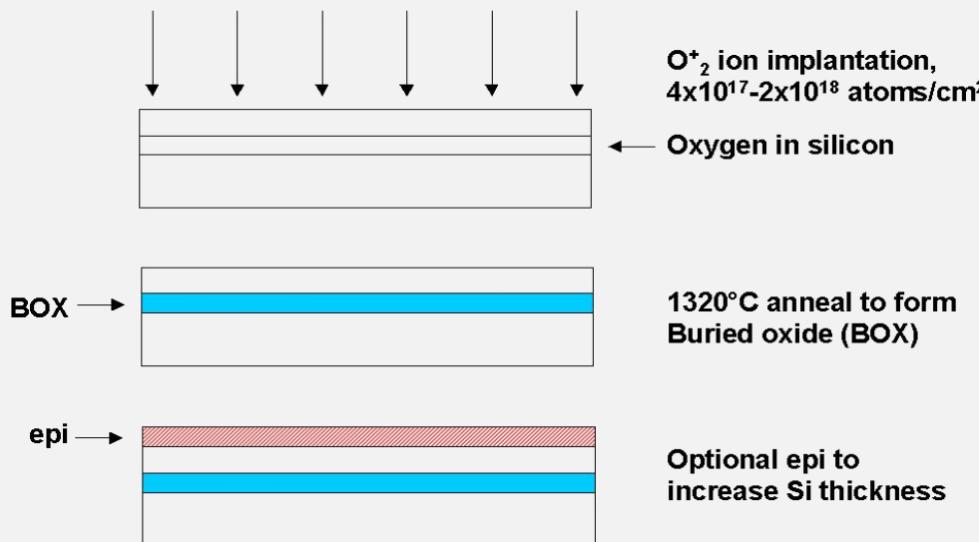
- + better performance with n-FET on SOI and p-FET in substrate

Disadvantage:

- high-temperature wafer bonding
- difficult (110) selective epitaxy

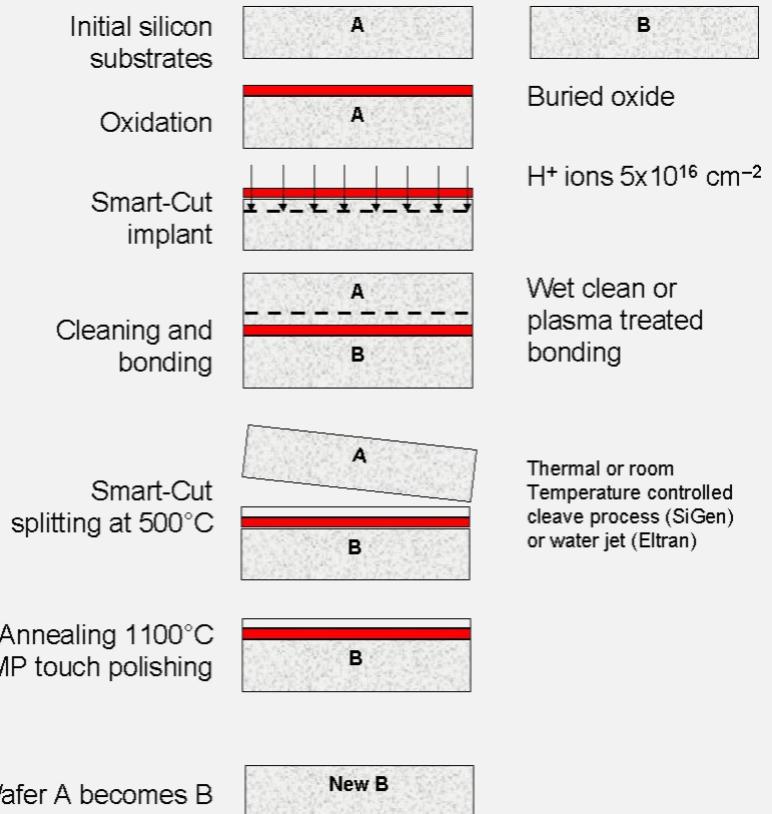
Usual fabrication of SOI-wafer

SIMOX (Separation by implanted oxygen)



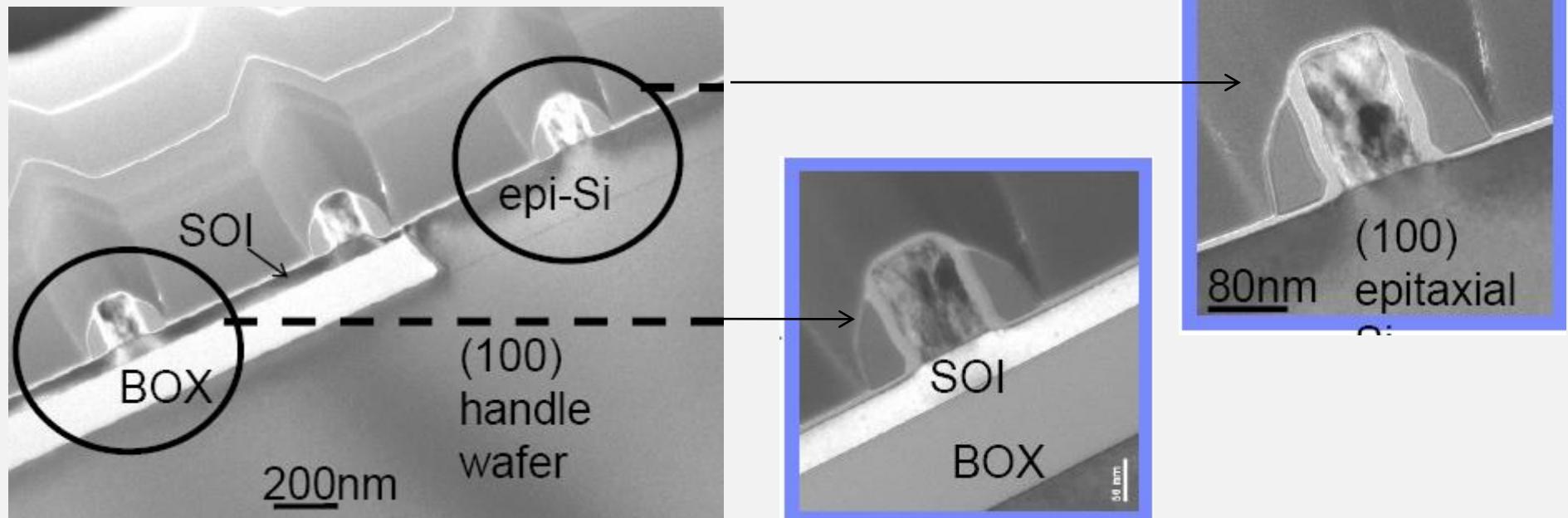
Wafer Bonding

Smart Cut

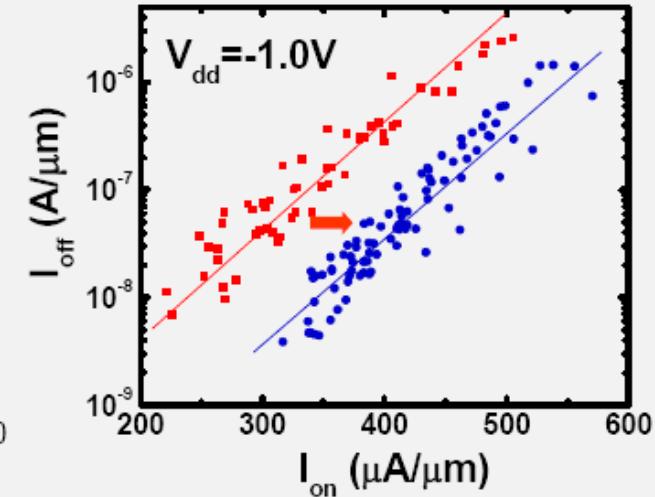
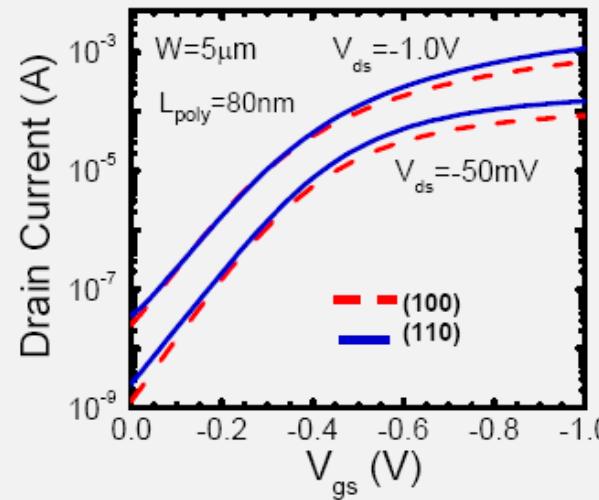
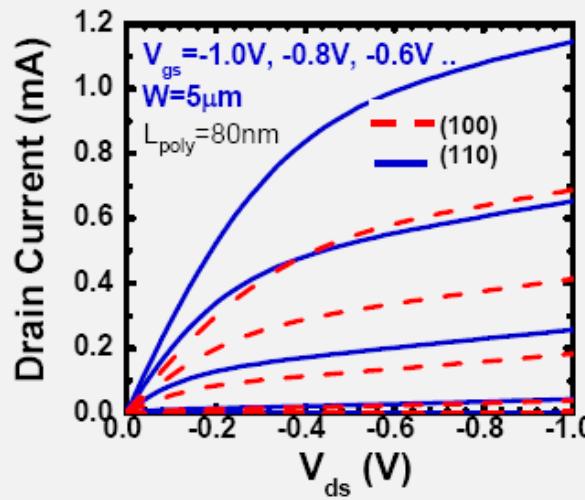


Basic technology steps:

1. wafer bonding (100) and (110)
2. reduce SOI thickness to ~50nm (CMP)
3. open BOX down to handling wafer
4. selective epitaxy
5. conventional CMOS fabrication



# pFET Performance Enhancement for HOT

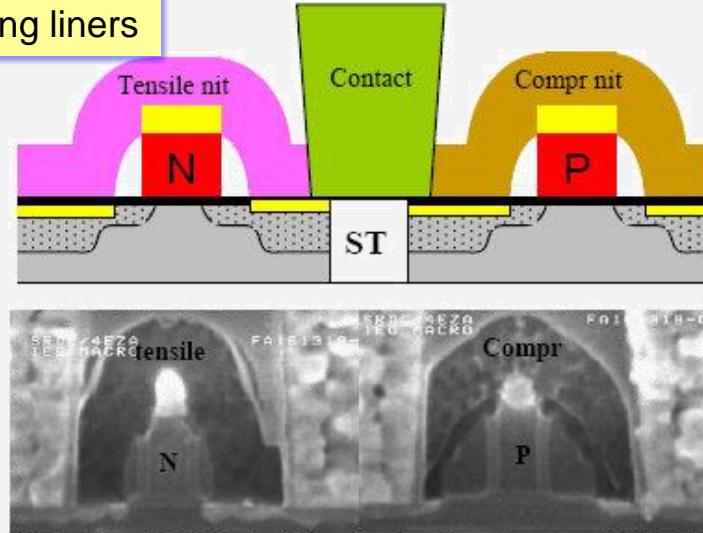


	$I_{on}$	$I_{dlin}$
$I_{off} = 100\text{nA}/\mu\text{m}$	+33%	+45%
$I_{off} = 10\text{nA}/\mu\text{m}$	+44%	+58%

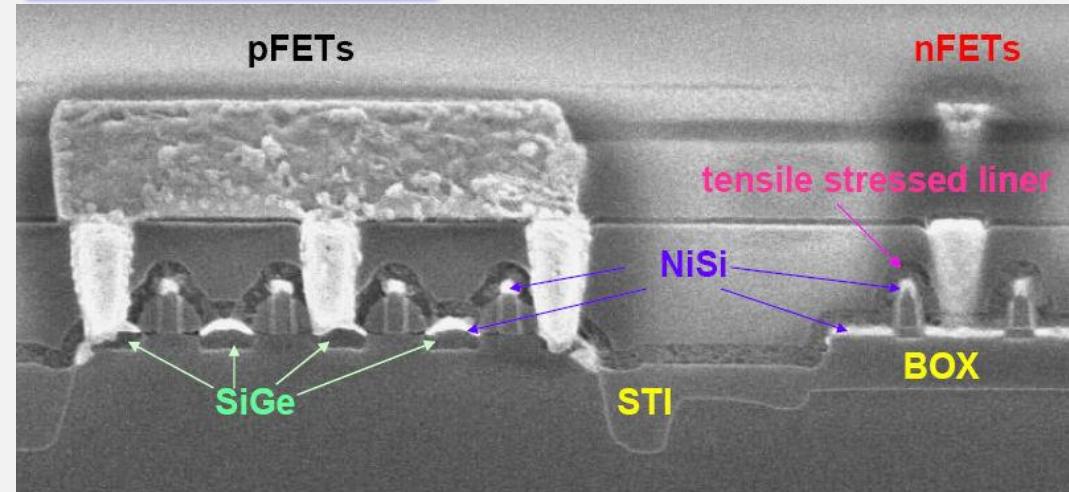
M. Yang et al., IEDM 2003

Depending on company's technology experience and risk willingness various combinations of the mobility enhancement techniques are possible:

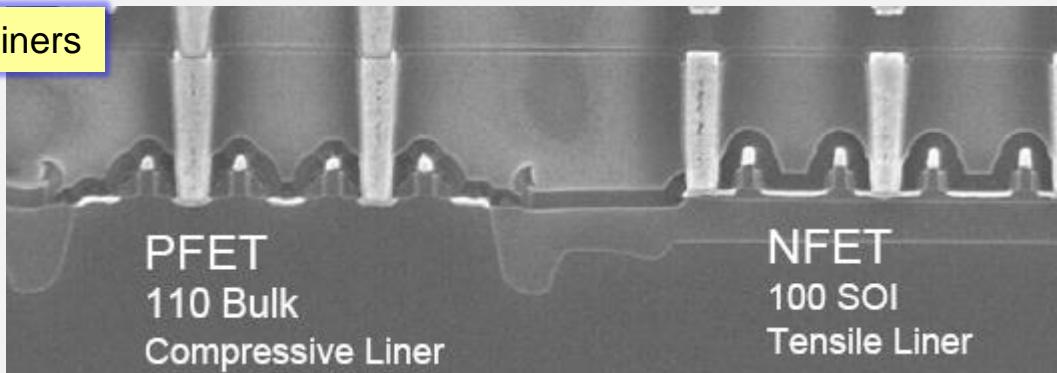
Using liners



Using SiGe and liners



Using HOT and liners



See research for vertical stacked hybrid channel devices chapter 6.8

## 6.1 Overview

Feature size, pitch, node

## 6.2 State of the Art MOSFETs

Electrical and geometrical parameters

## 6.3 Mobility Enhanced MOSFETs

Strain, SiGe, Liners, Directed Channels, Gate-Last

## 6.4 High-k Metal-Gate MOSFETs

Poly-Depletion, FuSi

## 6.5 Vertical MOSFETs

## 6.6 Halo MOSFETs

## 6.7 SOI-MOSFETs

## 6.8 Multi-Gate MOSFETs

Fully-depleted MOSFETs, FinFETs, Nanowire FETs

## 6.9 Advanced Memory Devices

In contacts of materials always a contact potential exists which depends more or less on the workfunction of the various materials  
This contact potential causes a depletion zone depending on charge carrier density

### Conventional CMOS: Highly doped Poly-Silicon is used as Gate electrode metallization

Advantages: Technological: \* self-adjust S/D formation because of high temperature stability (1100°C), no reaction with  $\text{SiO}_2$   
Electrical: using p+ and n+ poly gates symmetrical  $V_T$  on doped Si (p-MOS and n-MOS) can be realized  
important for good CMOS-inverter characteristics

Disadvantages

- 1) high resistivity of gate electrode even with silicide on -> additional design and technological effort on Gate lines in connection with thin gate oxides (3-1)nm as required from scaling roadmap
- 2) penetration of boron doping atoms from poly-Si Gate in/through oxide causes shift in threshold voltage  $V_T$
- 3) because poly-Si is just high doped ( $10^{20} \text{ cm}^{-3}$ ) and not a metal ( $10^{22} \text{ cm}^{-3}$ ) a depletion width exist of about 1nm at the interface of poly-Si to  $\text{SiO}_2$ .  
With scaling to  $t_{\text{ox}} \sim 1\text{nm}$  this depletion zone contributes in Gate capacitor to electrical oxide thickness, this reduces gate capacitance and therefore drive current  $I_{\text{on}}$

Possible solutions:

- 1) extremely high doping (ev.Ge) -> but creating charged defects -> mobility in channel and  $I_{\text{on}}$  decreases
- 2) opposite doping of poly-Gate -> no help for boron diffusion
- 3) new Gate material -> Metal -> poly-Si in accumulation, if MOS is in inversion  
-> but additional processes are needed (increasing cost, decreasing yield)  
-> what about adjustment of  $V_T$ ?  
-> what about manufacturing ?

these solutions are addressed in this section →

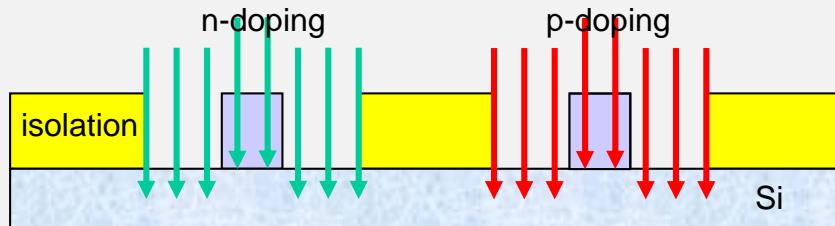
Conventional CMOS:

**Poly-Si Gate allows selfadjusting process with nanometer resolution:**

S/D extensions doping,  
spacer formation,  
S/D doping,  
Halo doping

Various processes for poly-Gate are possible:

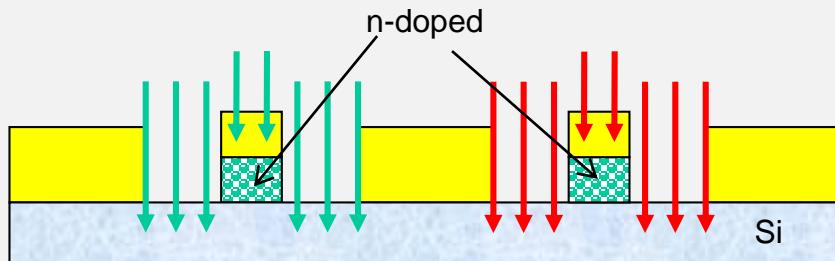
1



Standard process (1970 - 2007):

- 1) deposit undoped poly-Si for all devices
- 2) Gate patterning (1 mask)
- 3) S/D doping for n-MOSFETs and p-MOSFETs (2 masks)

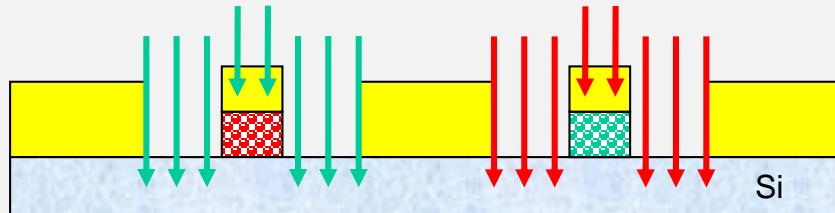
2



Sometimes used:

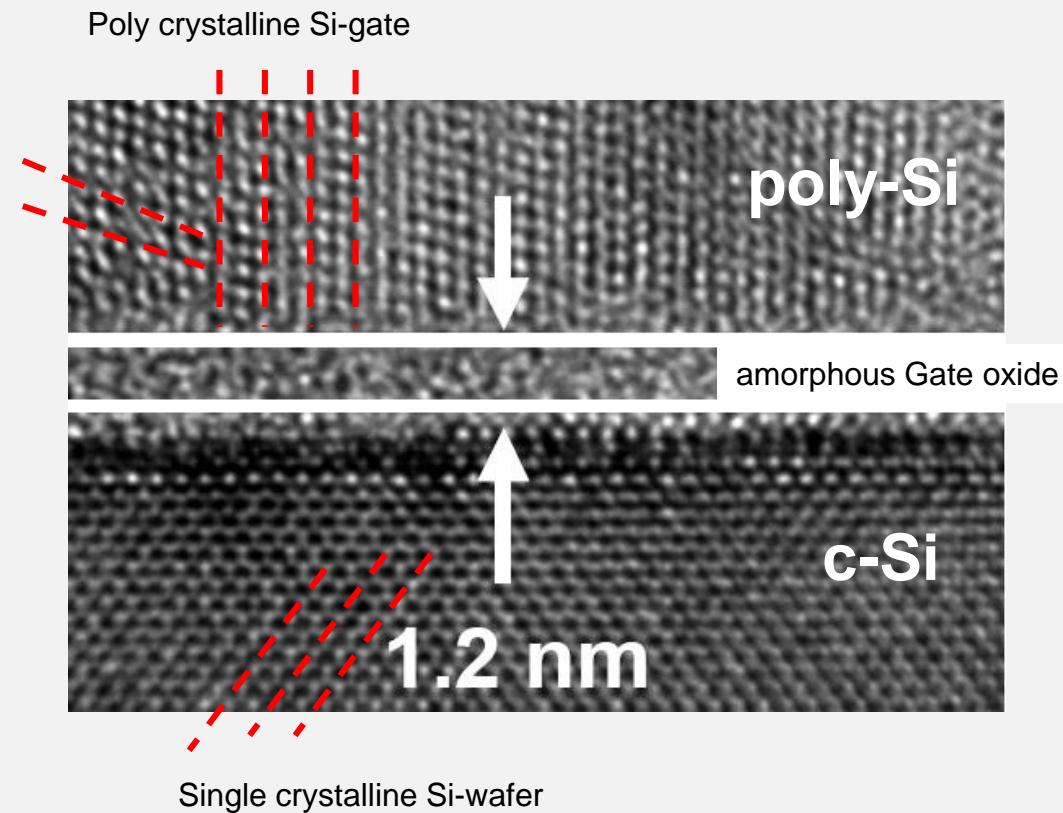
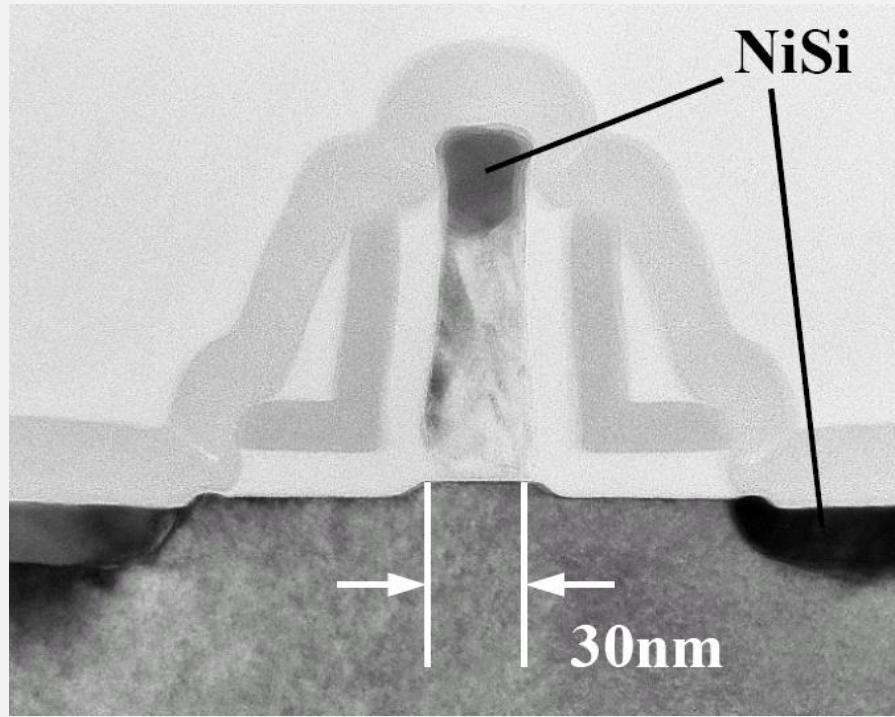
- 1) deposit single-doped poly-Si for all devices
- 2) Gate patterning (1 mask)
- 3) S/D doping for n-MOSFETs and p-MOSFETs (2 masks)

3

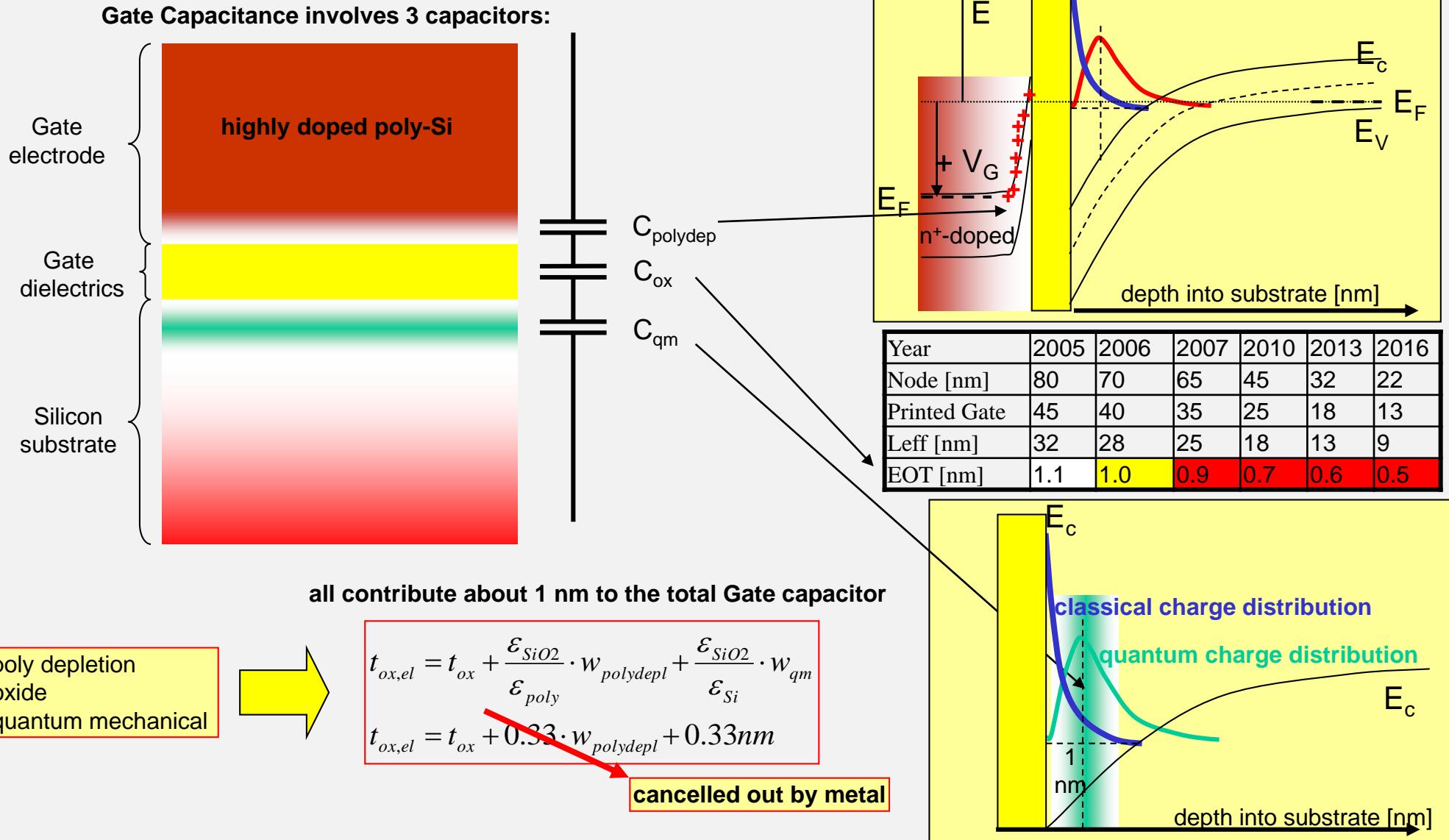


Not used yet:

- 1) deposit opposite doped poly-Si for n-Gates and p-Gates (2 masks)
- 2) Gate patterning (1 mask)
- 3) S/D doping for n-MOSFETs and p-MOSFETs (2 masks)



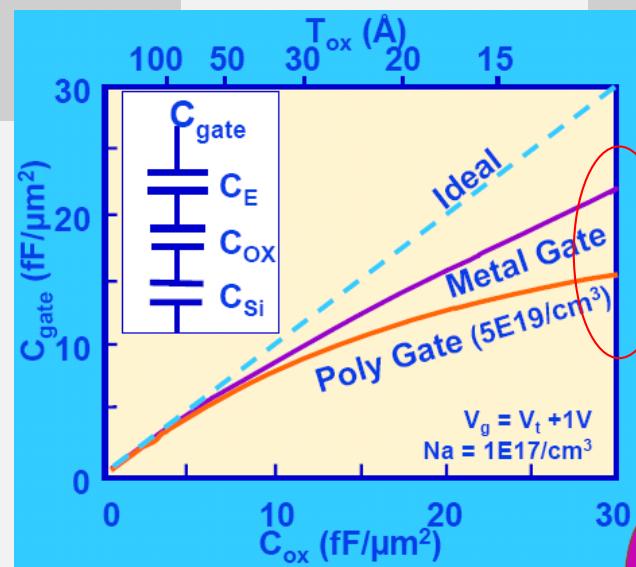
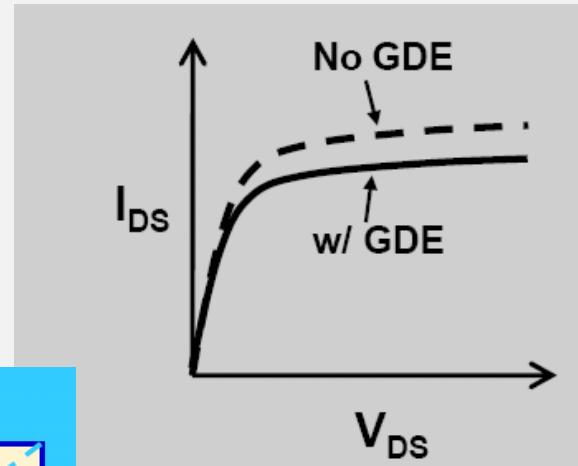
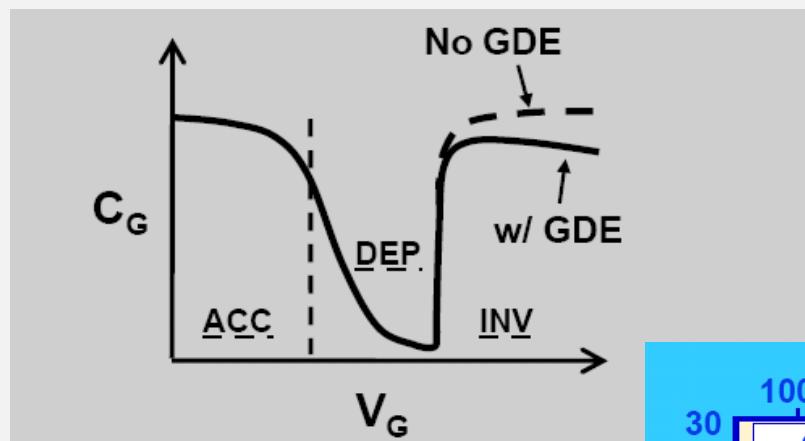
Source: Toshiba, IEDM 2005



## Gate Depletion Effects (GDE) are:

Reduction of Gate capacitance in On-state (inversion)  
-> increase of Short Channel effects (SCE)

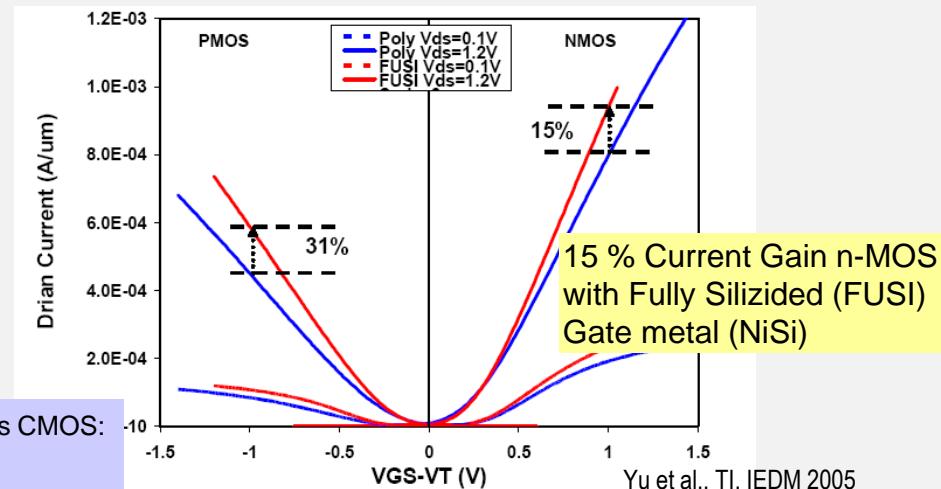
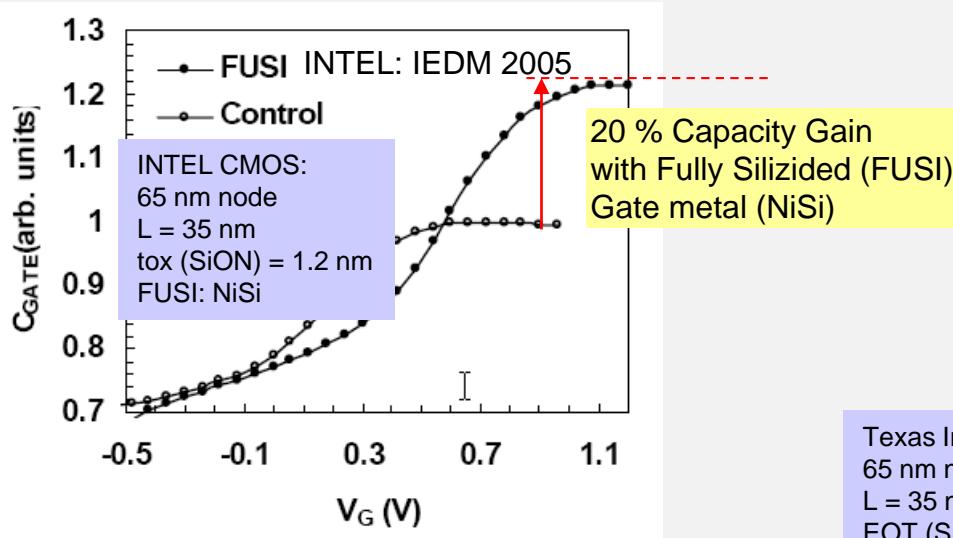
Voltage drop in Gate depletion zone  
-> lower Gate voltage on channel -> lower drive current  $I_{on}$



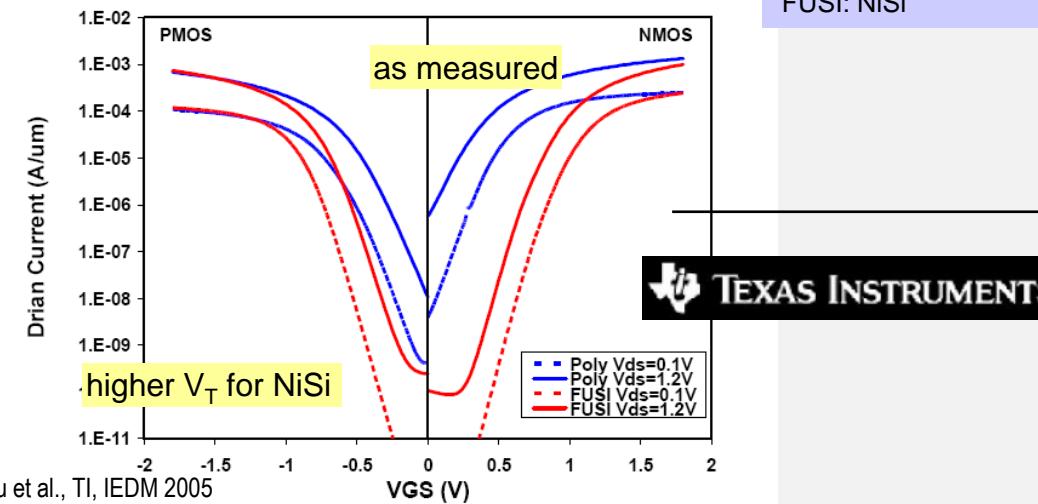
These effects are increased with thinner oxides



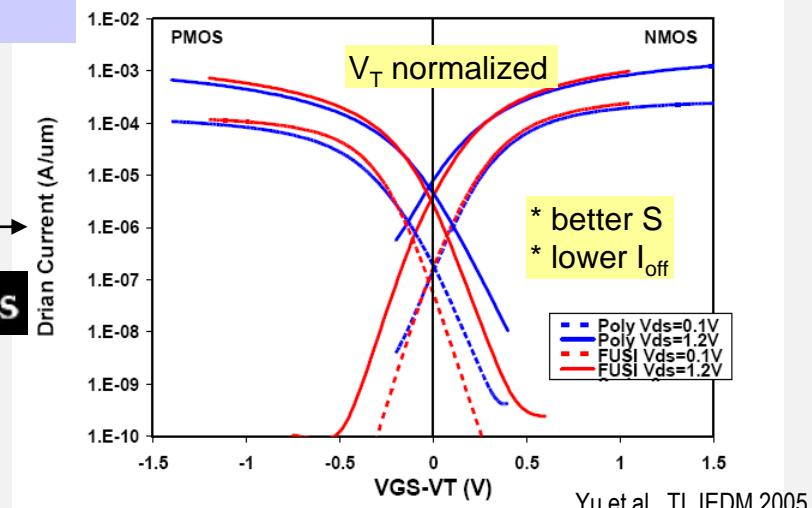
Metal Gate favorable at 65 nm node  
(2007, EOT < 1 nm)



Texas Instruments CMOS:  
65 nm node  
 $L = 35 \text{ nm}$   
 $\text{EOT (SiON)} = 1.7 \text{ nm}$   
**FUSI:** NiSi



TEXAS INSTRUMENTS

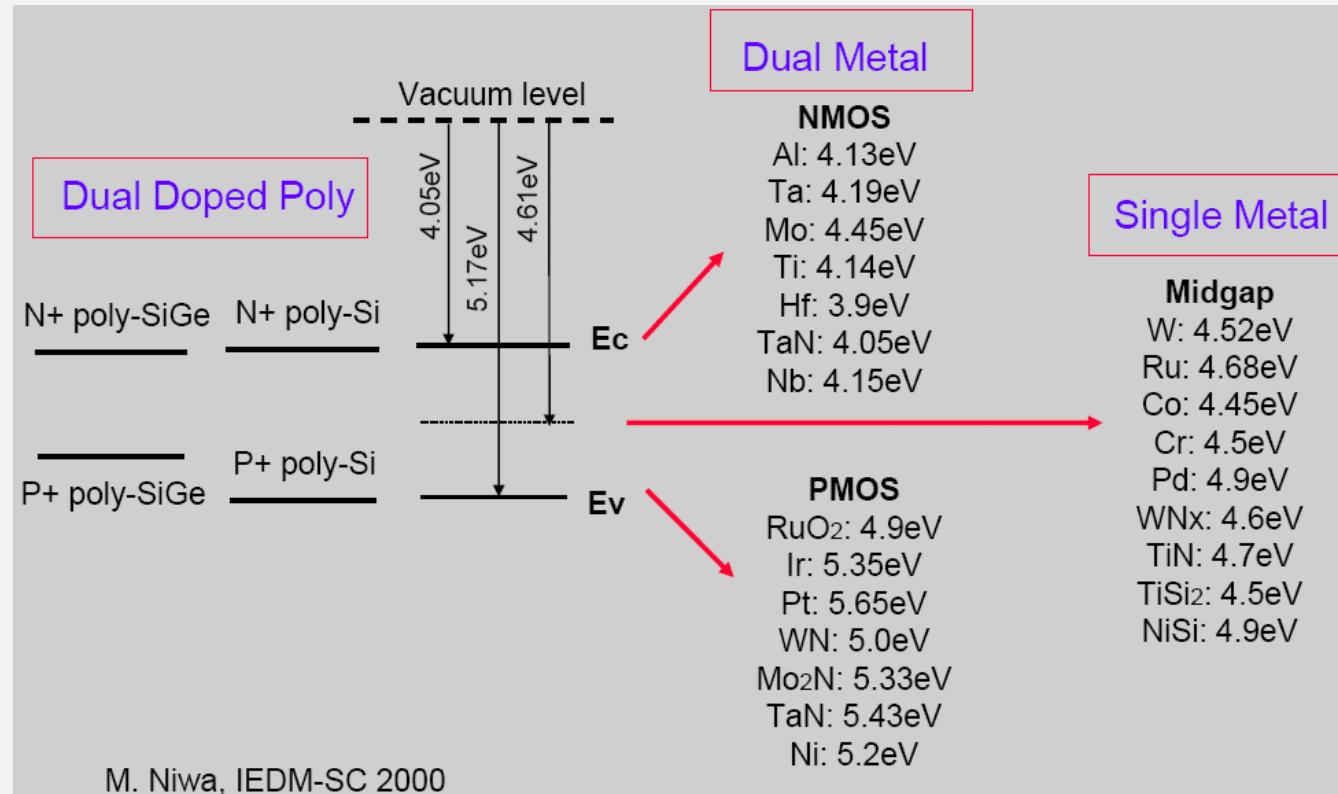


Metal Gate works !

But: What about  $V_T$  ?

What metals can be used to achieve symmetrical  $V_T$ ?

To overcome the problem of doping fluctuations and to achieve better gate control undoped channels will be needed.  
Therefore the barrier must be created by the workfunction of the gate metal.

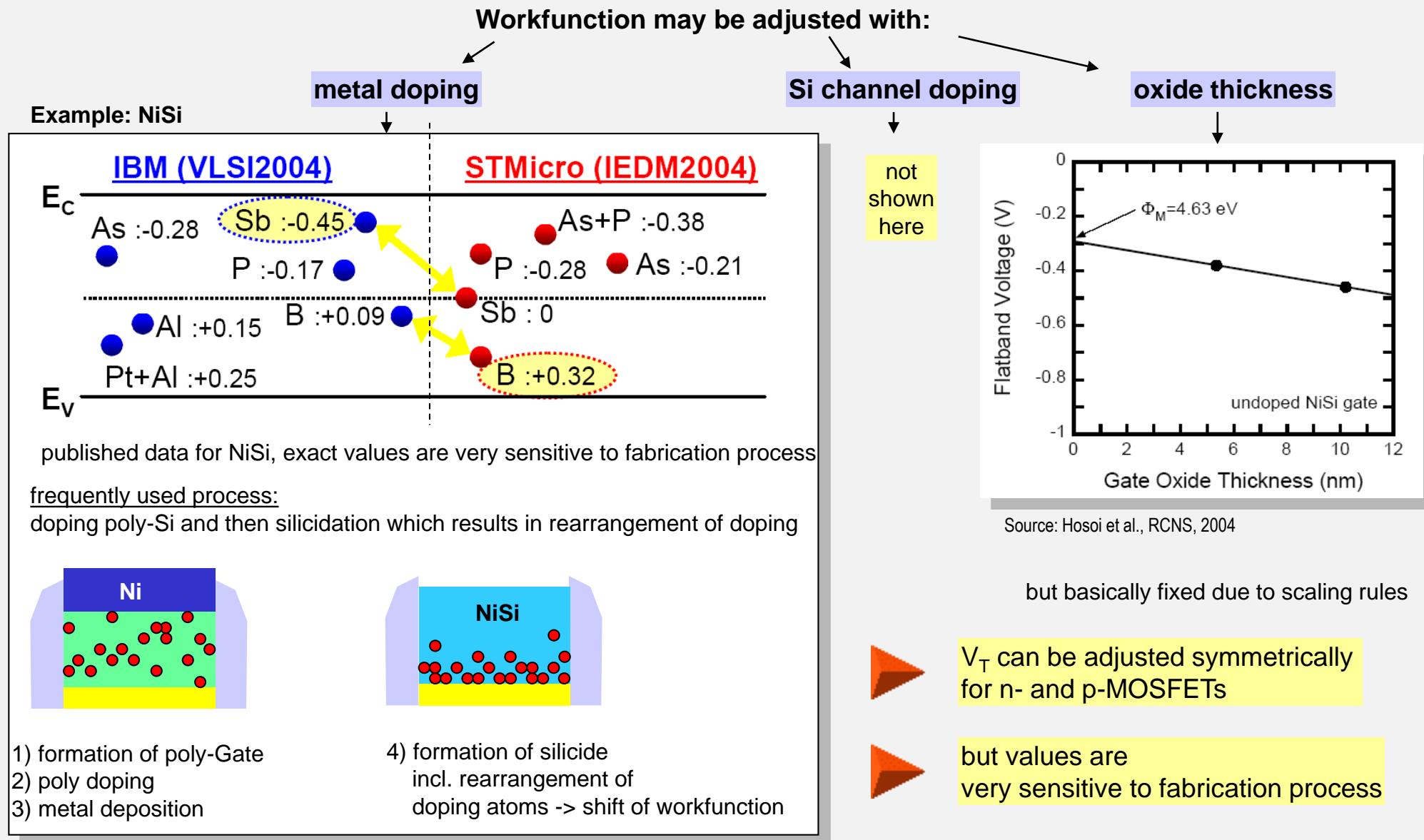


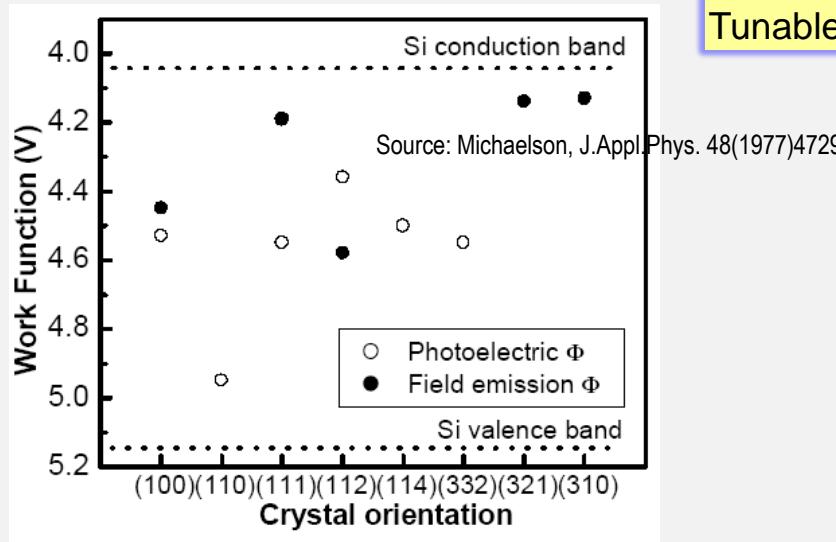
Basic value of work function may be chosen from reported bulk values

- but:
- 1) scaling  $V_T$  will need different metal  $\Omega_{MS}$  every future node -> no chance for technology
  - 2) measured workfunction may be dependent on process, charges (doping) and dielectrics



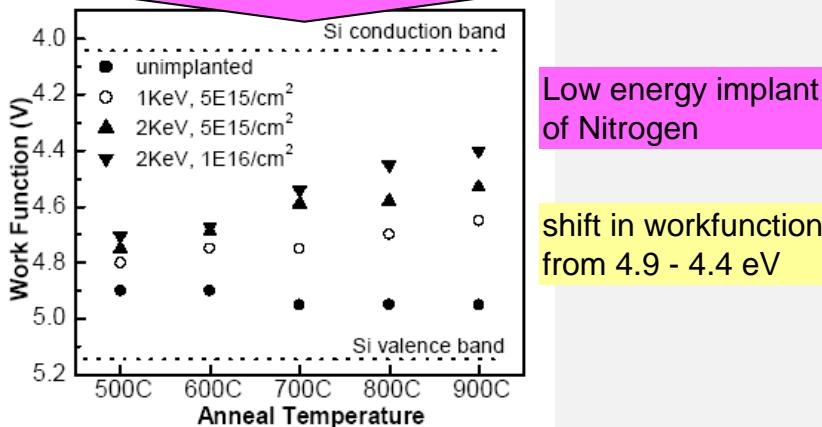
try to adjust workfunction with one metal





workfunction of Mo very sensitive to structural variations

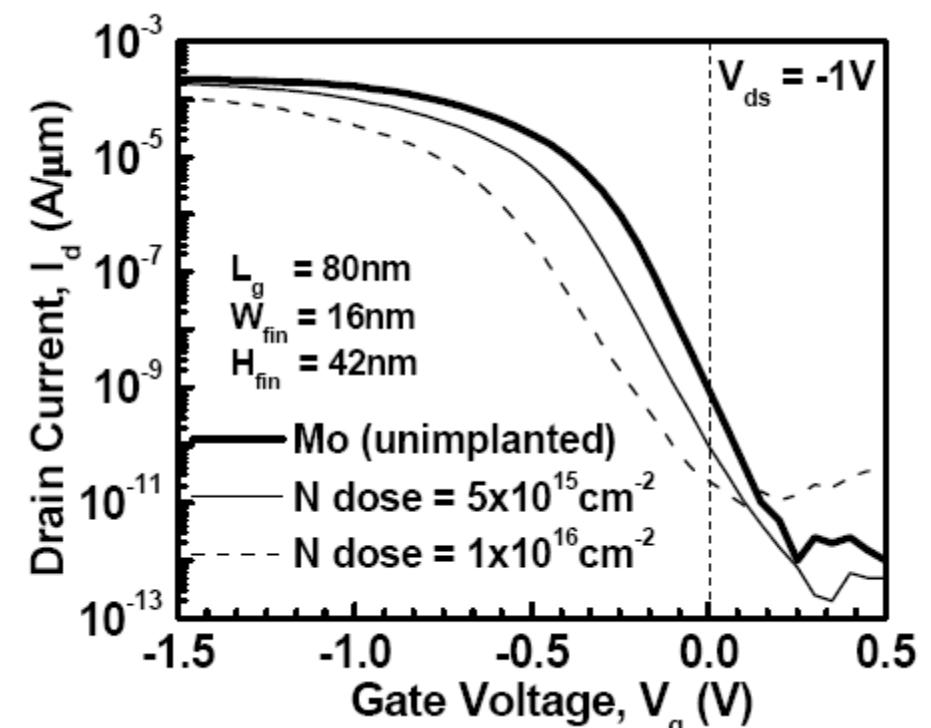
introduce structural modifications by implant and anneal



## Tunable Mo Gate Technology

Molybdenum:

- \* compatibility with CMOS processing
- \* low resistivity
- \* thermal expansion like Si

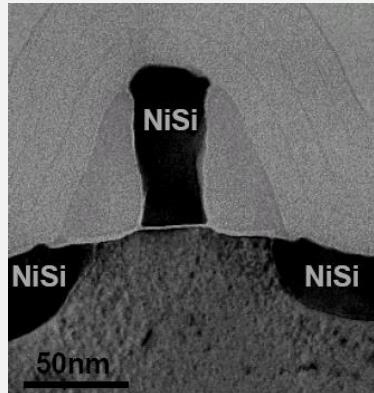


IV-characteristics for a Mo gated PMOS FinFET

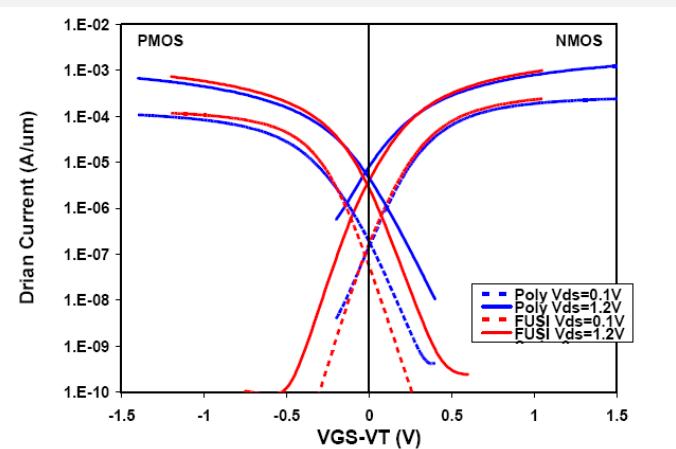
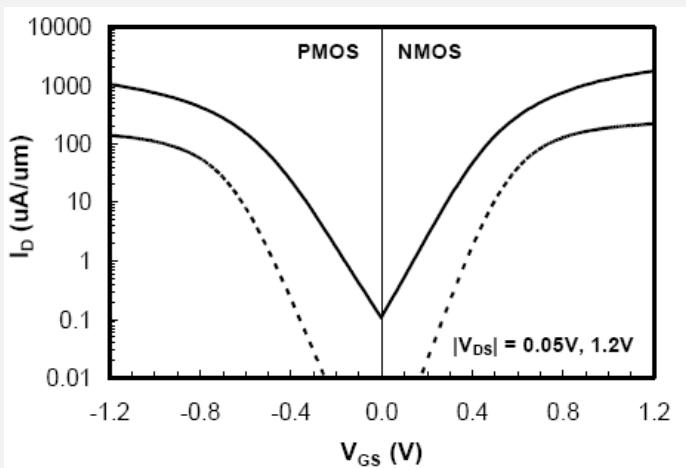
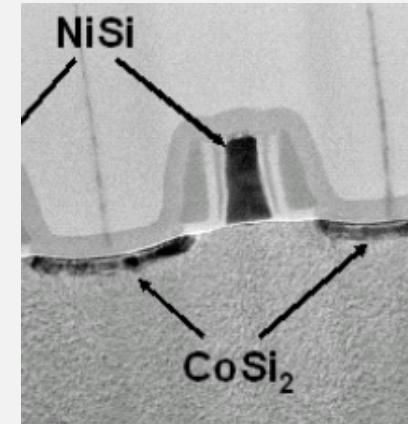
Implantation shifts the curve to more negative biases, indicating a lowering of gate work function

## Fully Silicide Gate in production

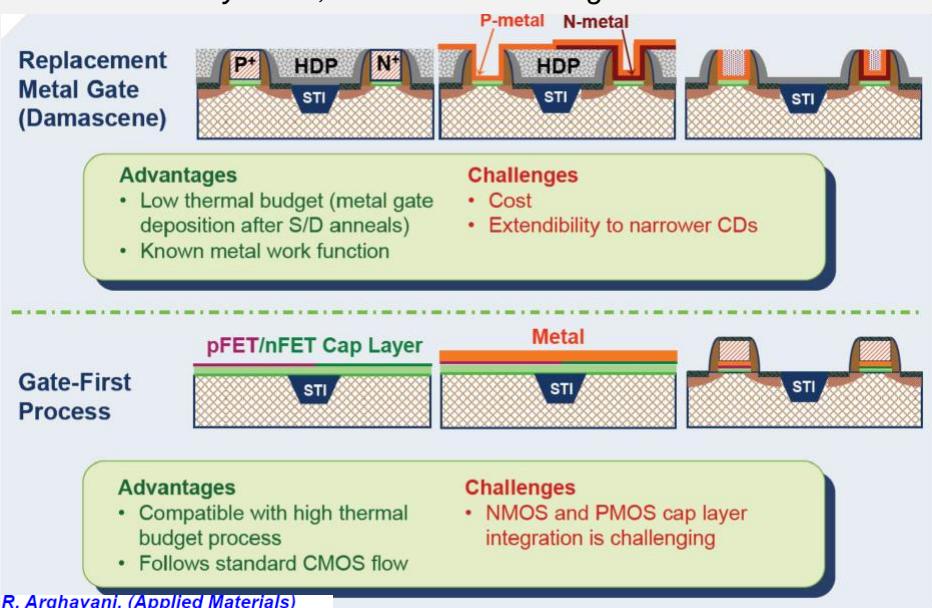
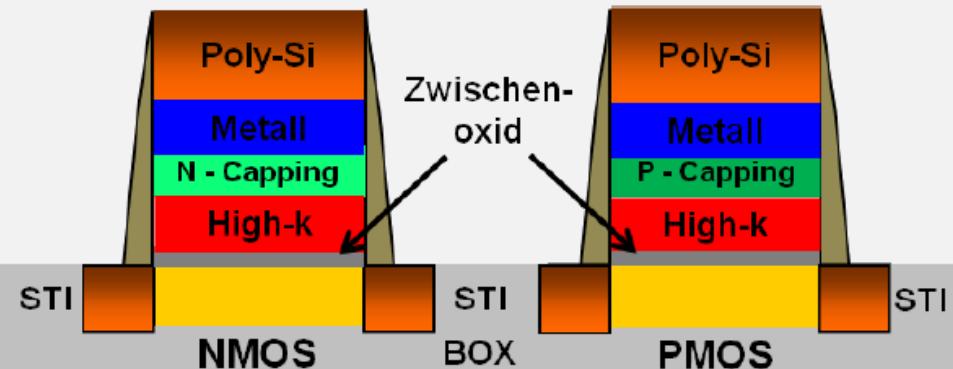
INTEL CMOS:  
65 nm node  
 $L = 35 \text{ nm}$   
tox (SiON) = 1.2 nm  
FUSI: NiSi



Texas Instruments CMOS:  
65 nm node  
 $L = 35 \text{ nm}$   
EOT (SiON) = 1.7 nm  
FUSI: NiSi



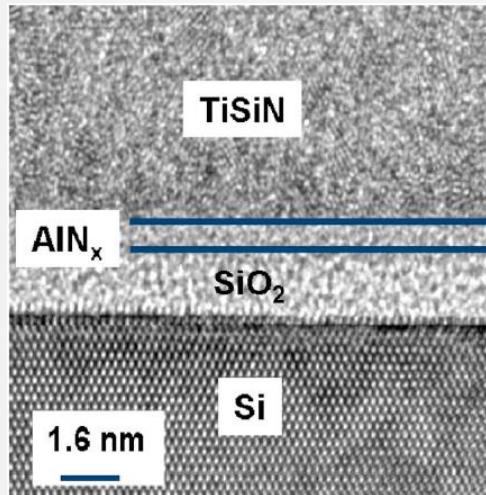
The workfunction values of the fully silicide gates are very sensitive to fabrication process.  
The workfunction of the metal gate is tailored by engineering the metal/dielectric interface. This leads to the usage of capping layers.



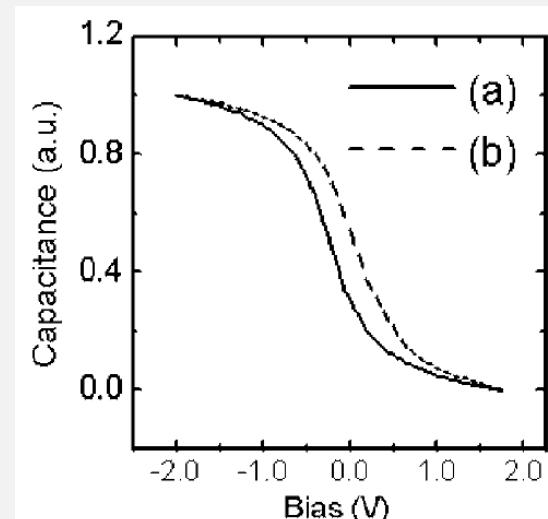
- ▶ Capping layer is a metal, around 1 nm thick
- ▶ Different capping layer for NMOS and PMOS
- ▶ Typical gate stack:  $\text{SiO}_2$ , High-k, capping layer, TiN, polysilicon or metal
- ▶ Used for Gate-First and Replacement-Gate
- ▶ In production since 45 nm technology node (INTEL)
- ▶ All companies are using capping layers to adjust the workfunction

## Metal gate work function engineering using $\text{AlN}_x$ interfacial layers

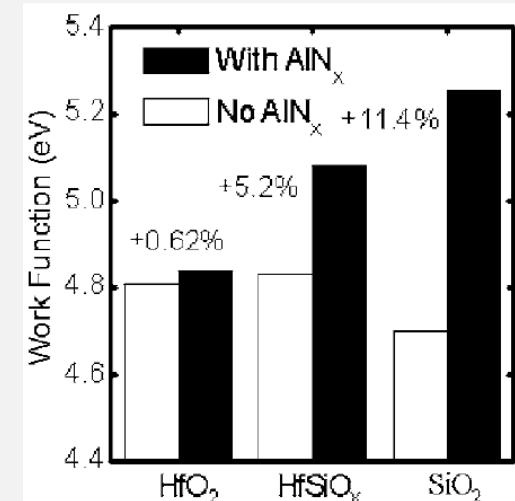
H. N. Alshareef,<sup>a),b)</sup> H. F. Luan, K. Choi, H. R. Harris, H. C. Wen, M. A. Quevedo-Lopez,<sup>a)</sup>  
P. Majhi, and B. H. Lee  
*SEMATECH, 2706 Montopolis Drive, Austin, Texas 78781* Applied Physics 2008



- AlN is amorphous
- Annealing



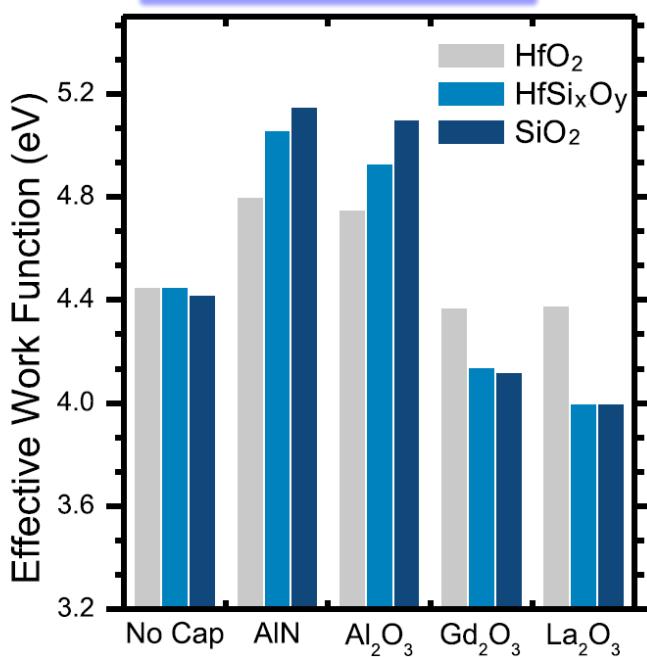
- a) No AlN layer
- b) With AlN layer



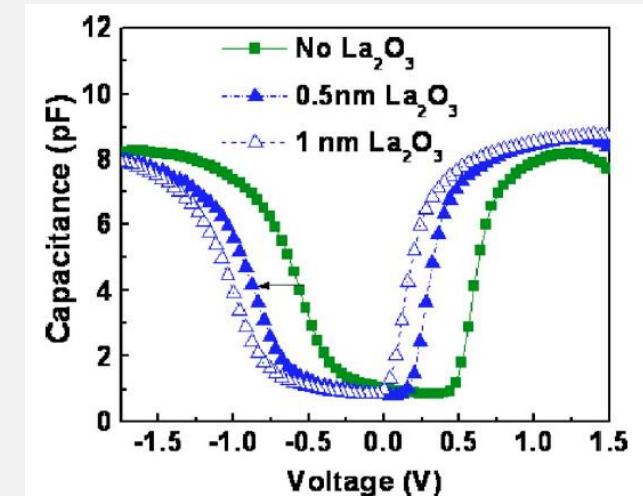
- Workfunction is shifting with the capping layer

- ▶ Capping layer AlN is shifting the workfunction depending on the gate dielectric
- ▶ AlN is making the workfunction more positive
- ▶ Promising candidate for the p-channel MOSFET

## Various capping layers



	n <sup>+</sup> -Si	p <sup>+</sup> -Si
$\Phi$ [V]	~ 4.1	~ 5.2



Source: H.N. Alshareef et al., Work function engineering using lanthanum oxide interfacial layers, Applied Physics 89, 2006

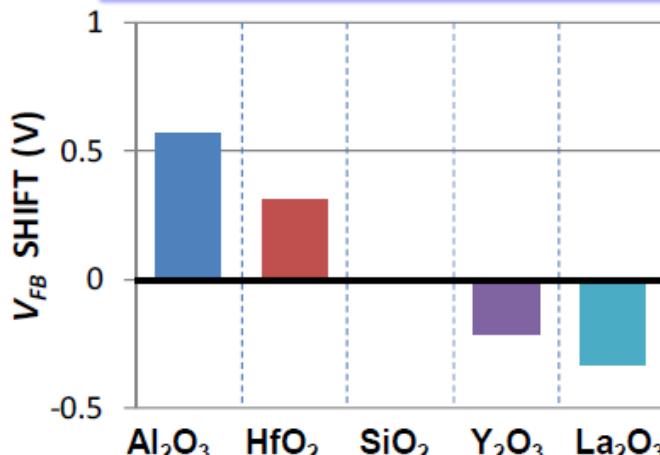
- Workfunction is shifting due to different capping layers
- Depending on the gate dielectric
- Lanthanum causes a negative  $V_{FB}$

► Capping layer AlN/Al<sub>2</sub>O<sub>3</sub> are suitable for PMOS

► Capping layer La<sub>2</sub>O<sub>3</sub> is suitable for NMOS

► Why is the capping layer shifting the workfunktion?

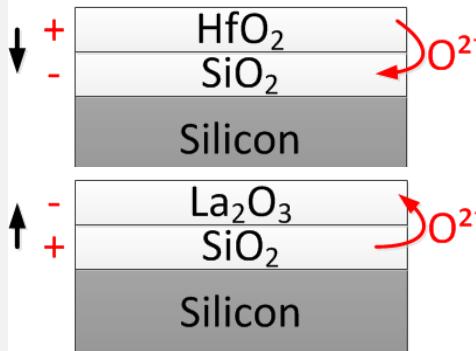
Different high-k materials on  $\text{SiO}_2$



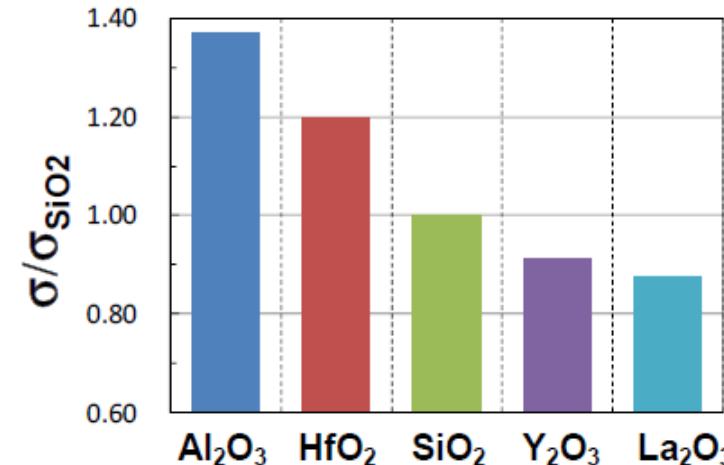
Source: Koji Kita et al., Intrinsic Origin of Electric Dipoles Formed at High-k/SiO<sub>2</sub> Interface, IEDM 2008

- For some oxides there is a positive  $V_{FB}$  shift
- For some oxides there is a negative  $V_{FB}$  shift

►  $V_{FB}$  shift due to the formation of a dipole layer at high-k/SiO<sub>2</sub> interface



Oxygen atoms per unit area



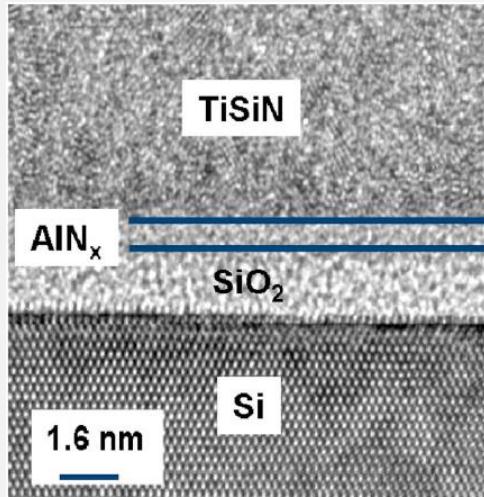
Koji Kita et al., Intrinsic Origin of Electric Dipoles Formed at High-k/SiO<sub>2</sub> Interface, IEDM 2008

- $\text{Al}_2\text{O}_3$  and  $\text{HfO}_2$  have a higher density of O<sub>2</sub> than SiO<sub>2</sub>
  - $\text{Y}_2\text{O}_3$  and  $\text{La}_2\text{O}_3$  have a lower density of O<sub>2</sub> than SiO<sub>2</sub>
- To adjust the oxygen atom density at the interface, oxygen atoms move from higher density material to lower density material

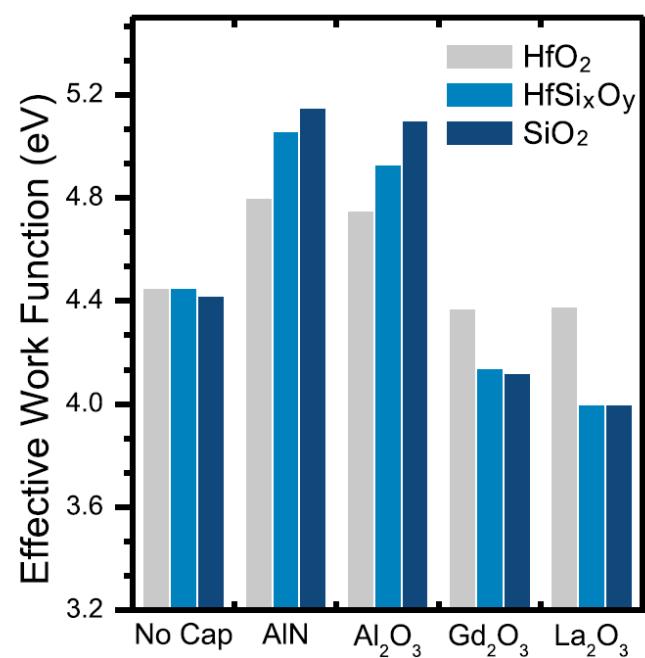
► Voltage drop at the interface layer for HfO<sub>2</sub> leads to higher workfunction

► Voltage drop at the interface layer for La<sub>2</sub>O<sub>3</sub> leads to smaller workfunction

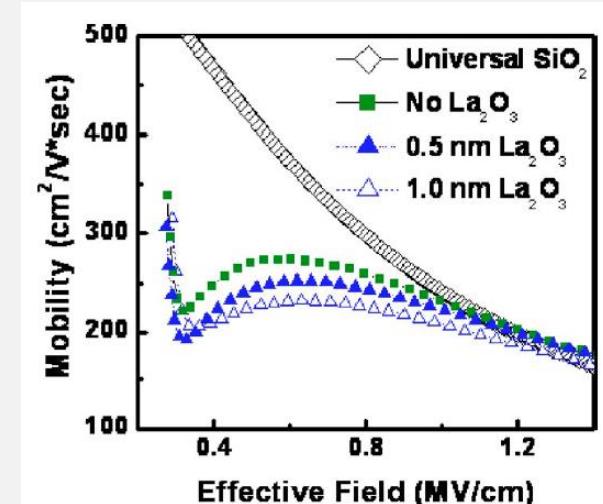
## High-k materials as capping layer



Source: H.N. Alshareef et al., Work function engineering using AlN interfacial layers, Applied Physics 89, 2006



Source: J.A. Carraveo-Frescas et al., Experimental and theoretical investigation of the effect of SiO<sub>2</sub> content in gate dielectrics on work function shift induced by nanoscale capping layers, Applied Physics 101, 2012



Source: H.N. Alshareef et al., Work function engineering using lanthanum oxide interfacial layers, Applied Physics 89, 2006

Through annealing, the capping metal reaches the SiO<sub>2</sub> /HfO<sub>2</sub> interface layer creating the dipole

The higher the amount of oxygen at the interface is the higher is the influence of the dipole

Disadvantage is a low mobility of the carriers due to higher scattering caused through the dipole (especially for PMOS)

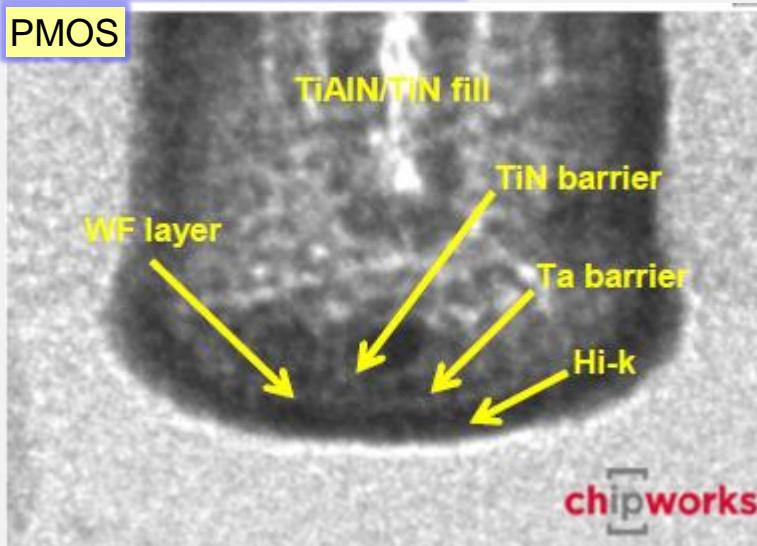
No knowledge about the reliability of lanthanum



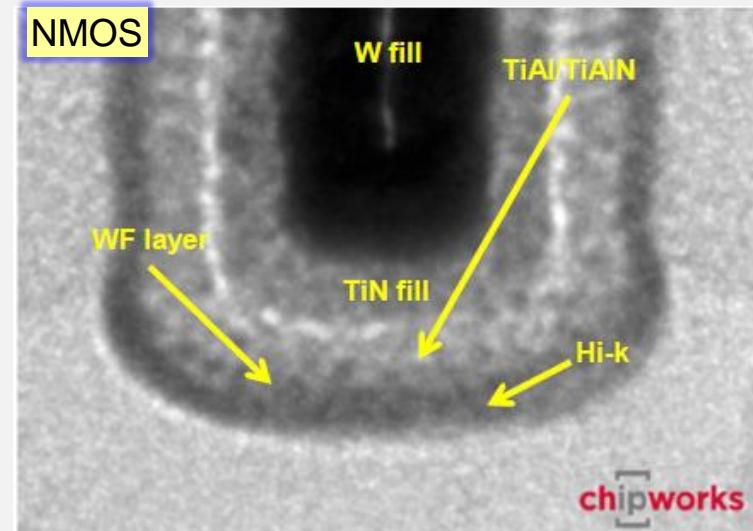
What is industry doing?

Gatestack Intel 14 nm 2014

PMOS



NMOS



- TiN is used as capping layer
- TiN:  $\Phi [V] = 5.1 - 4.7 \text{ eV}$ ; ( $p^+$ -Si:  $\Phi [V] = 5.1 \text{ eV}$ )
- No aluminum as capping layer needed
- TiN has a high thermo stability to  $\text{HfO}_2$

- TiAl/TiAIN is used as capping layer
- TiAlN:  $\Phi [V] = 4.3 - 5.1 \text{ eV}$ ; ( $n^+$ -Si:  $\Phi [V] = 4.1 \text{ eV}$ )
- Aluminum is reducing the workfunction of TiN
- Strong dependence on the process



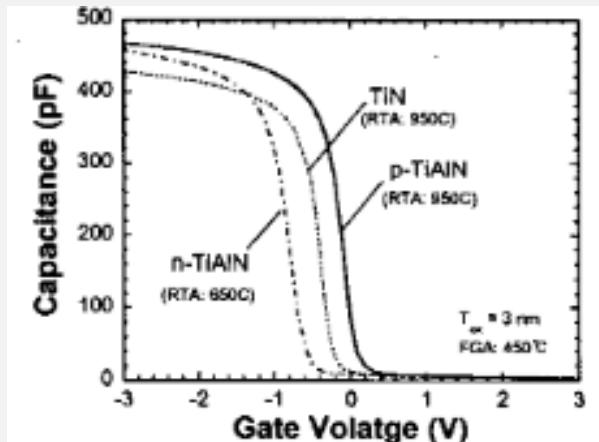
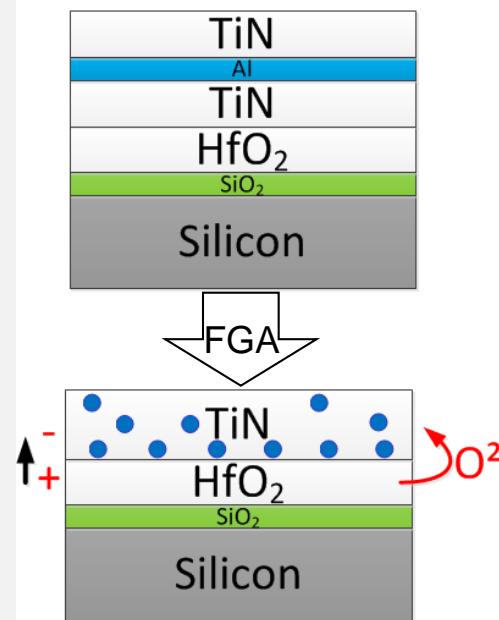
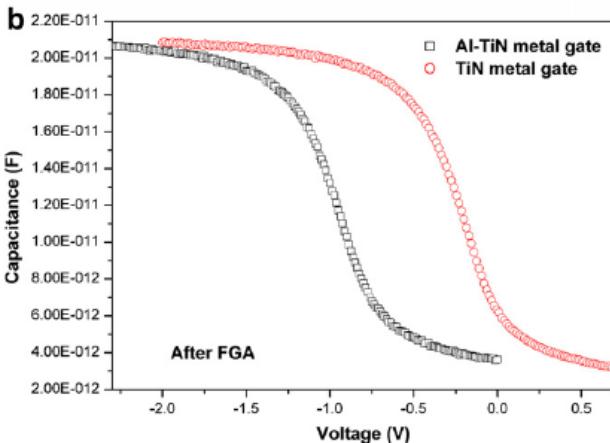
No mobility degradation



CMOS compatible



But very critical process window and high sophisticated process



- $\text{HfO}_2$ ,  $\text{TiN}$ ,  $\text{Al}$ ,  $\text{TiN}$  are deposited
  - Forminggas annealing ( $450^\circ\text{C}$ ) diffusion of  $\text{Al}$  towards the interface of  $\text{TiN}$  and  $\text{HfO}_2$
- ⇒ Dipole
- ⇒ Shift to negative Flatband Voltage

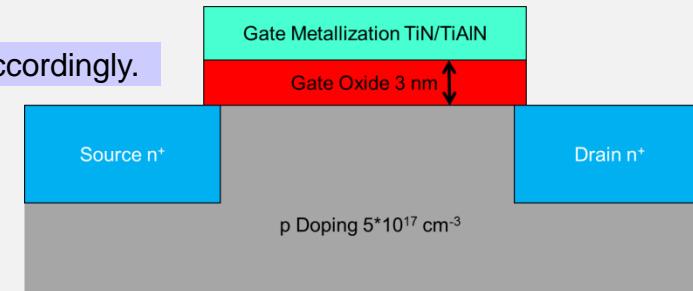
- To high N in  $\text{TiAlN}$  leads to positive  $V_{FB}$
- To high amount of aluminum leads to positive  $V_{FB}$
- Annealing with  $\text{O}_2 \Rightarrow \text{Al}_2\text{O}_3$  which leads to positive  $V_{FB}$

- ▶ TiN is suitable as capping layer
- ▶ NMOS aluminum has to diffuse to the interface high-k/TiN to become a negative shift of  $V_{FB}$
- ▶ Small changes in technology can lead to different workfunctions

Is the threshold voltage still symmetrical with the new capping layers?

For a quick calculation we take the values of exercise 3 and change the workfunction accordingly.

	TiN	TiAlN
$\Phi$ [V]	$\sim 4.8$	$\sim 4.4$



NMOS

$$\Phi_{MS} = \Phi_M - \left( X_{si} + \frac{E_g}{2q} + |\Psi_B| \right) = 4.4V - 4.05V - 0.56V - 0.45V = -0.66V$$

TiAlN Exercise 3

$$V_{Tn} = \left| \frac{Q''_{Depl}}{C''_{ox}} \right| + |2\Psi_B| + \Phi_{MS} = 0.34V + 2*0.45V - 0.66V = \underline{\underline{0.58V}}$$

Exercise 3

PMOS

$$\Phi_{MS} = \Phi_M - \left( X_{si} + \frac{E_g}{2q} - |\Psi_B| \right) = 4.8V - 4.05V - 0.56V + 0.45V = 0.64V$$

TiN Exercise 3

$$V_{Tp} = - \left| \frac{Q''_{Depl}}{C''_{ox}} \right| - |2\Psi_B| + \Phi_{MS} = -0.34V - 2*0.45V + 0.64V = \underline{\underline{-0.6V}}$$

Exercise 3

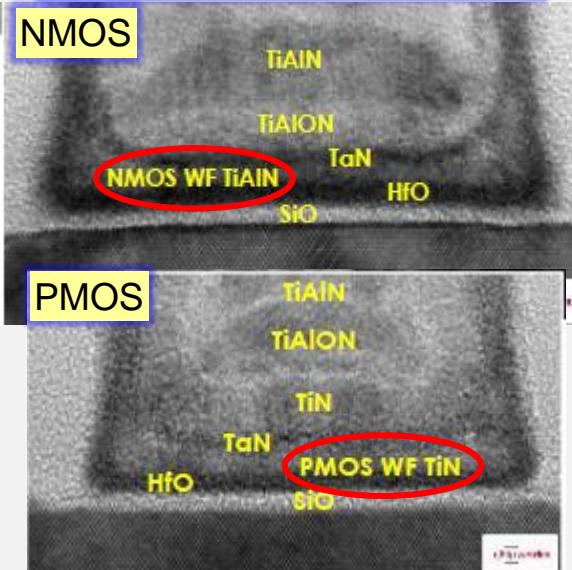


New metal gate stack leads to symmetrical threshold voltage

Channel doping must be eliminated to avoid dopant fluctuations and increase drive current  $I_{on}$

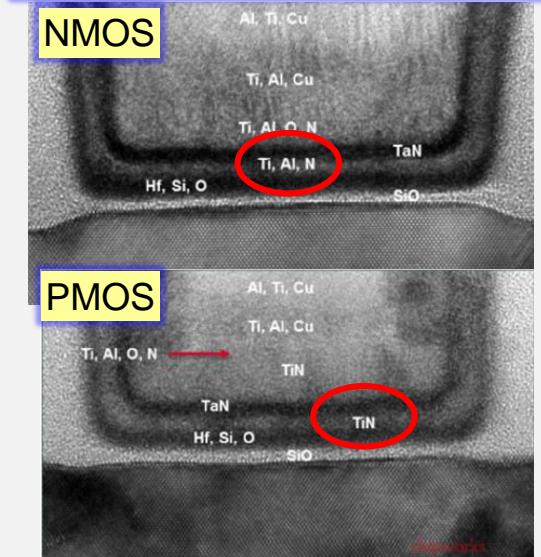
Poly-Si Gate depletion must be eliminated to avoid lowering of Gate capacitance and increasing short-channel effects

Gatestack TSMC 20 nm 2012



Metal Gate is needed

Gatestack Samsung 20 nm 2012



► Samsung/TSMC/Intel are using TiN for PMOS and TiAlN for NMOS as workfunction layer/capping layer

► Workfunction is very sensitive to fabrication and test chips have to be fabricated

► Most companies used until 20 nm technology node gate-first with fully silicid gate and capping layers

► Since the 20 nm, technology node all companies are using replacement-gate and capping layers

**SOFTPEDIA®** DESKTOP ▾ MOBILE ▾ WEB ▾ NEWS

FLASH SALE:

≡ Softpedia > News > Technology & Gadgets > CPU

Jan 20, 2011 13:16 GMT · By Sorin Nita · Share: 8+

**Yesterday, Common Platform members announced that the coalition will move to gate-last technology when it makes the transition to the 20nm manufacturing node, plans for 32nm and 28nm chip fabrication remaining unchanged at this time.**

The Common Platform Initiative is made out of IBM, GlobalFoundries, and Samsung Electronics, some of the world's largest semiconductor manufacturers, in an effort to introduce common process technology across all of their production facilities.

This helps in reducing research and development costs as the three companies can share those expenses between them.

Up until now, the companies were some of the most vehement supporters of the gate-first approach, GlobalFoundries reaffirming its support for the technology more than just a few times.

However, its position has now changed as the company decided to move to gate-last with their 20nm node.

The difference between the two technologies lies in the timing when the metal electrode is deposited, before or after the high temperature activation anneal(s) of the flow, gate-first, at least in theory, allowing customers to transition to a lower manufacturing node without having to redesign their chips.

According to Jon Carvill, VP of GlobalFoundries Communications, the change "is due to the design rules and desired scaling our customers want at 20nm. Gate Last was the best option to hit it and Gate First was the better option at 32/28nm."

"We have multiple customers engaged on Gate First 32nm/28nm as does Samsung (who made this disclosure with us yesterday)," concluded the company's rep.

Until Common Platform announced the switch to gate-last, Intel and TSMC were the major proponents of this technology.

# SolidState TECHNOLOGY®

Insights for Electronics Manufacturing

HOME SEMICONDUCTORS PACKAGING MATERIALS MEMS LEDS DISPLAYS SEMICON DESIGN

## Chipworks

Like Share 0 Tweet Share 0 G+1 0

## Common Platform Goes Gate-Last – at Last!

At the IBM/GLOBALFOUNDRIES/Samsung Common Platform Technology Forum on Tuesday, Gary Patton of IBM announced that the Platform would be moving to a gate-last high-k metal-gate (HKMG) technology at the 20-nm node.

At the 45- and 32-nm nodes there has been a dichotomy between gate-last as embodied by Intel, TSMC, and UMC, and gate-first, promoted by the Common Platform and others such as Panasonic. (Though, to be realistic, Intel's is the *only* HKMG we've seen so far, and the only 32-nm product.)



Since the 20 nm technology node all companies are using Gate-Last/Replacement-Gate

## 6.1 Overview

Feature size, pitch, node

## 6.2 State of the Art MOSFETs

Electrical and geometrical parameters

## 6.3 Mobility Enhanced MOSFETs

Strain, SiGe, Liners, Directed Channels, Gate-Last

## 6.4 High-k Metal-Gate MOSFETs

Poly-Depletion, FuSi

## 6.5 Vertical MOSFETs

## 6.6 Halo MOSFETs

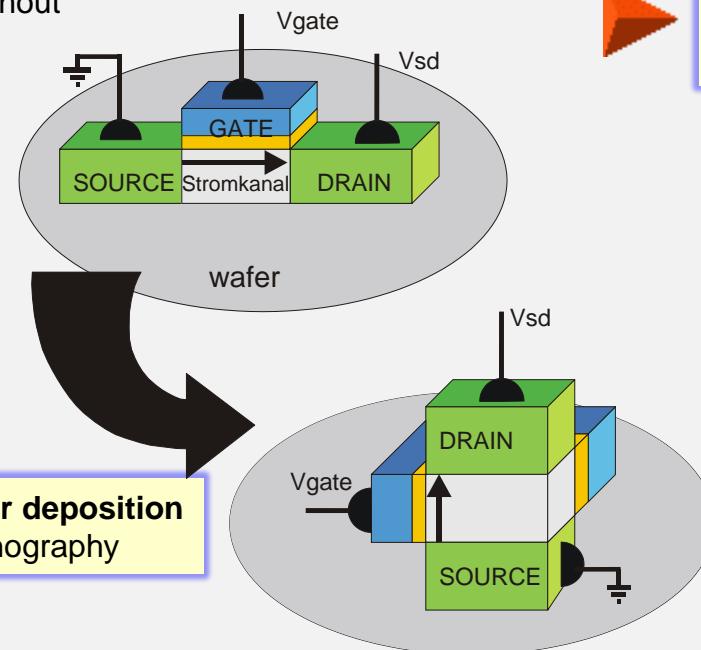
## 6.7 SOI-MOSFETs

## 6.8 Multi-Gate MOSFETs

Fully-depleted MOSFETs, FinFETs, Nanowire FETs

## 6.9 Advanced Memory Devices

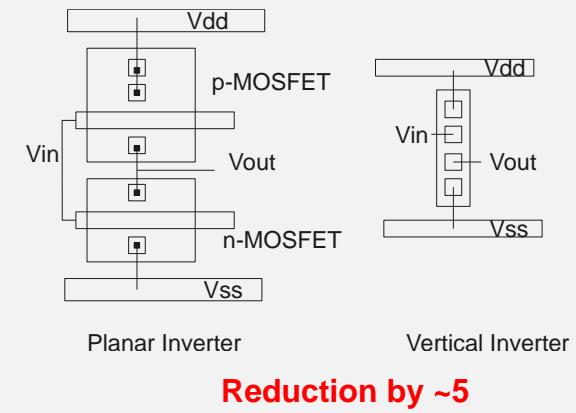
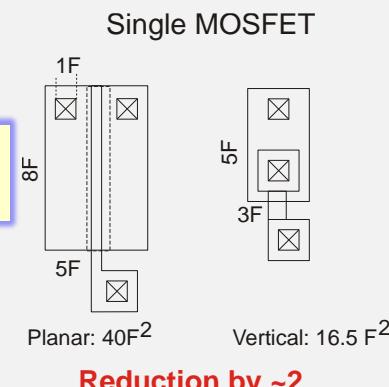
**Vertical MOSFETs** offer the possibility to extremely reduce channel length without sophisticated lithography:



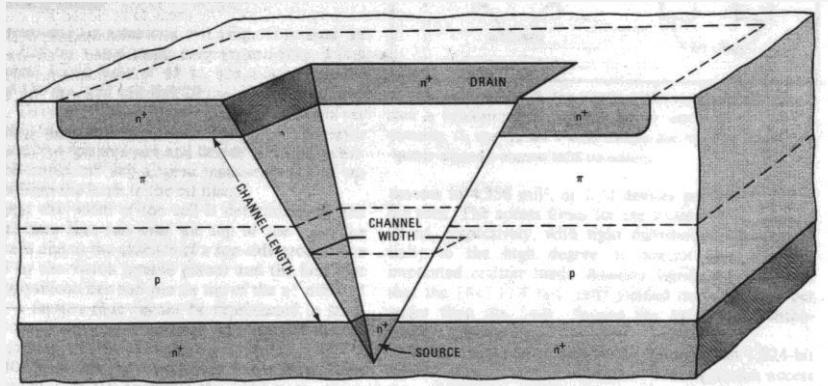
Stacking devices results in 3-dim integration and further reduces the footprint. The difficulties with vertical interconnects increase

channel length defined by layer deposition (nm-resolution), not by optical lithography

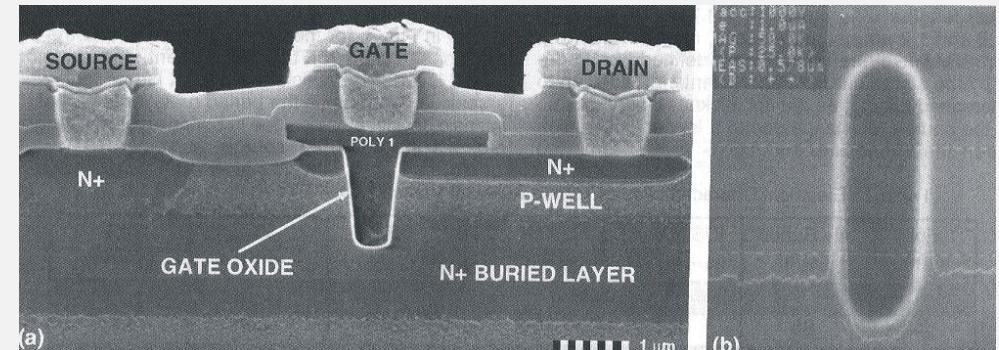
Using wrapped gates the MOSFETs have about factor 2 smaller footprints



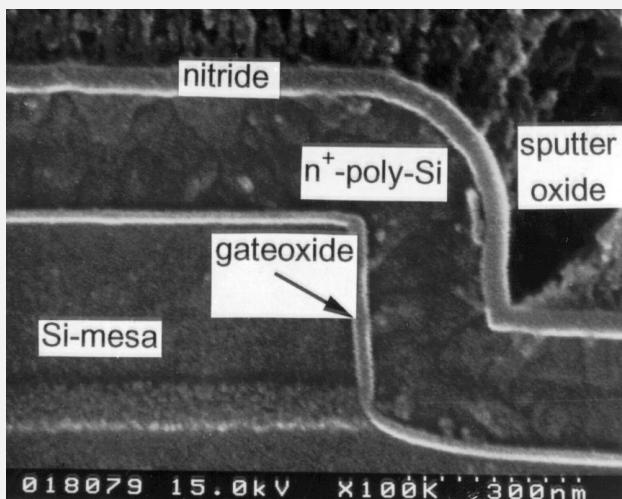
Historical:



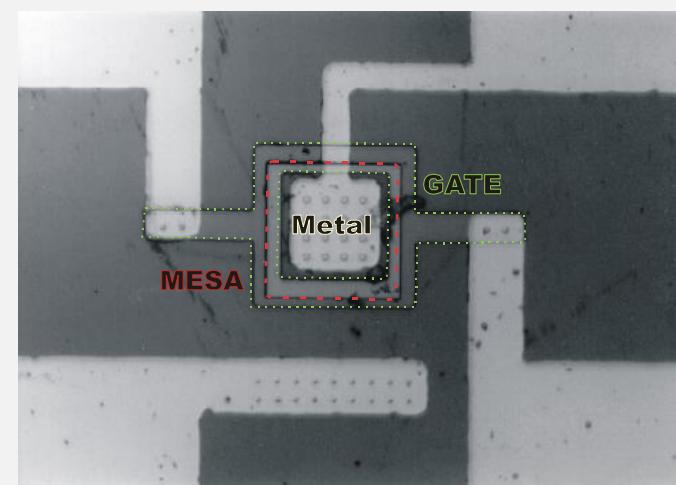
Anisotropic etched VMOS, ~ 1974  
1µm channel length, TTL-logic, 4kB-SRAM



example: Perera et al., IEDM'94, 851  
MOTOROLA



50-nm VMOS

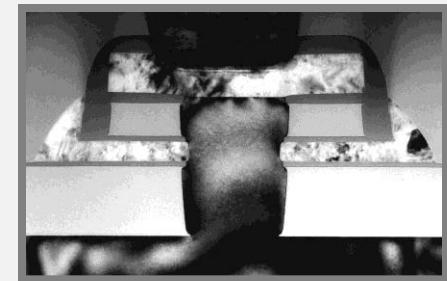
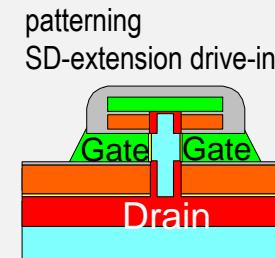
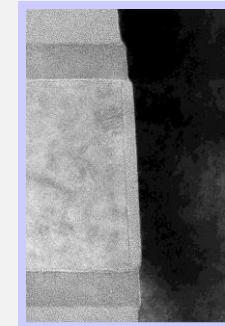
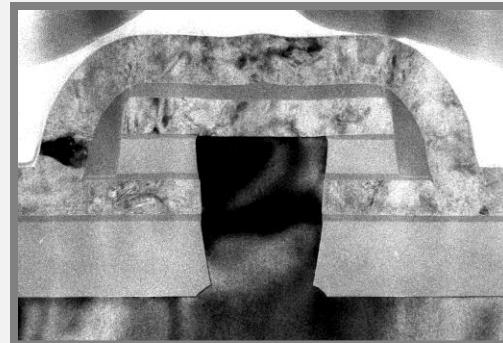
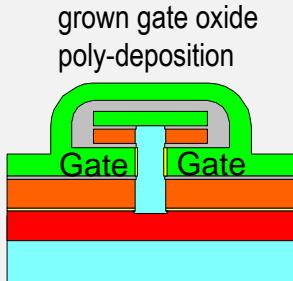
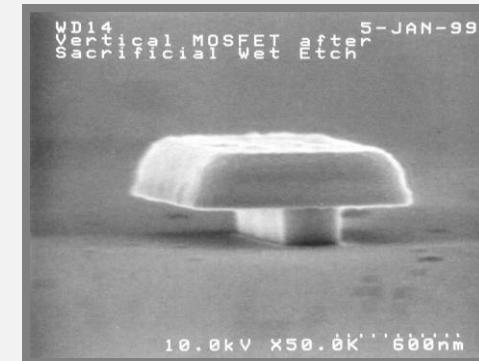
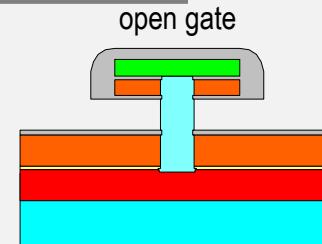
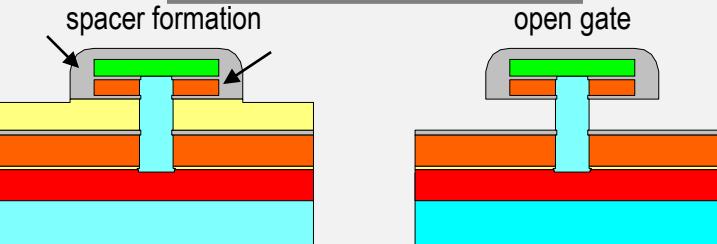
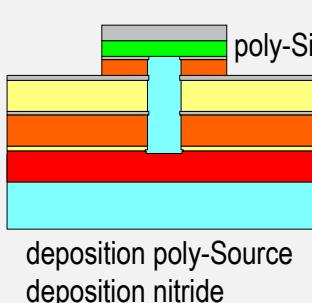
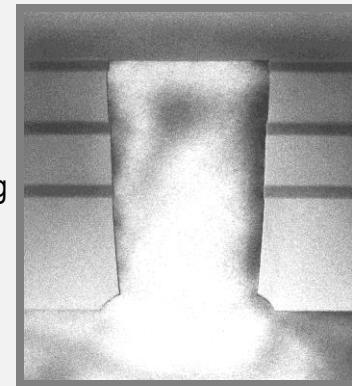
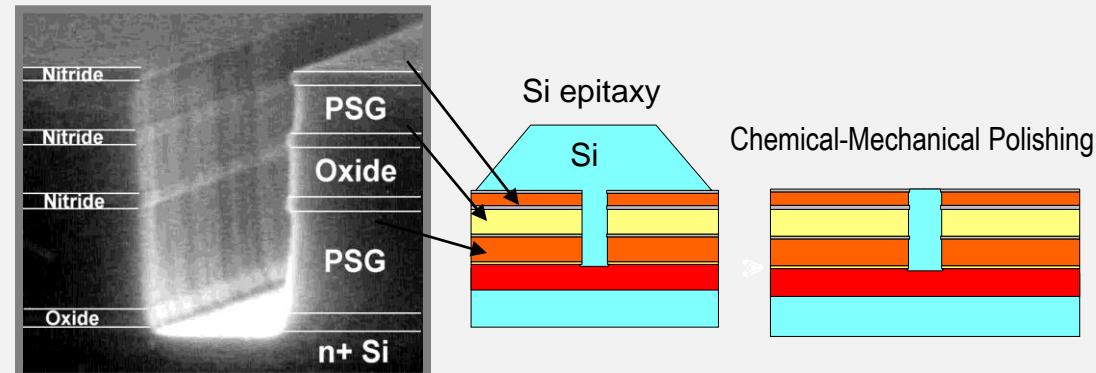
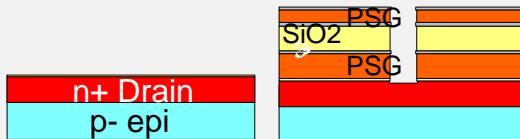


wrapped gate

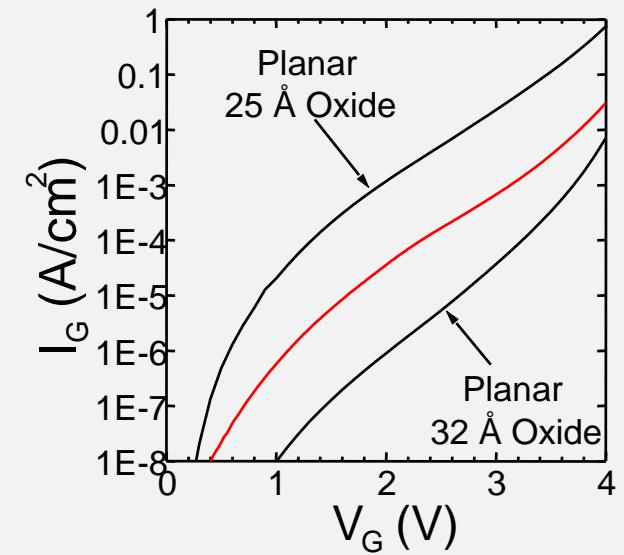
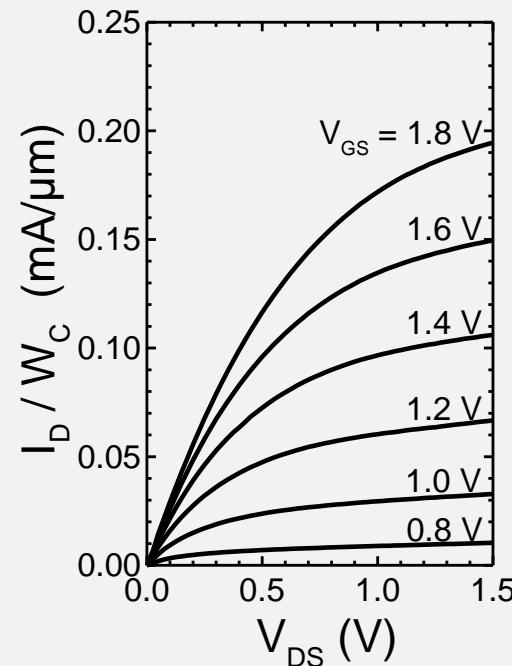
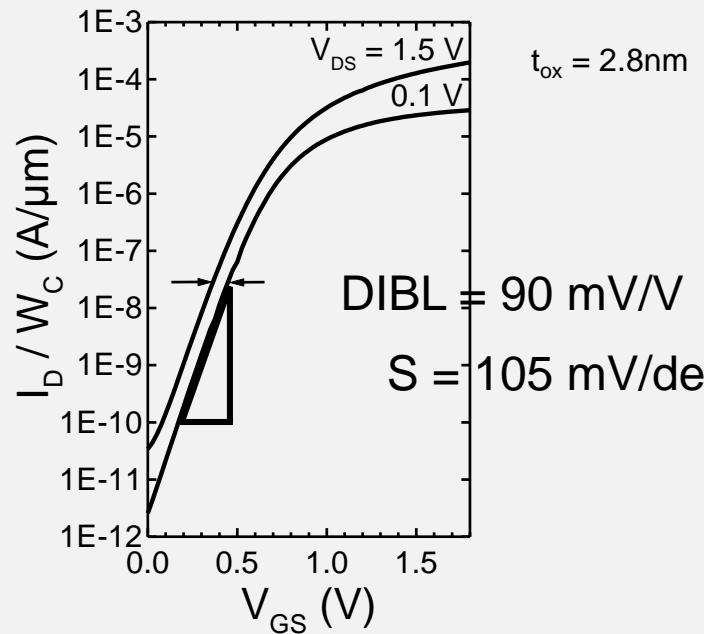
Source: Hansch et al, IEDM1997

## A Vertical Replacement Gate (VRG)-MOSFET

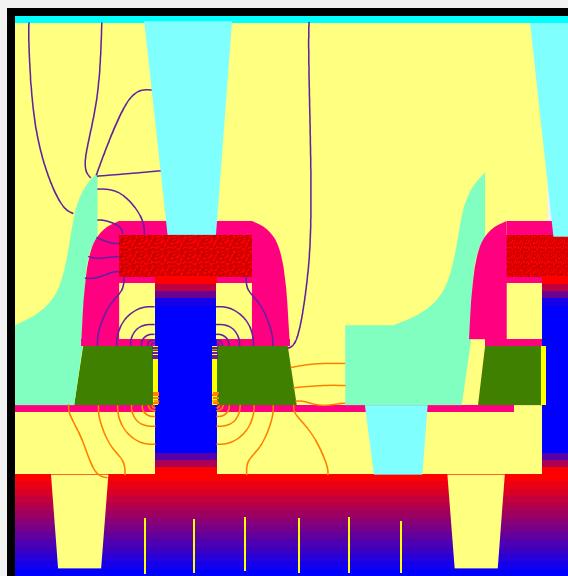
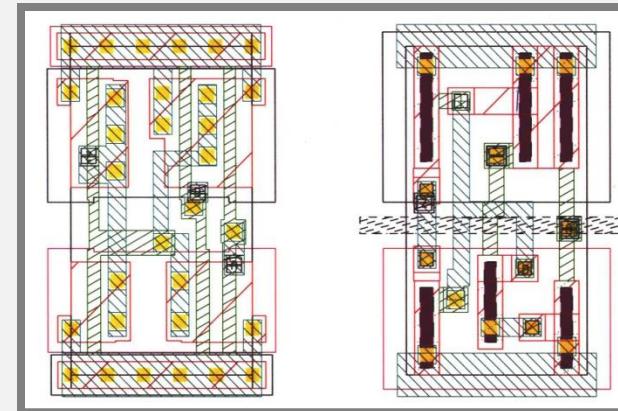
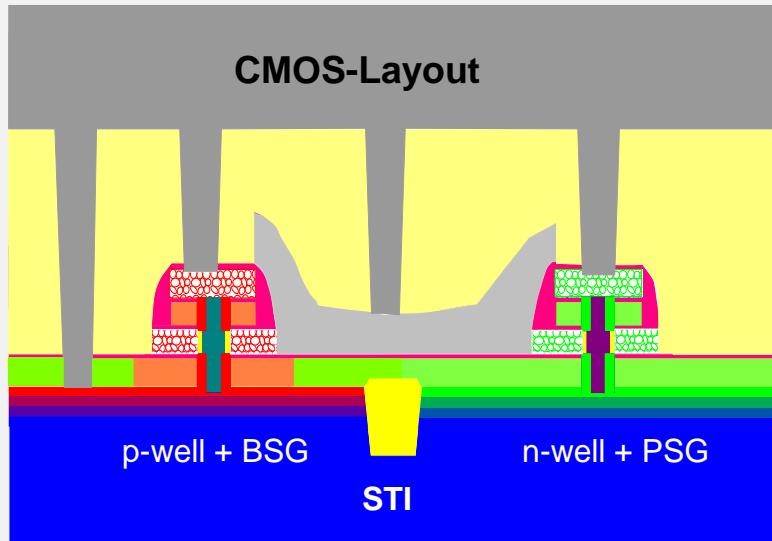
Gate replacement must be done,  
if temperature sensitive gate stacks  
(most high-k) will be used



Source: BELL-LABS, IEDM 1999



- All critical dimensions controlled without lithography
- Gate length = film thickness
- Gate oxide grown on Si
- Self-aligned SDEs formed by solid source diffusion
- Two sides of pillar drive in parallel
- Replacement-gate approach - enables alternative gate stacks
- Small parasitic capacitances should allow high-frequency operation
- Uses current manufacturing methods, materials, and tools



NAND gatter:

### VRG vs Planar:

- Similar packing density
- Up to double drive
- Parallel more dense
- Series less dense

### Overlap Capacitance

- Dominated by true overlap and local fringing
- Doubled width compensated by doubled drive current

### Substrate Capacitance

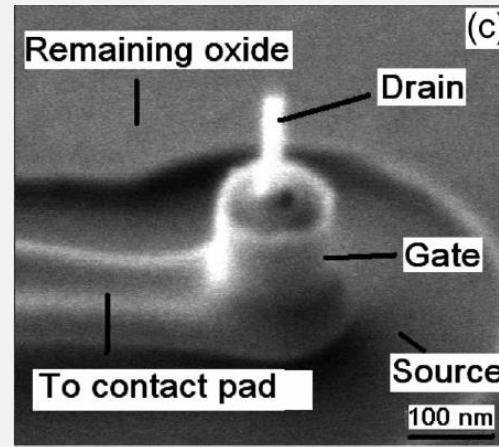
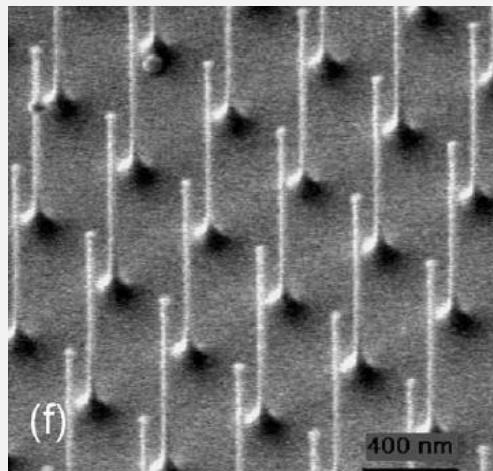
- Source: minimal
- Drain: reduced coefficient

An estimation shows that the parasitics are similar to SOI

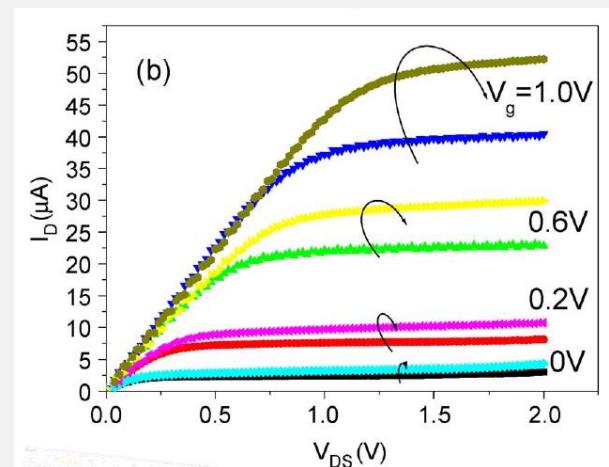
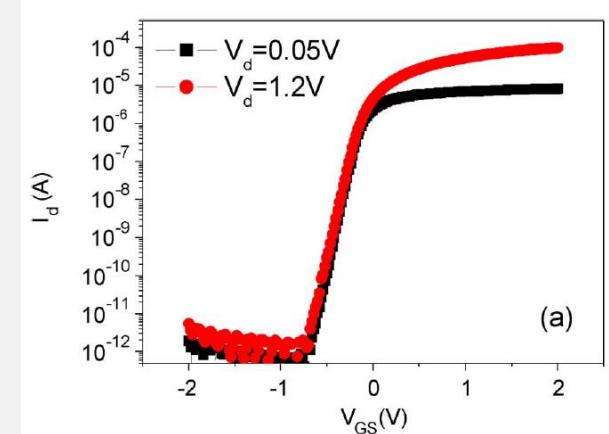
# Vertical Silicon-Nanowire Formation and Gate-All-Around MOSFET

IEEE 2008

B. Yang, K. D. Buddharaju, S. H. G. Teo, N. Singh, G. Q. Lo, and D. L. Kwong



$L_G = 150 \text{ nm}$   
 $D = 25 \text{ nm}$



Up to now no vertical MOSFETs are used in industry,  
mainly because planar MOSFETs could fulfill Moore's Law

Anyhow vertical MOSFETs are still under investigation.

## 6.1 Overview

Feature size, pitch, node

## 6.2 State of the Art MOSFETs

Electrical and geometrical parameters

## 6.3 Mobility Enhanced MOSFETs

Strain, SiGe, Liners, Directed Channels, Gate-Last

## 6.4 High-k Metal-Gate MOSFETs

Poly-Depletion, FuSi

## 6.5 Vertical MOSFETs

## 6.6 Halo MOSFETs

## 6.7 SOI-MOSFETs

## 6.8 Multi-Gate MOSFETs

Fully-depleted MOSFETs, FinFETs, Nanowire FETs

## 6.9 Advanced Memory Devices

In 1997, when the first "Roadmap" was created, the companies started to look several years in the future

Table 14 Memory and Logic Technology Requirements

Year of First Product Shipment Technology Generation	1997 250 nm	1999 180 nm	2001 150 nm	2003 130 nm	2006 100 nm	2009 70 nm	2012 50 nm
Min. Logic $V_{dd}$ (V) (desktop)	2.5–1.8	1.8–1.5	1.5–1.2	1.5–1.2	1.2–0.9	0.9–0.6	0.6–0.5
$V_{dd}$ Variation	≤ 10%	≤ 10%	≤ 10%	≤ 10%	≤ 10%	≤ 10%	≤ 10%
$T_{ox}$ Equivalent (nm)	4–5	3–4	2–3	2–3	1.5–2	< 1.5	< 1.0
Equivalent Maximum E-field (MV/cm)	4–5	5	5	5	> 5	> 5	> 5
Max $I_{off}$ @ 25°C (nA/ $\mu$ m) (For minimum L device)	1	1	3	3	3	10	10
Nominal $I_{on}$ @ 25°C ( $\mu$ A/ $\mu$ m) (NMOS/PMOS)	600/280	600/280	600/280	600/280	600/280	600/280	600/280
Gate Delay Metric (CV/I) (ps)*	16–17	12–13	10–12	9–10	7	4–5	3–4

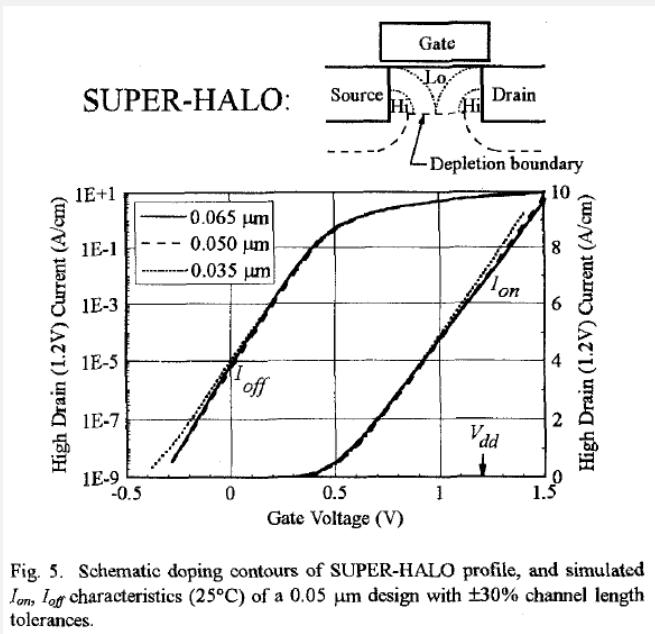
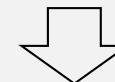


Fig. 5. Schematic doping contours of SUPER-HALO profile, and simulated  $I_{on}$ ,  $I_{off}$  characteristics (25°C) of a 0.05  $\mu$ m design with  $\pm 30\%$  channel length tolerances.

It was recognized:

- 1) because of leakage current the gate oxide can not be scaled,
- 2) the  $V_{DD}$  cannot be lowered, the electric fields will increase  
-> this will result in short-channel effects
- > the increase of S/D depletion zones under the gate must be stopped



First investigations (simulation) were given by:

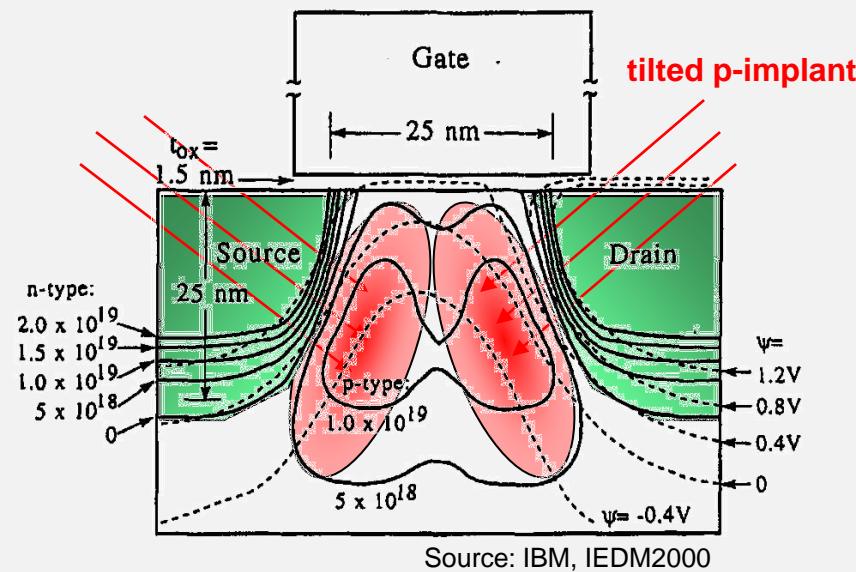
Taur et al. (IBM): "CMOS devices below 0.1  $\mu$ m: How high will performance go?", IEDM 1997, p.215-218

Introduction of high oppositely doped regions before Source/Drain

Halo Doping

or

Pocket Doping



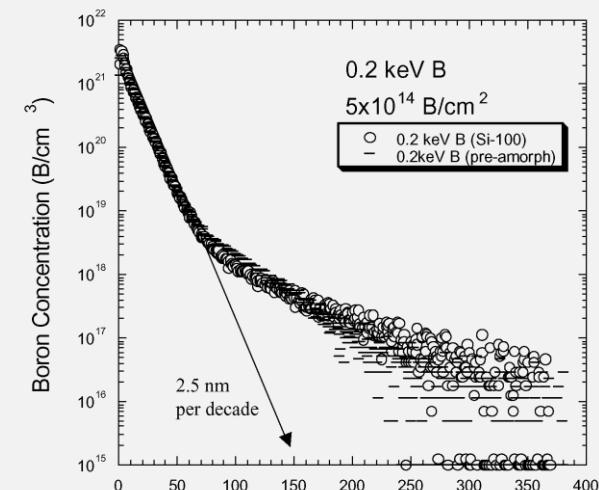
**Halo or pocket doping:**  
high, sharp bulk doping in front of Source/Drain  
to reduce Short Channel Effects

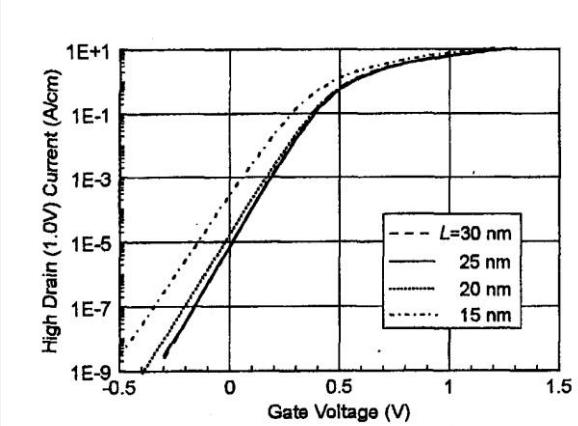
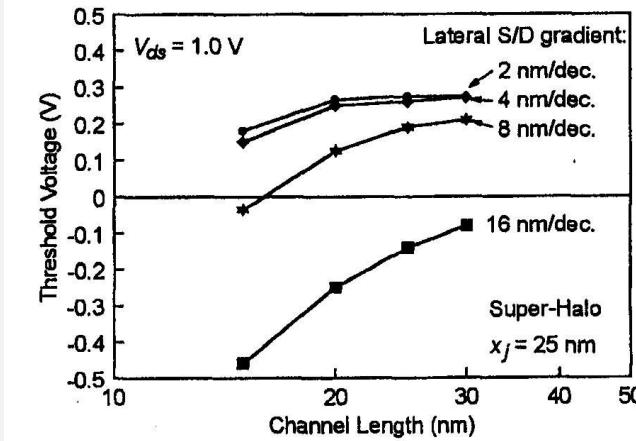
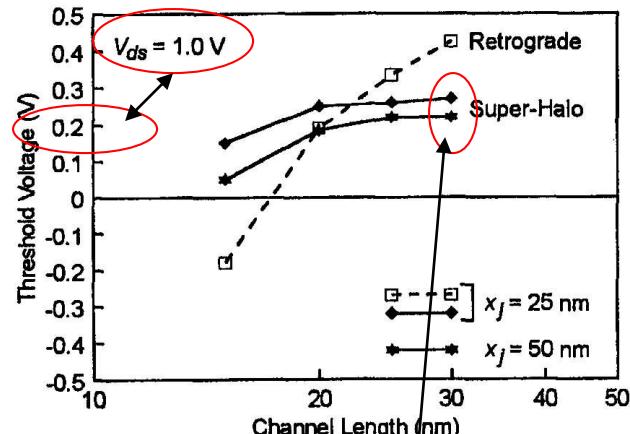
- ▶ stopping the width of Source/Drain depletion zones  
classical long channel limit:  $w_{SD} < 1/5 L$
- ▶ very abrupt implant profiles (~2nm)  
very sensitive high temperature treatment (-> diffusion)
- ▶ flash heating (-> 1000°C in 1 sec) for implant recrystallization

- ▶ decoupling of  $V_{DD}$ , threshold voltage  $V_{th}$  and  $I_{off}$
- ▶ problem: n+/p+ band-band tunneling starts -> additional  $I_{off}$

vertically and laterally nonuniform channel doping profile

- ▶ - optimization only via simulations
- ▶ - **classical scaling rules are overruled**  
(see  $t_{ox}=1.5\text{nm}$ ,  $V_{DD}=1.0\text{V}$  at 25nm channel length)





### Advantage:

with HALO-doping the depth of Source/Drain junctions is not very critical

- leaving scaling rules
- Source/Drain resistance keeps low
- S/D extensions (if any) with low resistance
- decoupling of  $V_{DD}$  and  $V_{th}$

### important:

sharp doping profiles (nm/dec) are needed

- questionable from implant process  
(statistics in range straggling, under-implant)
- flash heating (1000°C for a few sec is needed)  
(recrystallization of implant damage, but no diffusion)

### Electrical properties:

- $S \sim 200 \text{ mV/dec}$
- $\rightarrow$  for  $I_{on}/I_{off} \sim 10^5$  1V swing is needed
- >  $V_{DD}=1\text{V}$  and  $V_{th}=0.2\text{V}$



Feasibility of 25 nm channel length (proposed for year 2012) planar bulk MOSFETs was shown in 1998



Halo doping is used till 22/20 nm technology node (Samsung/TSMC planar technology)



For fully depleted channels like FinFET Halo doping not necessary

## 6.1 Overview

Feature size, pitch, node

## 6.2 State of the Art MOSFETs

Electrical and geometrical parameters

## 6.3 Mobility Enhanced MOSFETs

Strain, SiGe, Liners, Directed Channels, Gate-Last

## 6.4 High-k Metal-Gate MOSFETs

Poly-Depletion, FuSi

## 6.5 Vertical MOSFETs

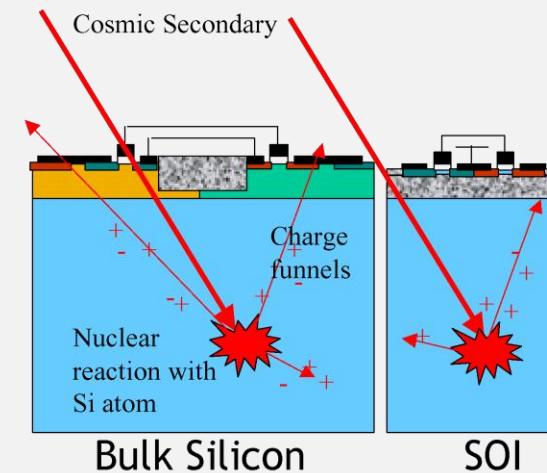
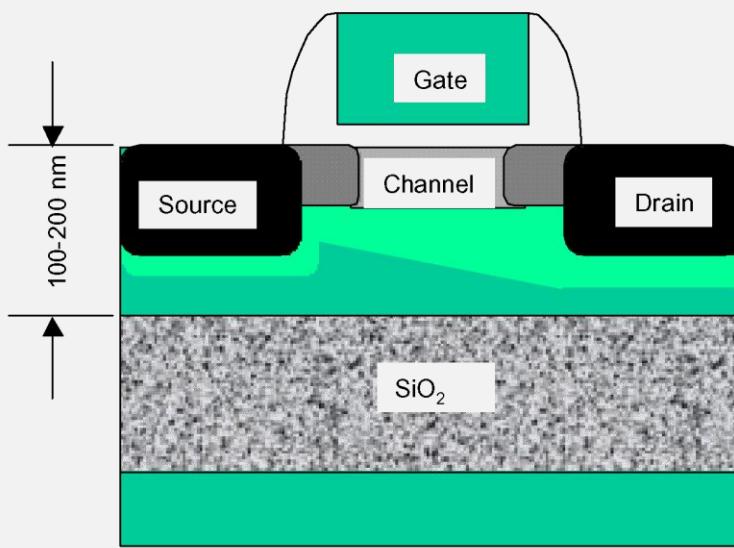
## 6.6 Halo MOSFETs

## 6.7 SOI-MOSFETs

## 6.8 Multi-Gate MOSFETs

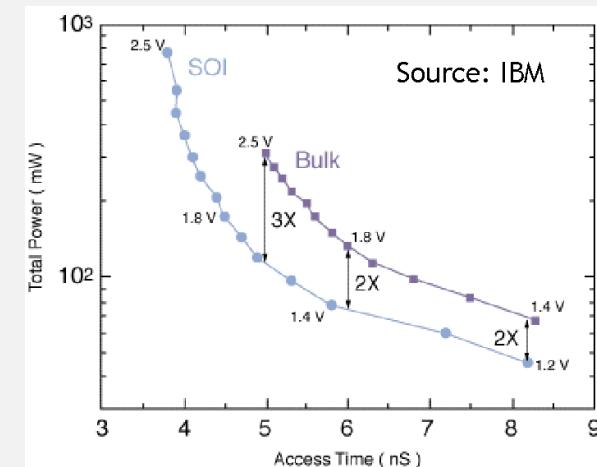
Fully-depleted MOSFETs, FinFETs, Nanowire FETs

## 6.9 Advanced Memory Devices



**Advantage of SOI:**

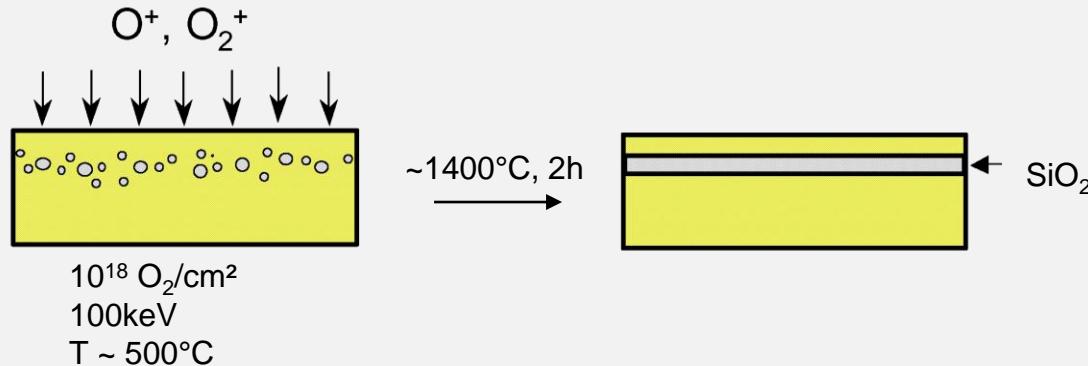
- ▶ easy lateral isolation
- ▶ radiation tolerance
- ▶ lower leakage current -> lower power consumption
- ▶ lower parasitic capacitance -> higher speed



Classical SOI-MOSFETs show various advantages in performance compared to classical bulk-MOSFET, but are more expensive

Several techniques are established to fabricate SOI-wafers:

1) SIMOX: Separation by Implanted Oxygen



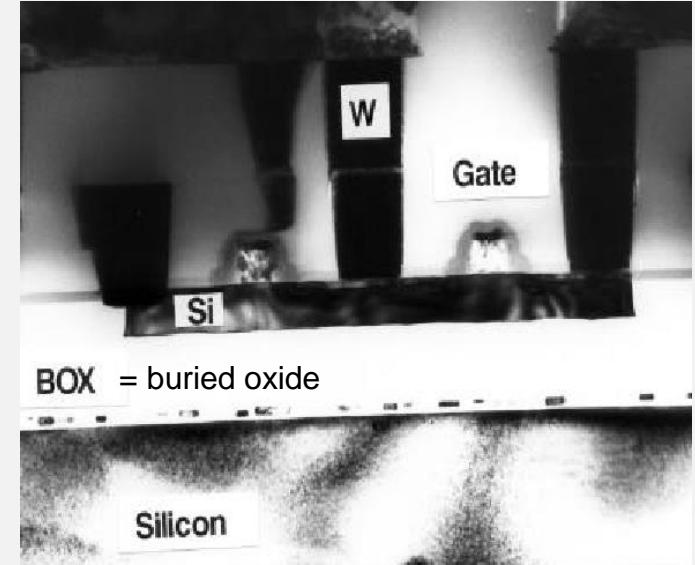
2) BESOI: Bonded etch-back silicon on insulator

3) other separation techniques, like Nanocleave, Smartcut,....

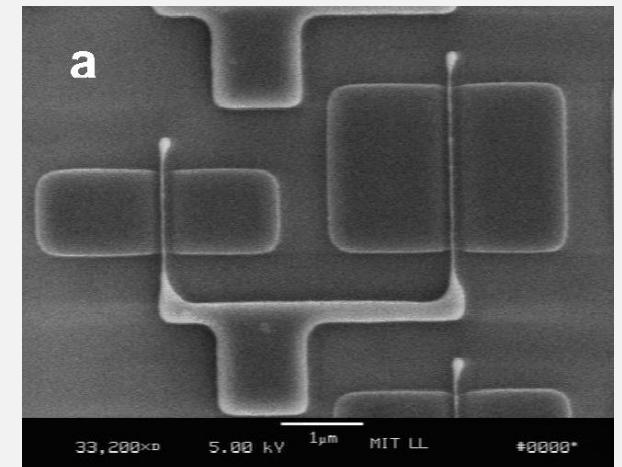


Wafer costs are 10 x standard wafer

Silicon layers down to 10nm are demonstrated



MOSFETs fabricated on SOI



1

The position of Fermi-Level can be calculated (chap. 3.13):

$$\Psi_{bulk} = kT \cdot \ln \left( \frac{N_{doping}}{n_i} \right)$$

2

The barrier in a pn-junction can be calculated:

$$V_{bi} = \Psi_n + \Psi_p = \frac{kT}{q} \cdot \left( \ln \left[ \frac{N_D}{n_i} \right] + \ln \left[ \frac{N_A}{n_i} \right] \right) = \frac{kT}{q} \ln \left[ \frac{N_A \cdot N_D}{n_i^2} \right]$$

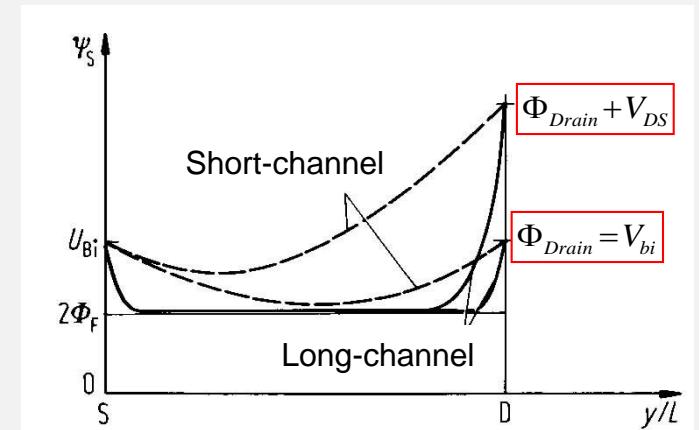
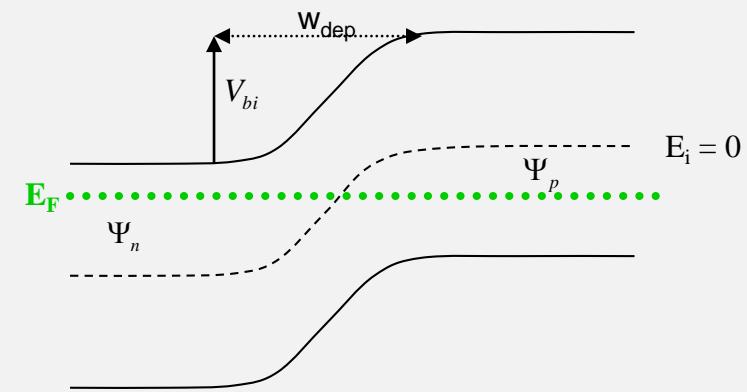
and the width of the (one-side abrupt) depletion zone will be:

$$w = \sqrt{\frac{2\epsilon_0 \epsilon_{Si} \cdot (V_{bi} \pm V_{ext})}{qN_{dop}}}$$

3

Calculating the threshold voltage, we receive:

$$V_t = V_{FB} + 2\Psi_{bulk} + \frac{\sqrt{2\epsilon_0 \epsilon_{Si} \cdot qN_{dop} \cdot (\Phi_{Drain} + V_{ext})}}{C_{ox}} \cdot \left( 1 - \frac{\epsilon_0 \epsilon_{Si}}{qN_{dop}} \cdot \frac{(\Phi_{Drain} + V_{DS})}{L_{eff}^2} \right)$$



We receive for the threshold voltage:

$$V_t = V_{t, \text{long}} - SCE - DIBL$$

$$SCE = 0.64 \cdot \frac{\epsilon_0 \epsilon_{Si}}{\epsilon_{ox}} \cdot \left( 1 + \frac{x_j^2}{L_{el}^2} \right) \cdot \frac{t_{ox,el}}{L_{el}} \cdot \frac{w_{dep}}{L_{el}} \cdot \Phi_{\text{drain}} \quad \Rightarrow \quad SCE \approx 2 \cdot EI \cdot \Phi_{\text{drain}}$$

$$DIBL = 0.80 \cdot \frac{\epsilon_0 \epsilon_{Si}}{\epsilon_{ox}} \cdot \left( 1 + \frac{x_j^2}{L_{el}^2} \right) \cdot \frac{t_{ox,el}}{L_{el}} \cdot \frac{w_{dep}}{L_{el}} \cdot V_{DS} \quad \Rightarrow \quad DIBL \approx 2.5 \cdot EI \cdot V_{DS}$$

The geometrical factors we define as **Electrical Integrity EI**:

$$EI = \left( 1 + \frac{x_j^2}{L_{el}^2} \right) \cdot \frac{t_{ox,el}}{L_{el}} \cdot \frac{w_{dep}}{L_{el}}$$

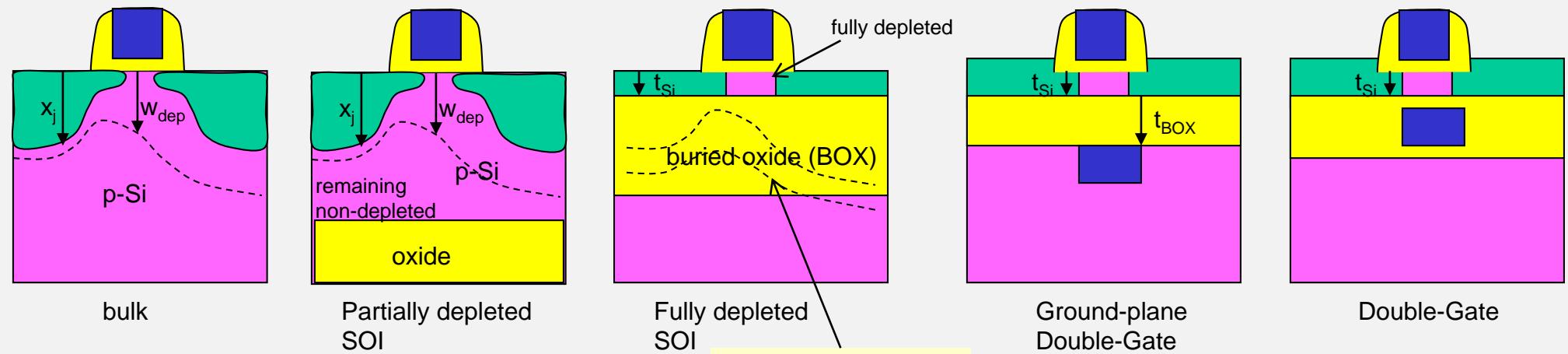
From EI-equation we can follow:

- SCE and DIBL will be small, if:  $L_{el} \rightarrow \text{large}$  of course, because we have a long channel device
- $t_{ox} \rightarrow \text{small}$  better gate control  $\rightarrow$  less reduction of threshold voltage
- $x_j$  and  $w_{dep} \rightarrow \text{small}$  we already learned from scaling (chap.4)

Around the year 2000 it was recognized:

If we use thin Si on insulator (SOI), then the thickness of  $x_j$  and  $w_{dep}$  will be defined by silicon thickness

see next page



$$EI = \left(1 + \frac{x_j^2}{L_{el}^2}\right) \cdot \frac{t_{ox,el}}{L_{el}} \cdot \frac{w_{dep}}{L_{el}}$$

$$EI = \left(1 + \frac{x_j^2}{L_{el}^2}\right) \cdot \frac{t_{ox,el}}{L_{el}} \cdot \frac{w_{dep}}{L_{el}}$$

$$EI = \left(1 + \frac{t_{Si}^2}{L_{el}^2}\right) \cdot \frac{t_{ox,el}}{L_{el}} \cdot \frac{t_{Si} + \lambda t_{BOX}}{L_{el}}$$

remaining contribution of partially depleted BOX

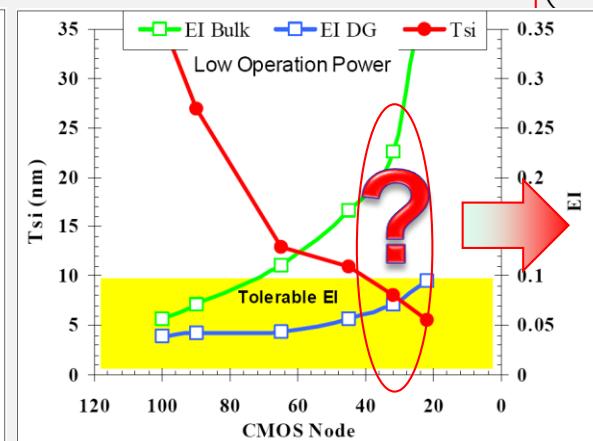
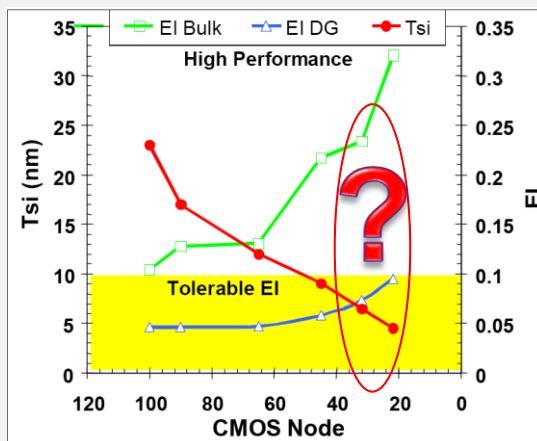
$$EI = \left(1 + \frac{t_{Si}^2}{L_{el}^2}\right) \cdot \frac{t_{ox,el}}{L_{el}} \cdot \frac{t_{Si} + t_{BOX}}{L_{el}}$$

$$EI = 0.5 \cdot \left(1 + \frac{t_{Si}^2/4}{L_{el}^2}\right) \cdot \frac{t_{ox,el}}{L_{el}} \cdot \frac{t_{Si}/2}{L_{el}}$$

From roadmap small SCE and DIBL are requested

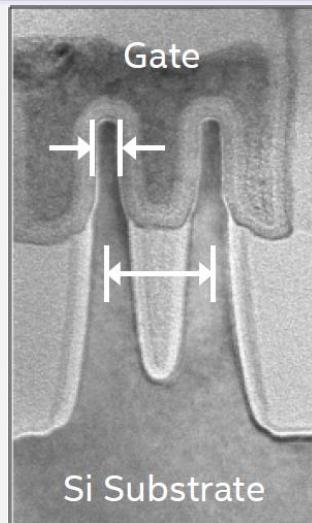
Requirement:  
 $EI < 10\% \rightarrow DIBL < 25\%$   
(see eq. DIBL)

Only Double-Gate can fulfill requirements



Double-Gate (FinFET) since 2012 (22nm)

2014 Node (FinFET): 14 nm



- Fin width 8 nm
- Fin pitch 42 nm
- Assumption:
  - $L_{el} = 14 \text{ nm}$
  - $t_{ox,el} = 0.8 \text{ nm}$

Calculation of the Electrical Integrity

$$EI = 0.5 \cdot \left( 1 + \frac{t_{Si}^2 / 4}{L_{el}^2} \right) \cdot \frac{t_{ox,el}}{L_{el}} \cdot \frac{t_{Si} / 2}{L_{el}}$$

$$EI = 0.5 \cdot \left( 1 + \frac{(8\text{nm})^2 / 4}{(14\text{nm})^2} \right) \cdot \frac{0.8\text{nm}}{14\text{nm}} \cdot \frac{8\text{nm} / 2}{14\text{nm}} = 0.0088$$

Calculation of DIBL

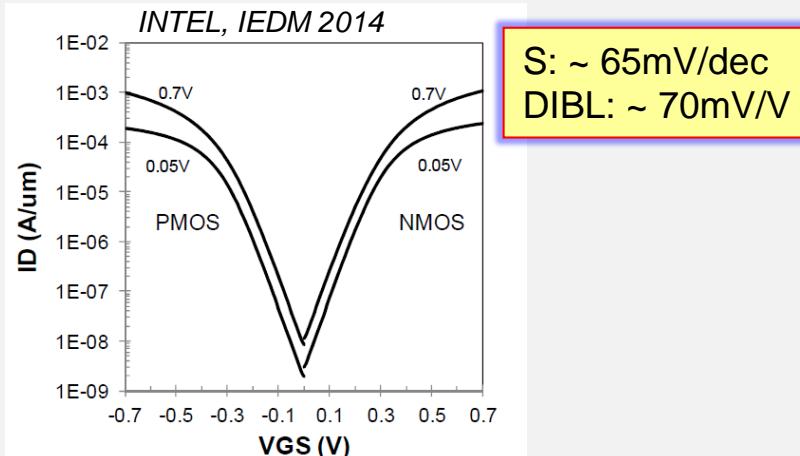
$$DIBL \approx 2.5 \cdot EI \cdot V_{DS}$$

$$DIBL \approx 2.5 \cdot 0.0088 \cdot 0.7V \approx 15.4mV/V$$



State-of Art FinFET is suppressing Short-channel effects

The actual DIBL value is in the range of the theory value



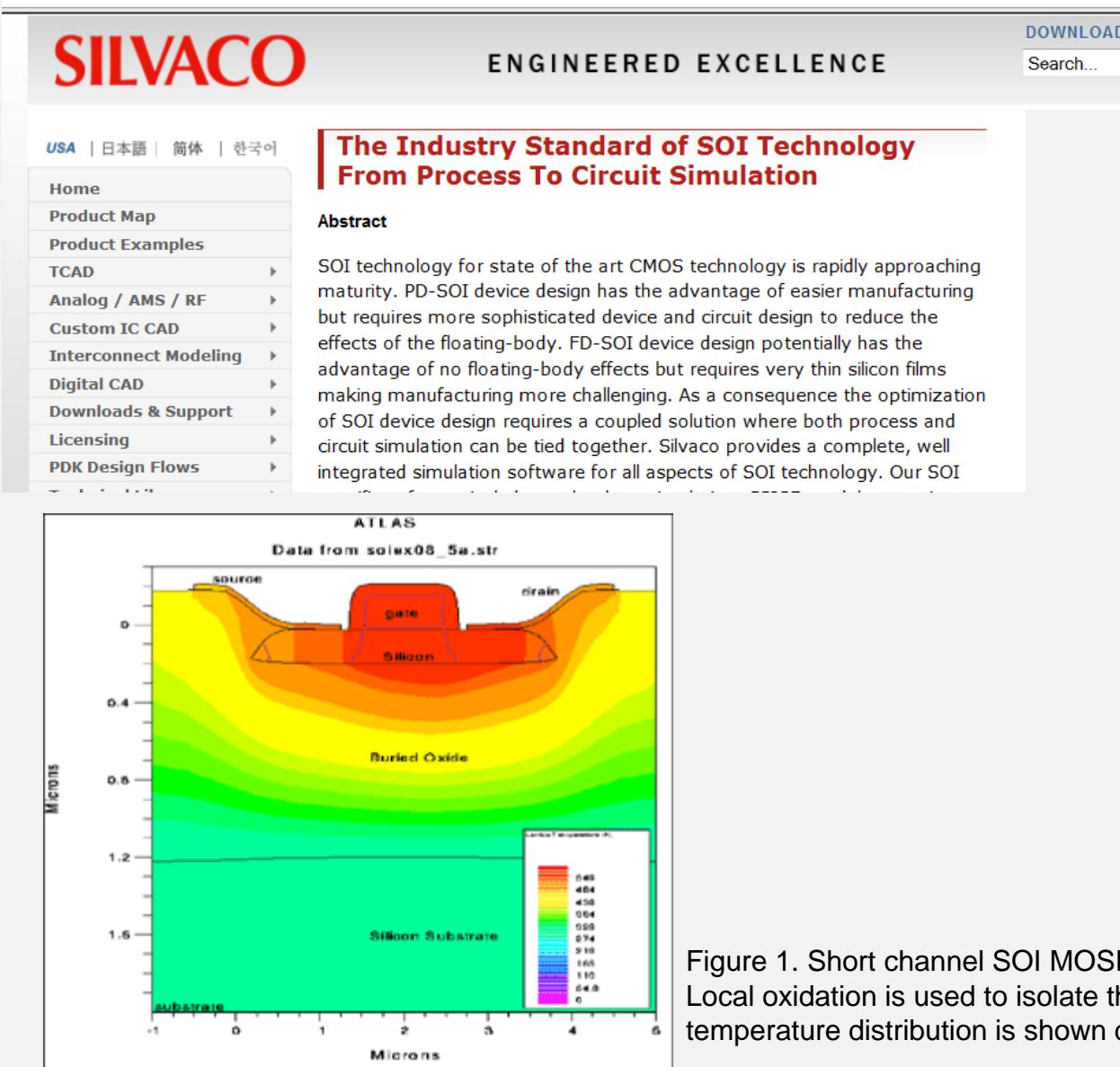


Figure 1. Short channel SOI MOSFET created using **ATHENA**. Local oxidation is used to isolate the active layer. Here temperature distribution is shown due to self heating.

## 6.1 Overview

Feature size, pitch, node

## 6.2 State of the Art MOSFETs

Electrical and geometrical parameters

## 6.3 Mobility Enhanced MOSFETs

Strain, SiGe, Liners, Directed Channels, Gate-Last

## 6.4 High-k Metal-Gate MOSFETs

Poly-Depletion, FuSi

## 6.5 Vertical MOSFETs

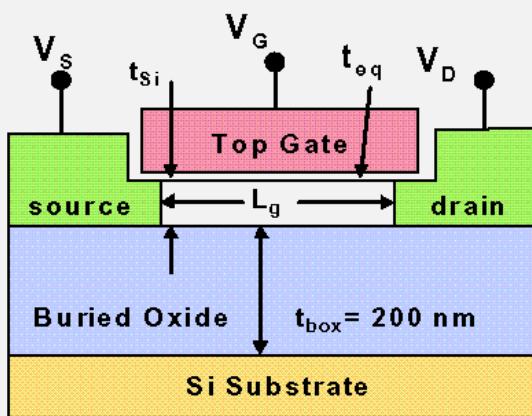
## 6.6 Halo MOSFETs

## 6.7 SOI-MOSFETs

## 6.8 Multi-Gate MOSFETs

Fully-depleted MOSFETs, FinFETs, Nanowire FETs

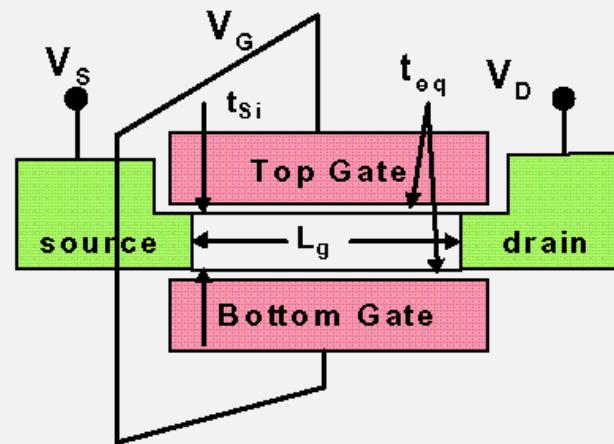
## 6.9 Advanced Memory Devices



**Single-gate SOI**

Advantage:

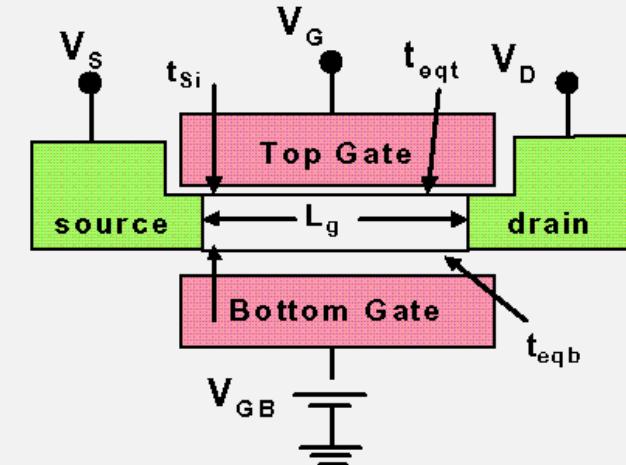
- reduction of leakage current through bulk



**Double-gate**  
2 equal gates

Advantage:

- for very thin Si-layers < 5nm the  $V_{th}$  roll-off (**fully depleted "bulk"**, all Si in inversion) can be avoided down to 25nm channel length



**Double-Gate, ground-plane**

2 different gates

Advantage:

- in addition to double-gate the threshold voltage can be adjusted

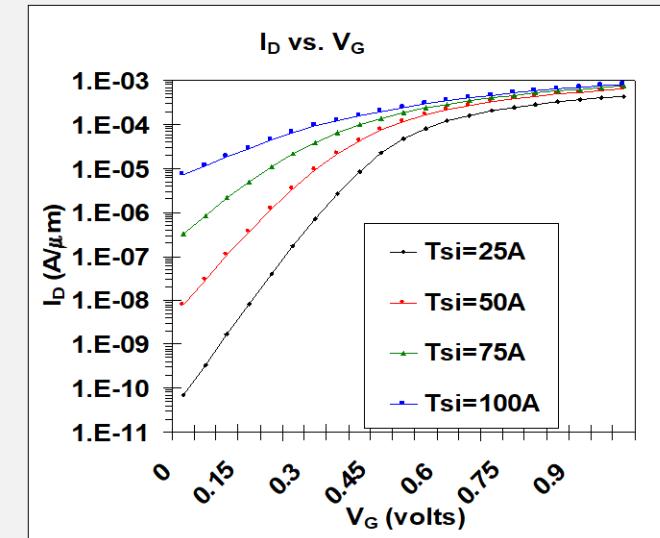
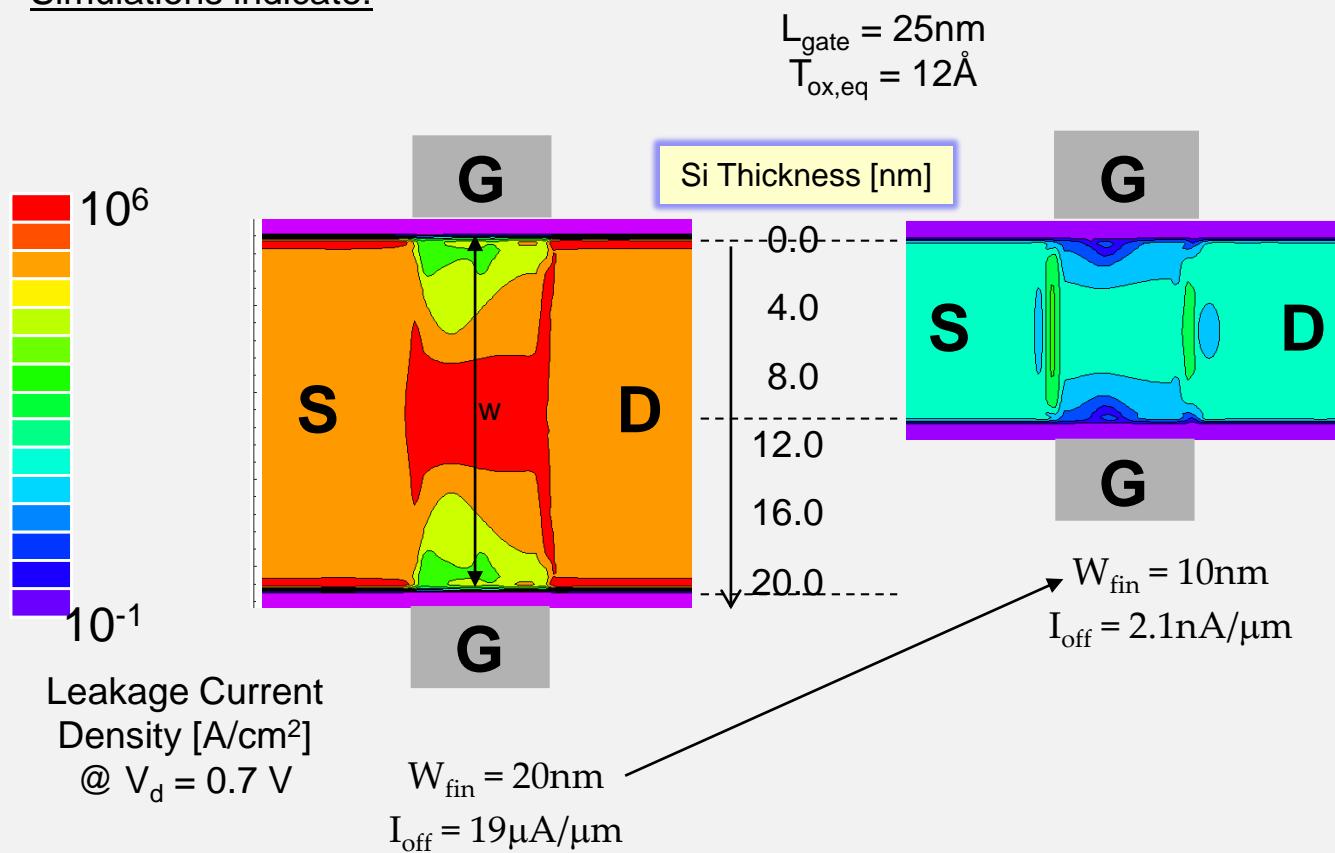


In planar technique the fabrication of a back-gate (oxide with nm thickness and good interface quality) is not easy



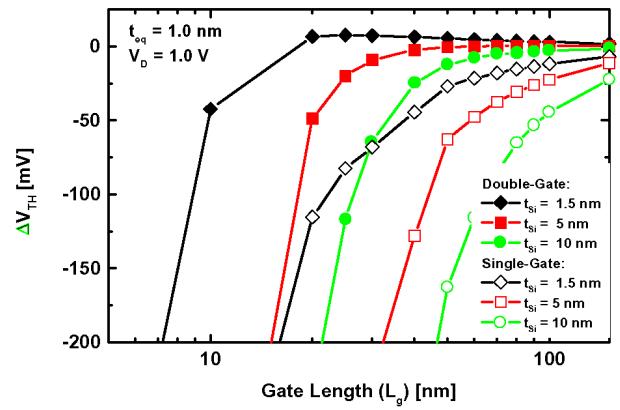
from STMicroelectronics an approach "Silicon on Nothing, SON" exists

Simulations indicate:

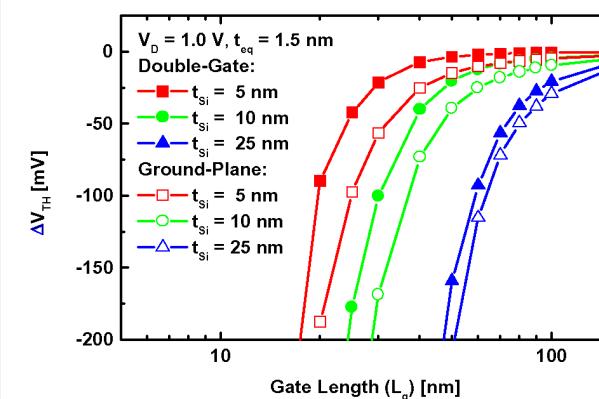


Leakage current is greatly reduced with thinner Si-body

## Simulations

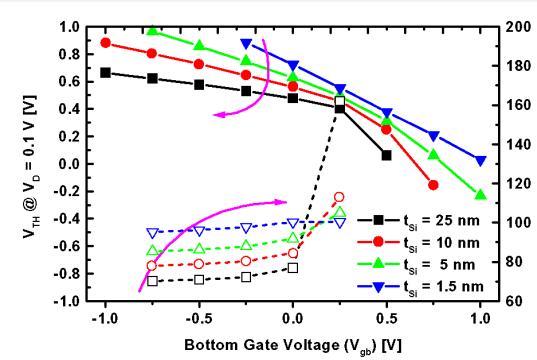
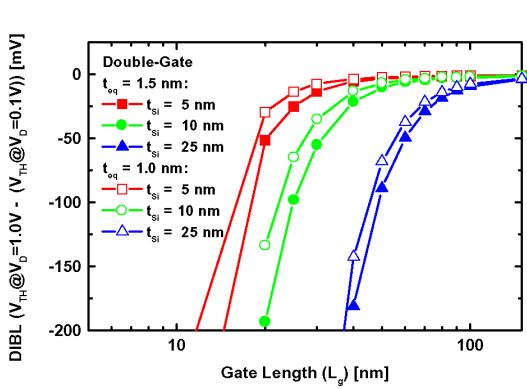


Threshold voltage roll-off for double-gate and single-gate MOSFET



Threshold voltage roll-off for double-gate and ground-plate MOSFET is similar

## Simulations



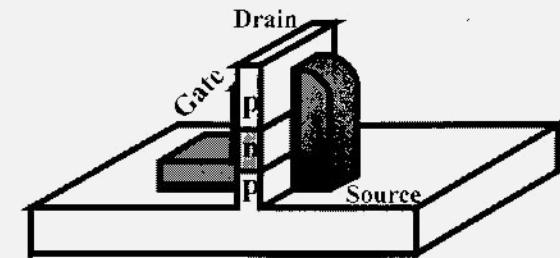
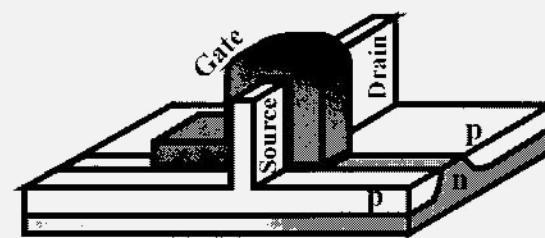
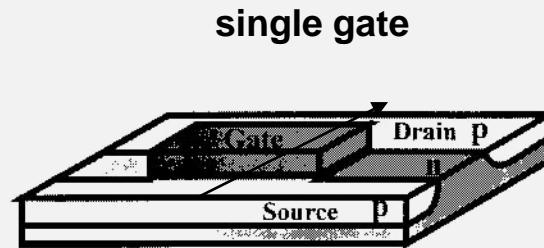
Tunable threshold voltage with ground-plate gate.  
At larger  $V_{GP}$  an additional inversion channel is created

single-gated MOSFETs will not meet short-channel effect requirements

Short-channel effects are reduced in thinner silicon layers -> fully depleted

oxide thickness is not very sensitive -> overruling of scaling rules

## Concepts

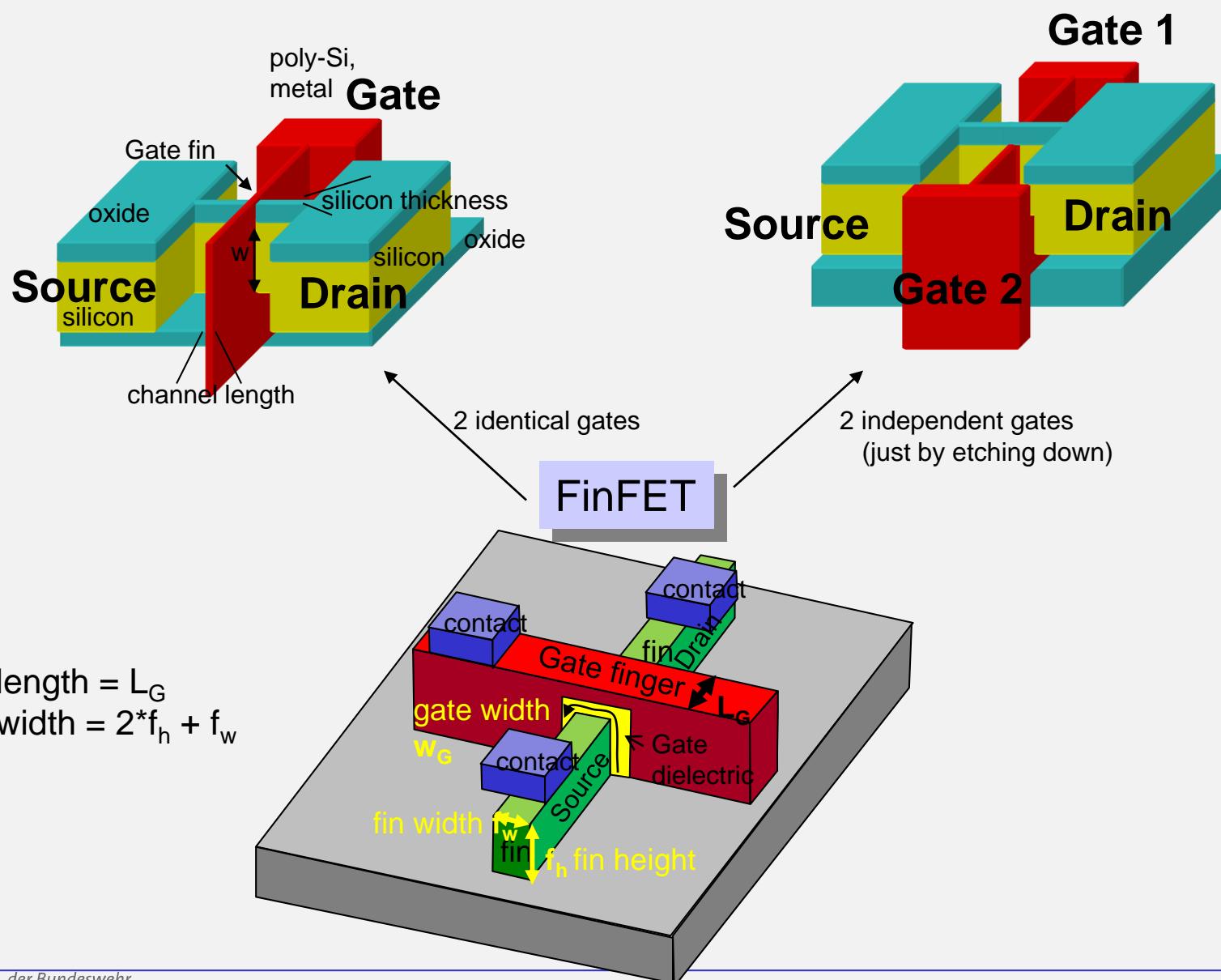


**planar**

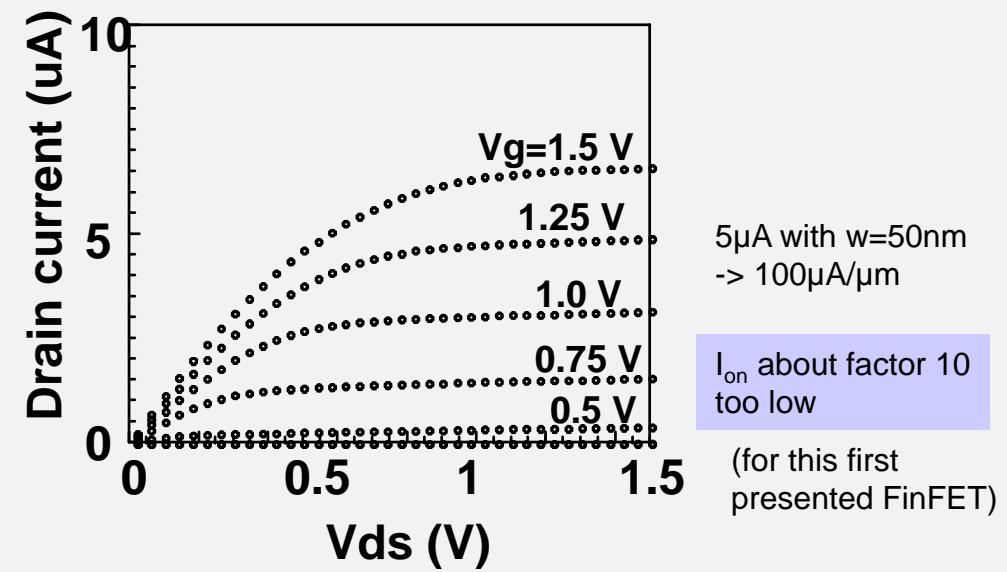
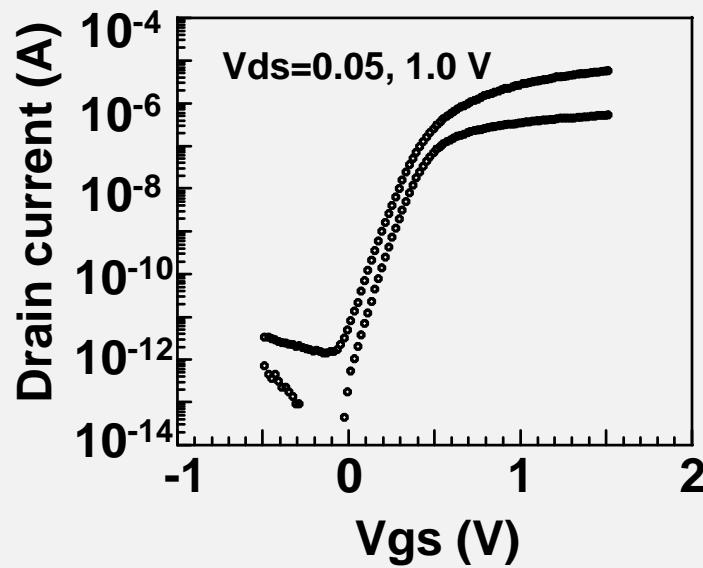
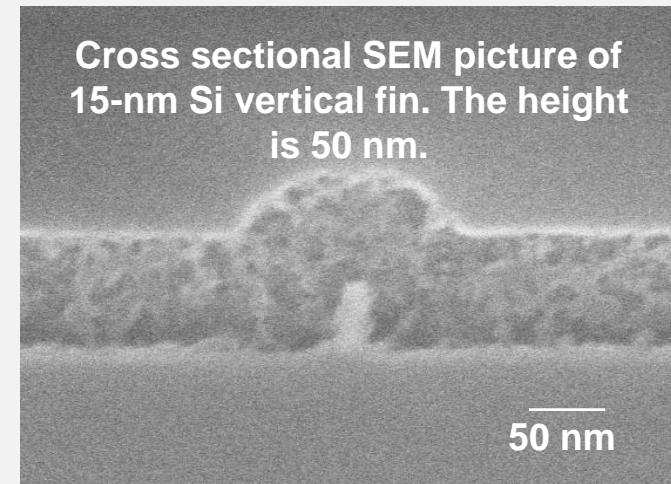
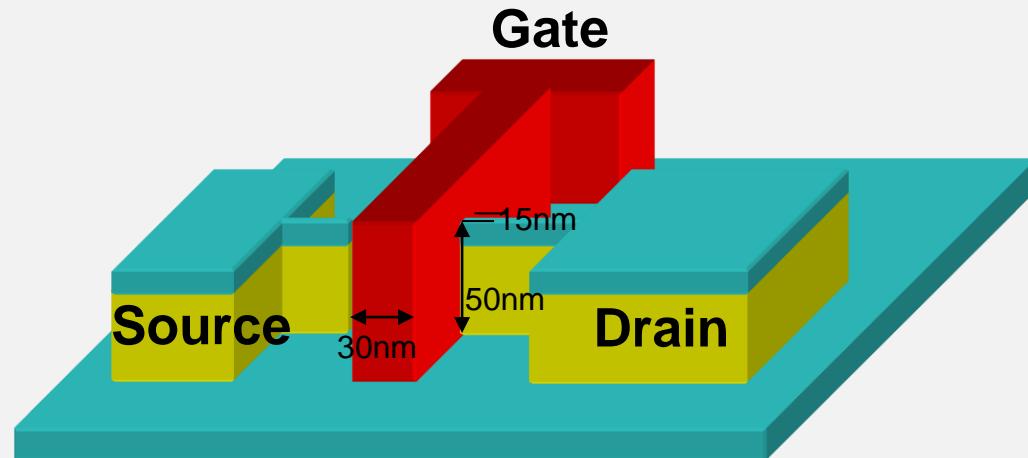
**vertical**

**vertical pillar**

- ↳ **planar** : planar structure and planar current flow
- ↳ **vertical** : vertical structure but planar current flow
- ↳ **vertical pillar** : vertical structure and current flow

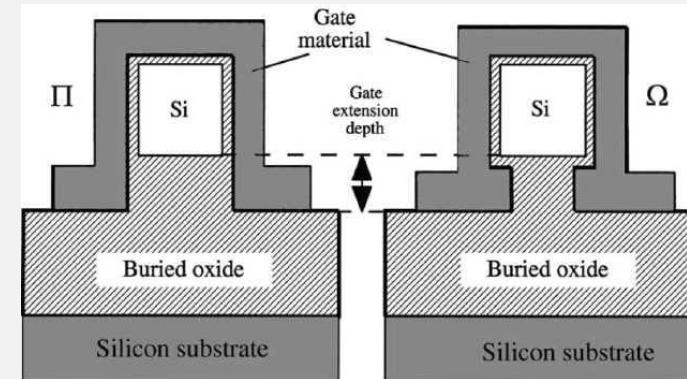


Hisamoto et al., IEDM 1998



In early days several geometrical designs due to special process flow were developed and various names were proposed, since some of the early names were protected by patent laws:

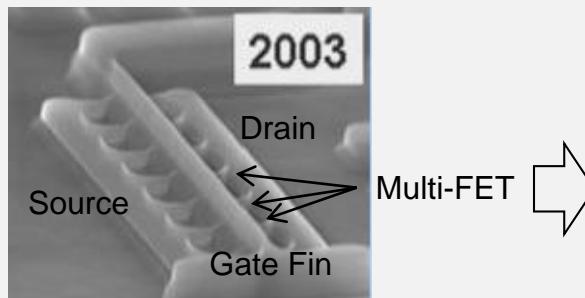
Pi - Gate      Omega - Gate



one S/D body with one FIN

- > resulting in 2 sidewall gates -> double-gate
- > or in 1 top-gate and 2 sidewall gates -> Tri-gate

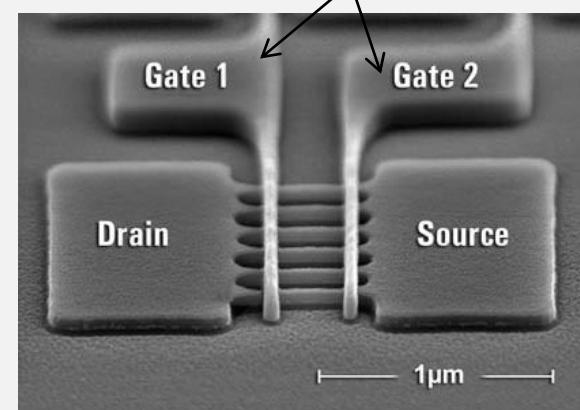
Multi-Transistor FinFET



multi S/D bodies with one FIN

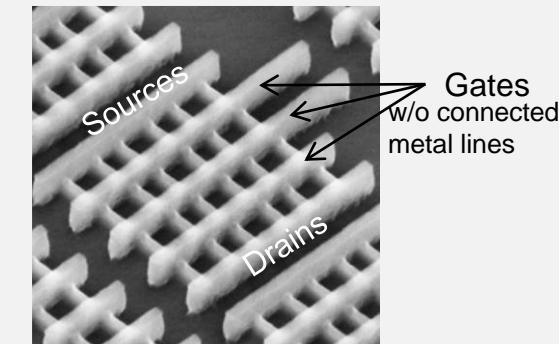
- > resulting in 2 sidewall gates -> double-gate
- > or in 1 top-gate and 2 sidewall gates -> Tri-gate

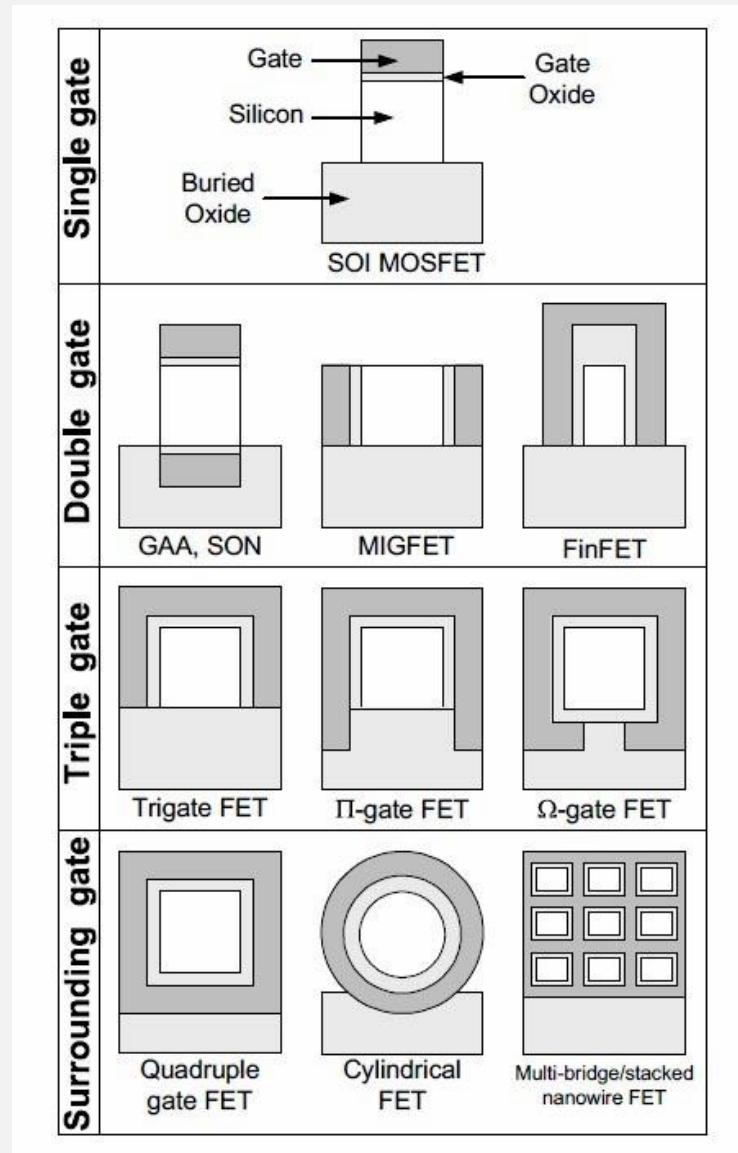
Multi-Transistor Multi-Gate FinFET



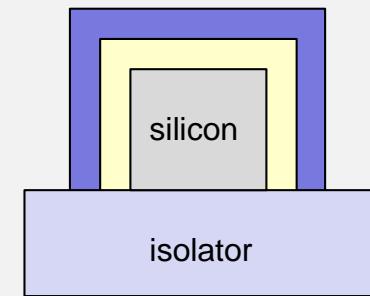
Easy design of functions like AND, XOR, SRAM

Multi-Transistor Multi-Gate FinFET



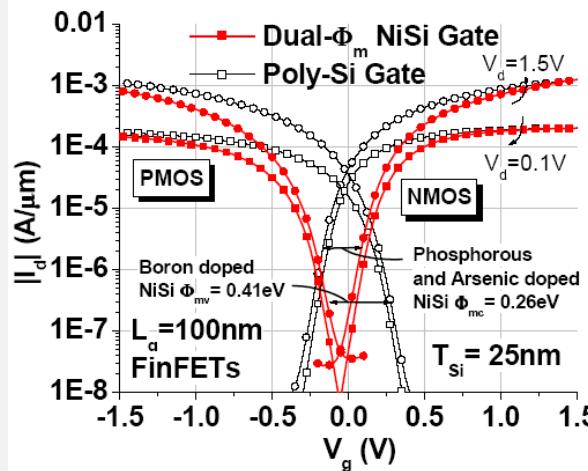
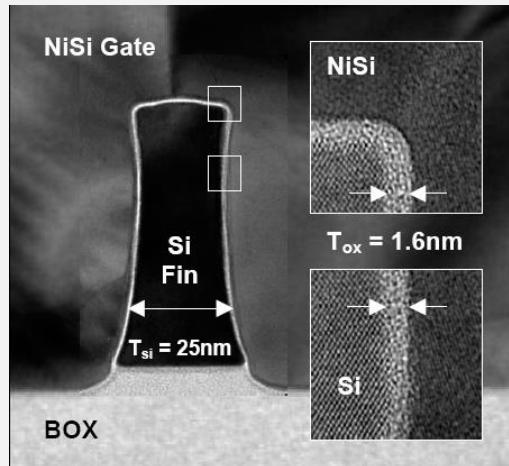


looking for the cross section various other names were invented:



**IBM** IEDM 2002

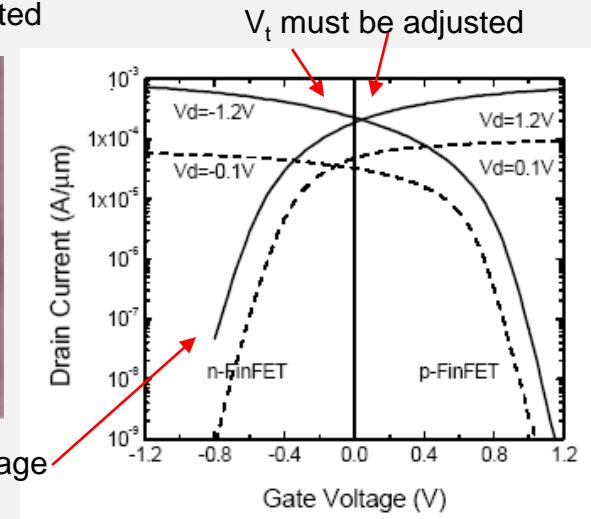
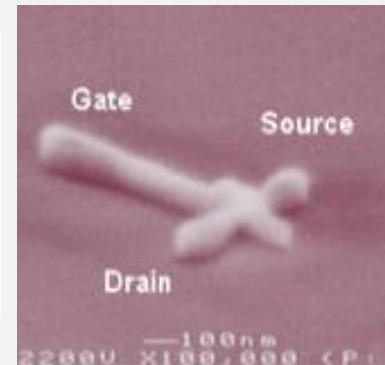
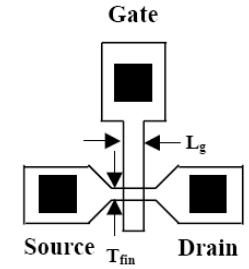
$L > 100\text{nm}$ ,  $T_{Si} = 25\text{nm}$



Demonstration of Metal Gate

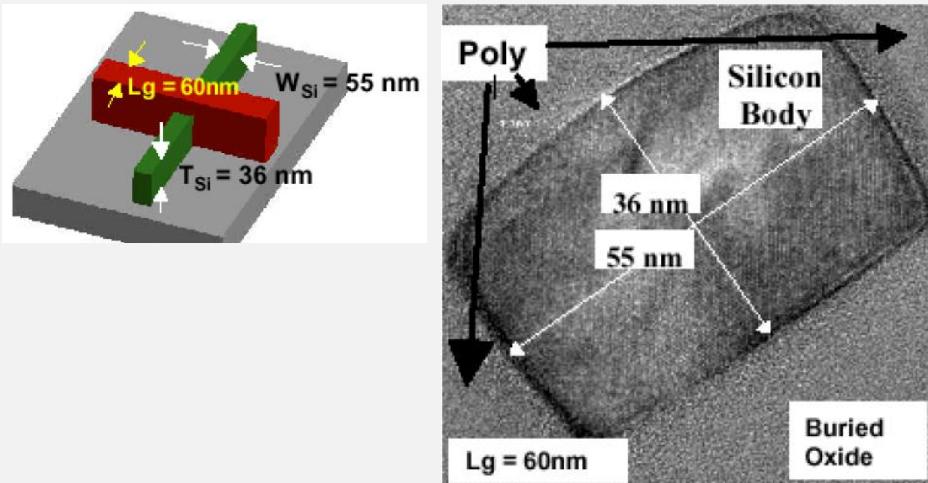
**AMD** IEDM 2002

$L = 10\text{nm}$  shortest channel length demonstrated  
 $T_{Si} = 12\text{nm}$

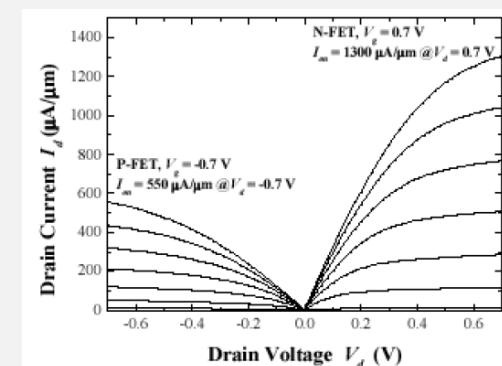
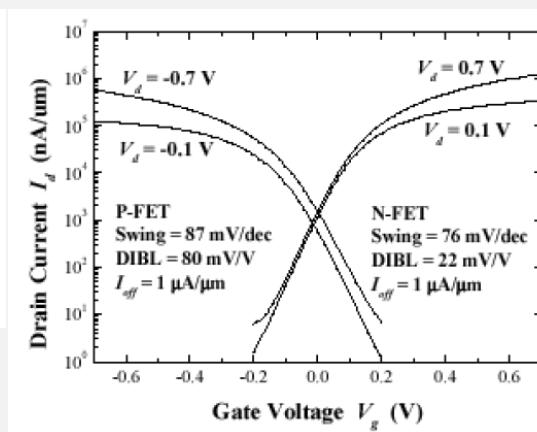
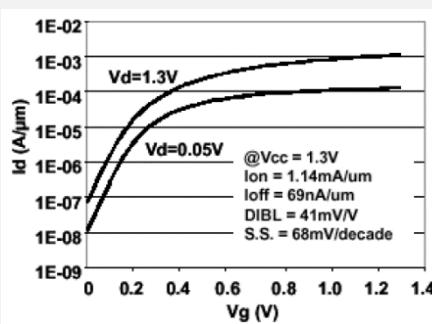
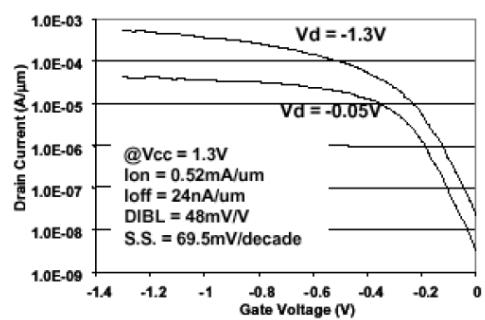
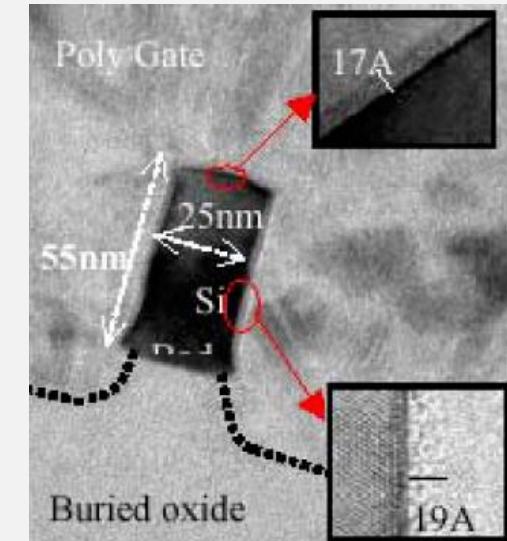
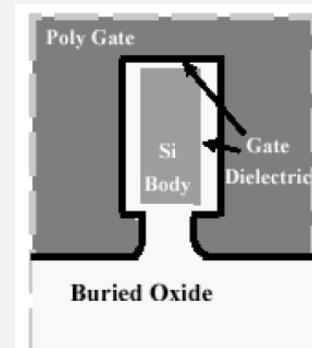


Transistor Parameter	This Work FinFET	This Work FinFET	Ref. [5] FinFET (IEDM'01)	Ref. [7] FinFET (VLSI'02)	Ref. [3] FinFET (DRC'01)	Ref. [6] FinFET (IEDM'01)	Ref. [4] FinFET (DRC'01)
$L_g$ (nm)	10	55	20	35	60	60	100
$V_{dd}$ (V)	1.2	1.2	1	1	1	1.5	1.5
$T_{ox}(\text{phys})$ (A)	17	17	21	24	25	16	22
Gate Electrode	Poly-Si	Poly-Si	N+ SiGe	in-situ N+	N+ SiGe	Poly-Si	Poly-Si
Swing, N (mV/dec)	125	64	75	78	70	66	72
Swing, P (mV/dec)	101	68	90	96	80	65	N/A
DIBL, N (mV/V)	71	11	100	N/A	40	N/A	57
DIBL, P (mV/V)	120	27	140	N/A	90	N/A	N/A

INTEL, Tri-Gate (SSDM 2002)

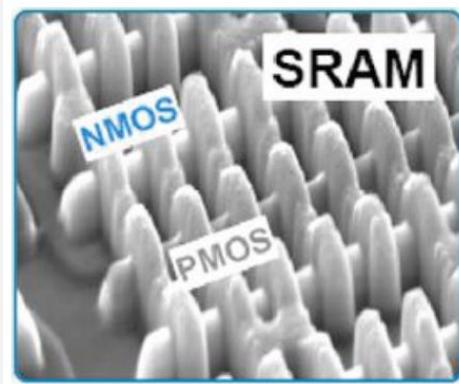
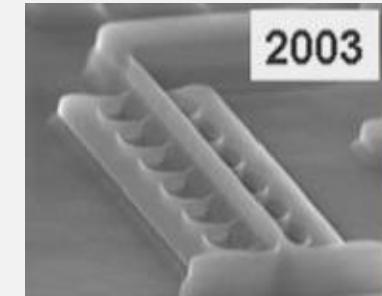
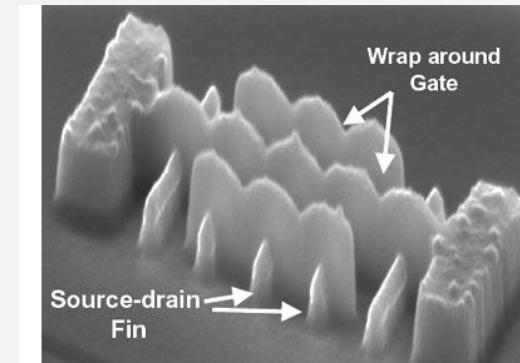


TSMC,  $\Omega$ -Gate (IEDM 2002)



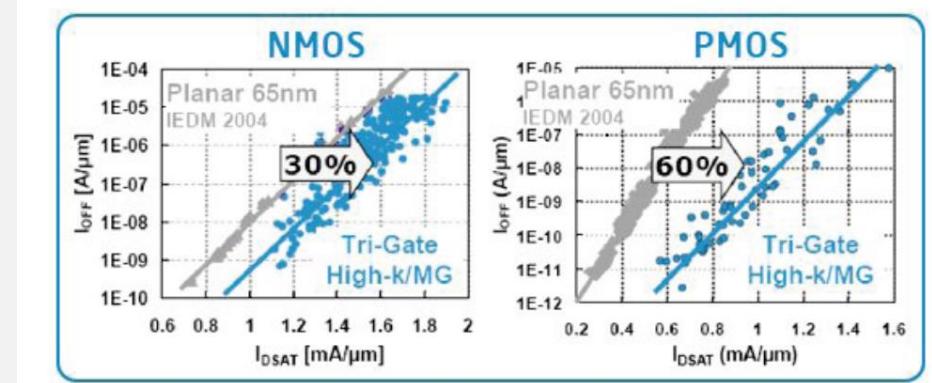


Single-Gate-FET



SRAM

the tri-gate architecture provides more device width for a given cell footprint compared to the standard planar transistor, thus providing a higher read current because total current is a direct function of the total device width



Source: Technology@Intel Magazine 2006

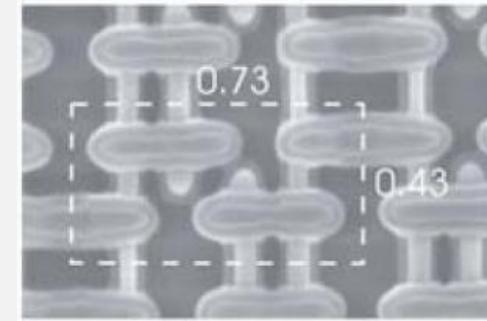
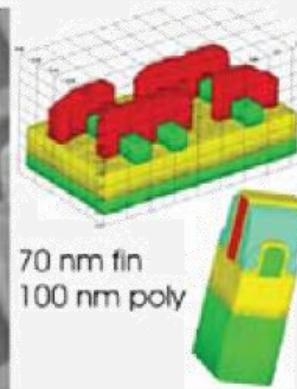
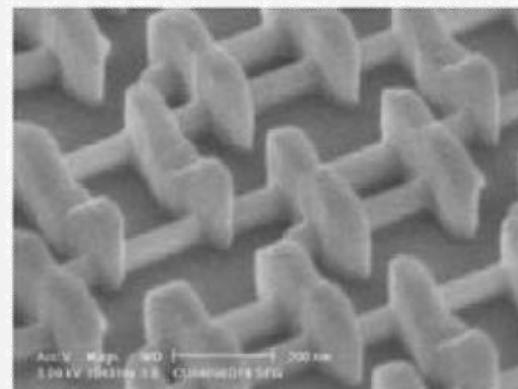
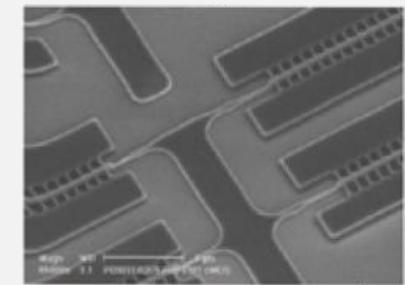
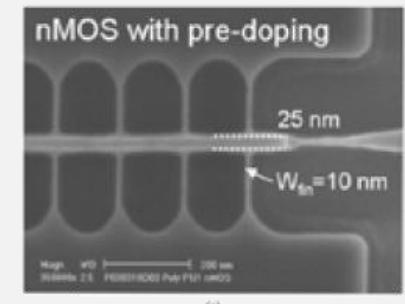
## FinFET Circuits – Ring Oscillator & SRAM

### Ring oscillator:

- H<sub>2</sub> anneal to smooth fin
- Fin pre-doping (phos., nFET) to set V<sub>T</sub>

### SRAM:

- Implant shadowing by multiple fins (fin height)
- Double contact hole print/etch for fin topography

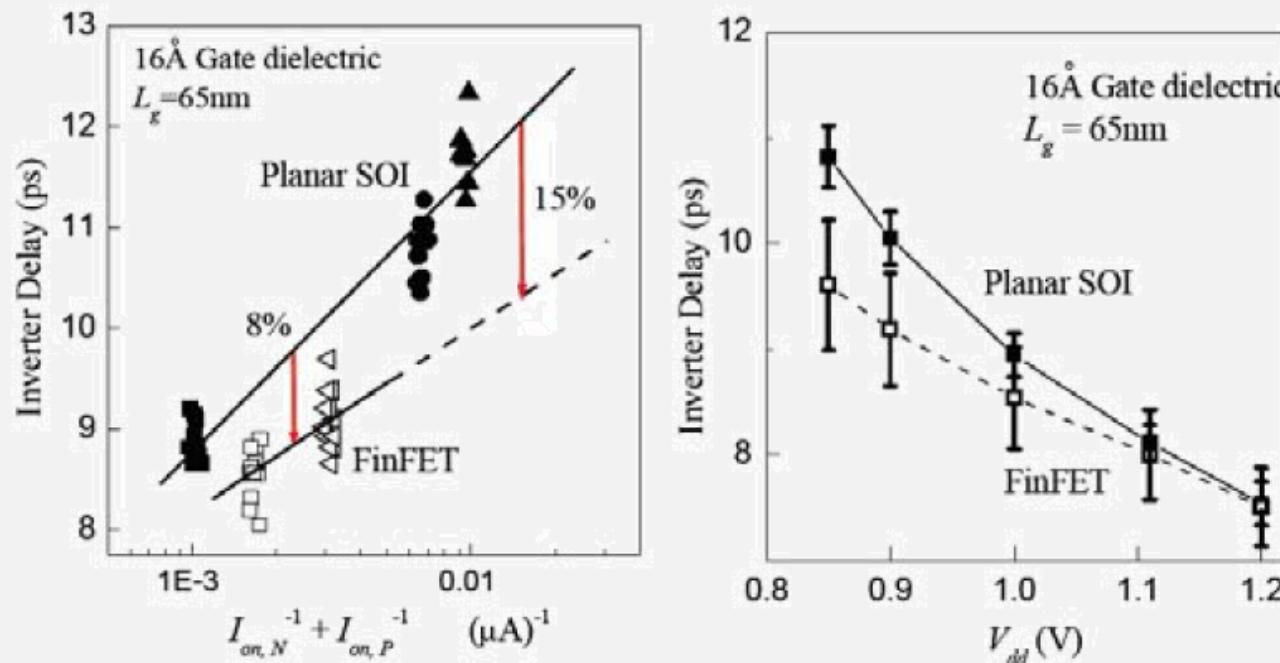


A. Nackaerts, ... and S. Biesemans, "A 0.314 $\mu$ m<sup>2</sup> 6T-SRAM cell build with tall triple-gate devices for 45nm node applications using 0.75NA 193nm lithography," *IEDM Tech. Dig.*, pp. 269 - 272, December 2004.

N. Collaert, ... and S. Biesemans, "A functional 41-stage ring oscillator using scaled FinFET devices with 25-nm gate lengths and 10-nm fin widths applicable for the 45-nm CMOS node," *IEEE Electron Device Lett.*, Vol. 25, pp. 568 - 570, August 2004.

## FinFET Ring Oscillators

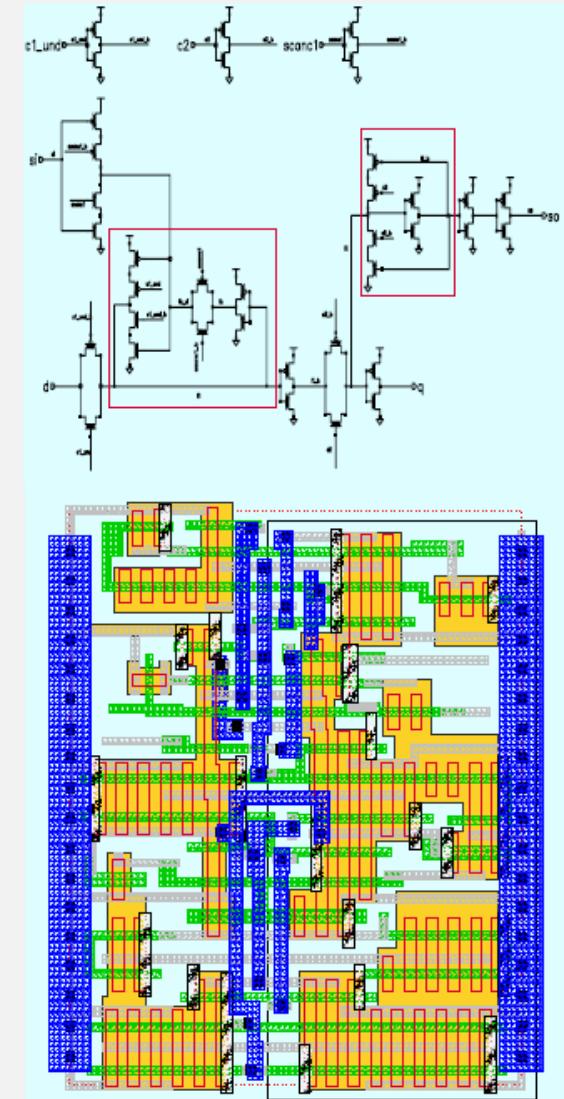
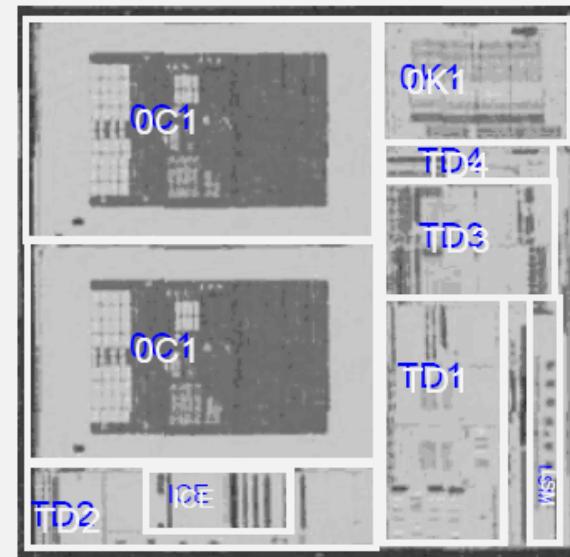
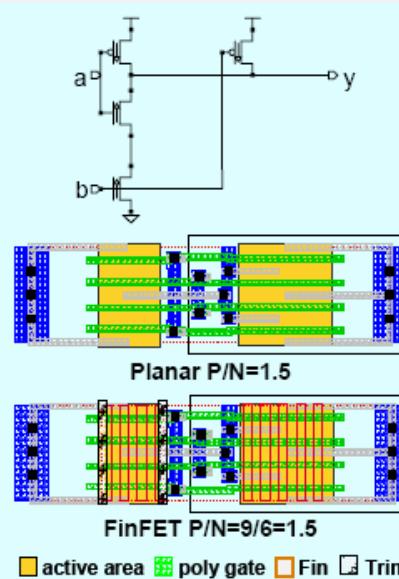
- **8 – 15% faster than planar SOI**
  - Reduced gate to body capacitance
- **FinFET integrated with FD/PD-SOI**



F. Yang, ..., and C. Hu, "A 65nm node strained SOI technology with slim spacer," *IEDM Tech. Dig.*, pp. 627 - 630, December 2003.

# FinFET Microprocessor

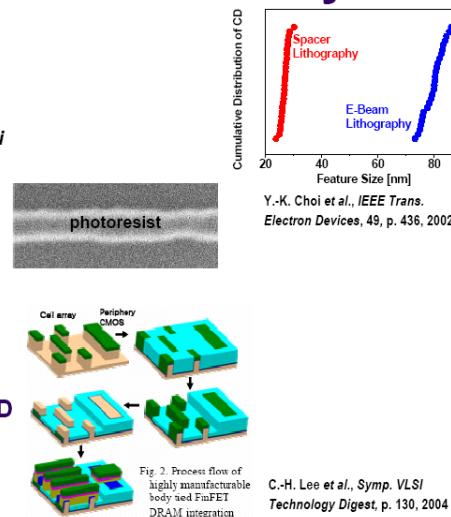
- Map planar SOI design to FinFET
  - Automatic layout conversion (~ 90%)
    - Device width tuning
    - EDS devices



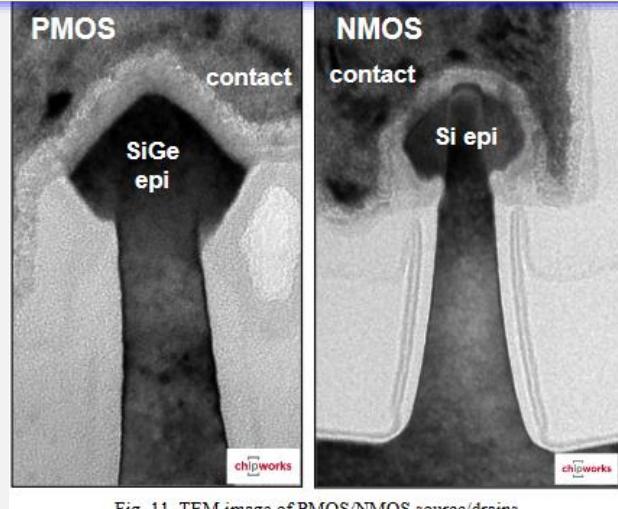
T. Ludwig, I. Aller, V. Gernhoefer, J. Keinert, E. Nowak, R.V. Joshi, A. Mueller, S. Tomaschko, "FinFET technology for future microprocessors," *IEEE SOI Conf.*, pp. 33 – 34 (2003).

## Design for Manufacturability

- Spacer process can be used to achieve sub-lithographic & uniform  $T_{Si}$
- Line-edge roughness is not an issue for FinFET gate lithography
- FinFETs can be made on bulk-Si wafers
  - lower cost
  - good recrystallization of S/D regions after implantation
  - improved thermal conduction

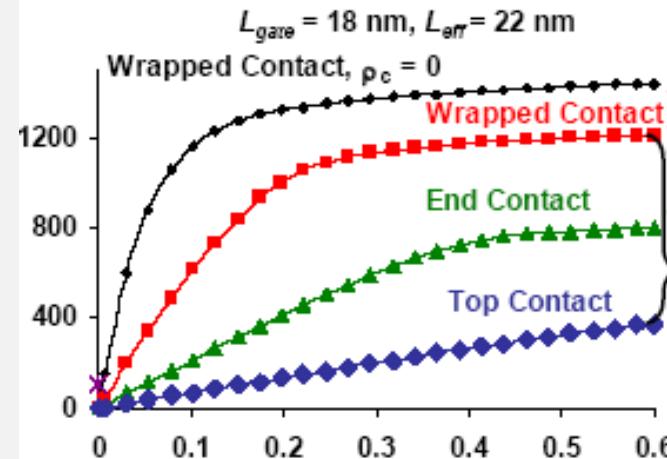


Intel 14nm : Wrapped contact for S/D



## Impact of S/D Contact Structure

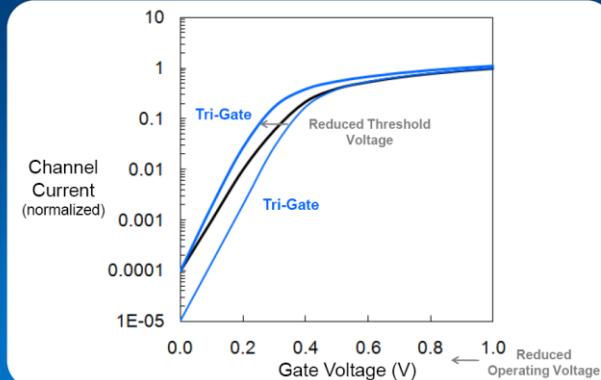
H. Kam and T.-J. King, 2004 Silicon Nanoelectronics Workshop



• Parasitic resistance (& capacitance) will limit  $I_{dsat}$

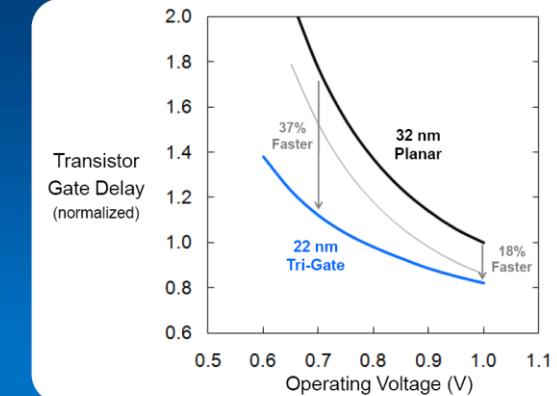
Source: J. Dick: Moore's Law Continues into 1x-nm Era, IEEE 2016

## Transistor Operation



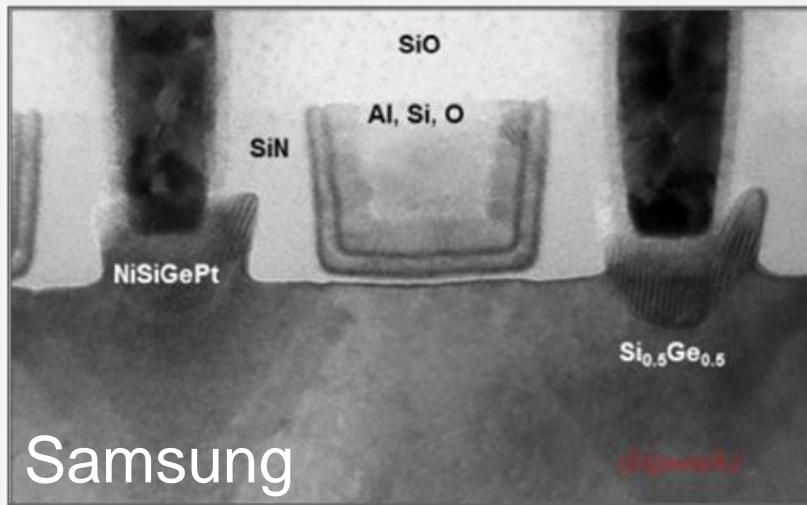
The steeper sub-threshold slope can also be used to target a lower threshold voltage, allowing transistors to operate at lower voltage to reduce power and/or improve switching speed

## Transistor Gate Delay



22 nm 3-D Tri-Gate transistors provide improved performance at high voltage and an *unprecedented* performance gain at low voltage

2012 Node : 20 nm

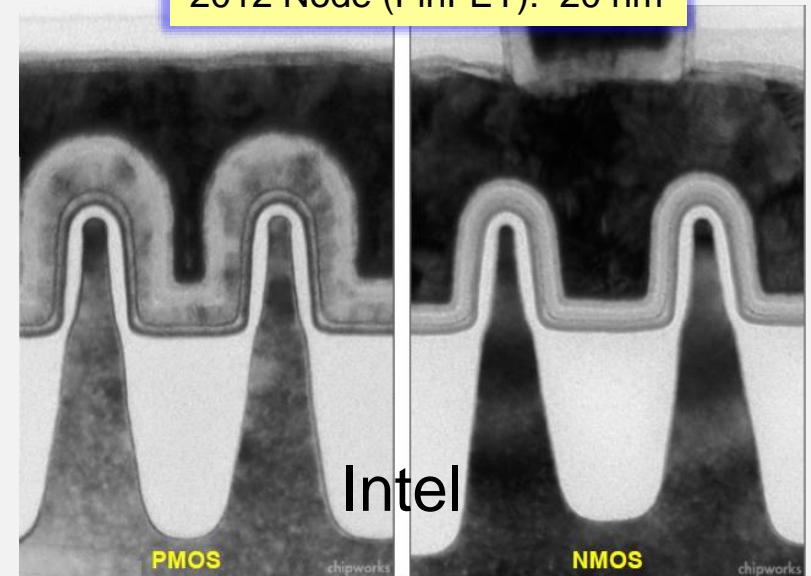


Samsung

Source: J. Dick: Moore's Law Continues into 1x-nm Era, IEEE 2016

2012 Node (FinFET): 20 nm

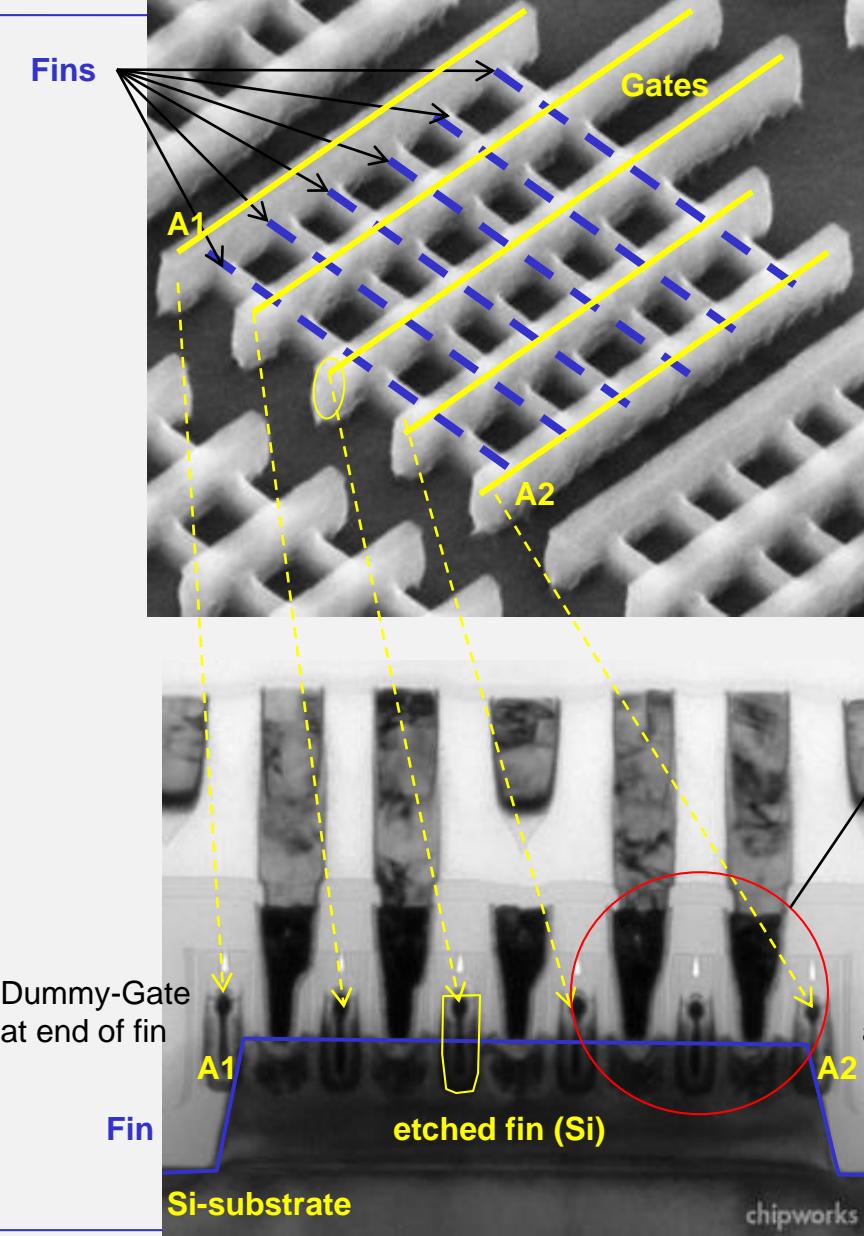
Source: Chipworks



Dr.-Ing. J.Biba  
AdMOS, 6-127

Fig.1

## Multi-Gate FETs



**Fig.2: cross section along the fin**

(this cross-section is a different structure (only 5 gates) compared to the top view above (Fig.1) with 6 gates)

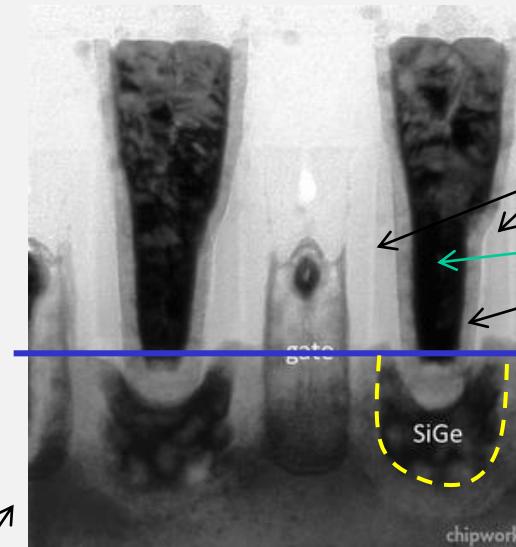
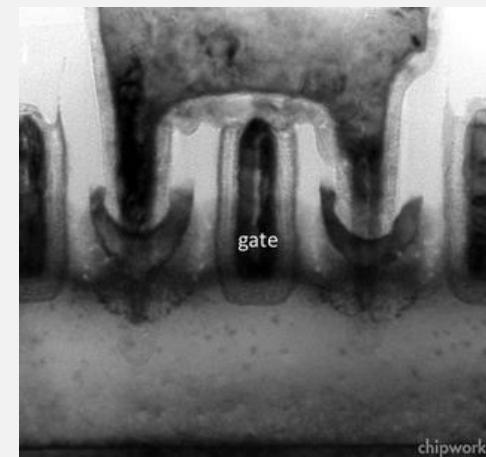
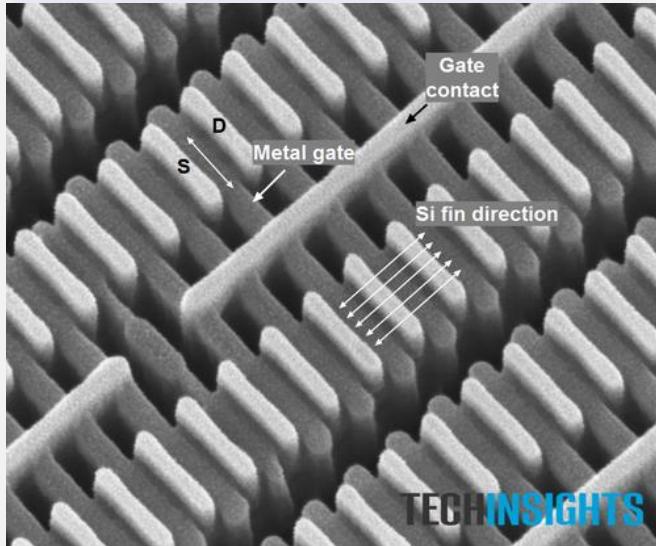


Fig.4

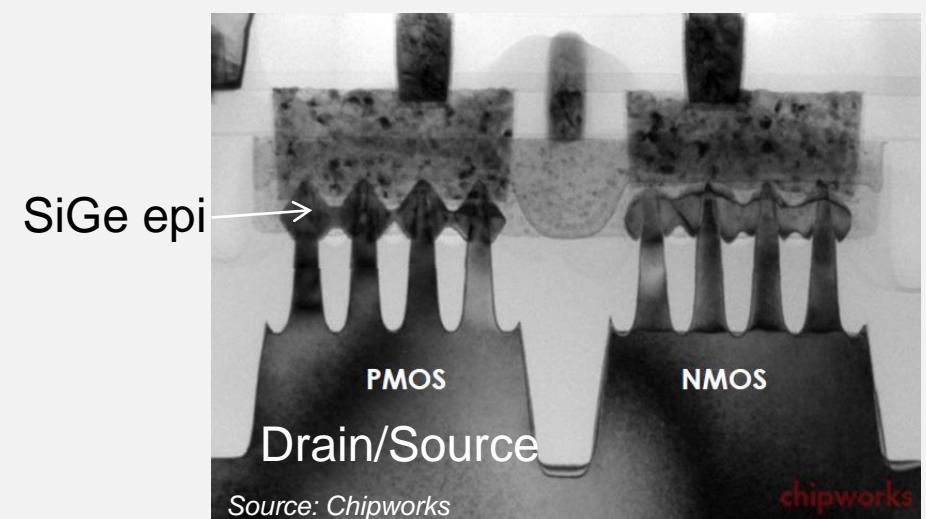
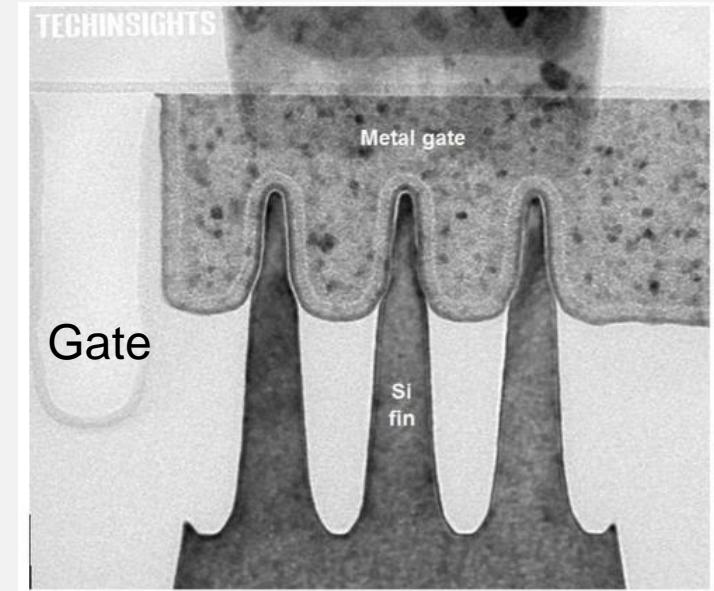
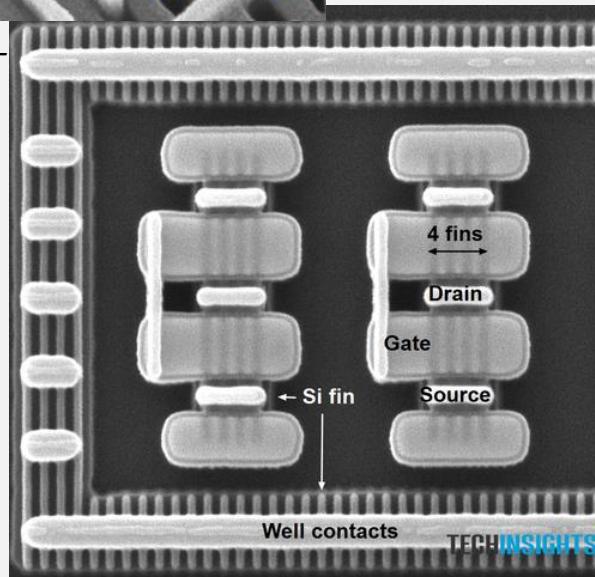
## cross-section of NMOS



2015 Node: 14 nm by Samsung



Samsung 14 nm FinFET  
Source: TechInsights



Source: Chipworks

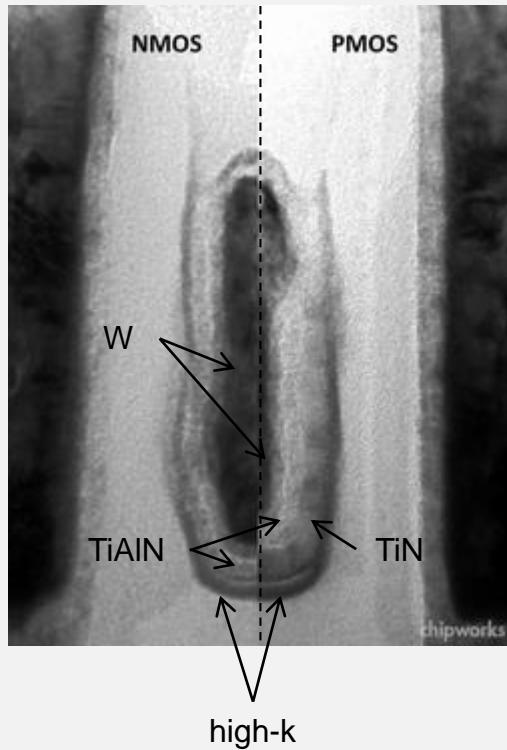
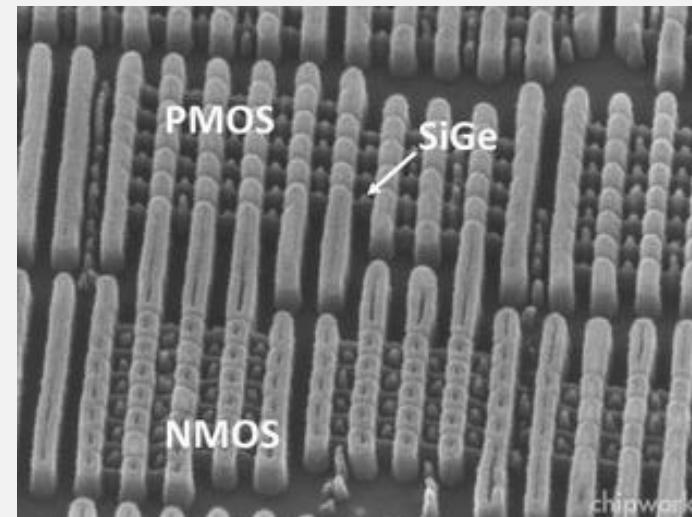


Figure 13 is a composite image of NMOS and PMOS gates so that the differences are highlighted. The dark line surrounding the gate structures is the Hf-based high-k, and within that are the two work-function materials, likely TiN for PMOS and TiAlN for NMOS. (The columnar structure of the PMOS TiN is visible in the right half of the image.)

The fill has been changed from TiAl in the earlier parts to tungsten. It is more prominent in the NMOS gates than the PMOS, because the PMOS structure includes both work-function metals, whereas the TiN has been etched out of the NMOS gates. At the 45-nm node Intel used tensile tungsten in the contacts to apply channel stress – have they transposed this to the gates in the 22-nm process?



With FinFET proper working devices with acceptable SCE and DIBL down to technology node 5 nm (year 2020) look reasonable

Technological problems will be:

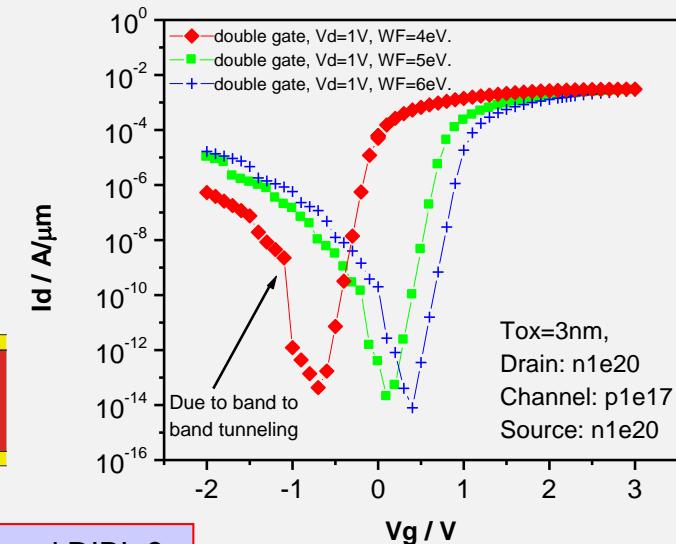
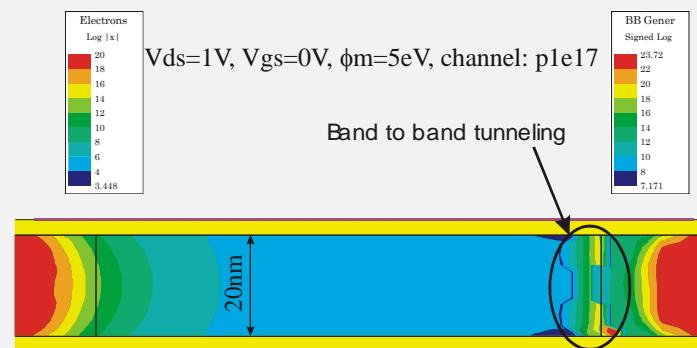
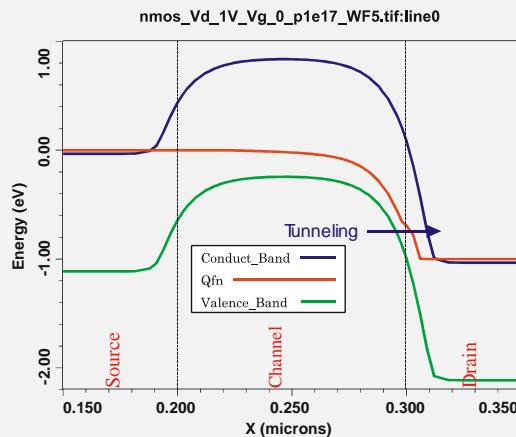
gate oxide < 1nm → high-k necessary

ultra-thin silicon < 5nm must be vertically patterned → oxide quality on etched, vertical sidewall ?

for symmetrical  $V_t$  new gate metals must be used

Electrical problems:

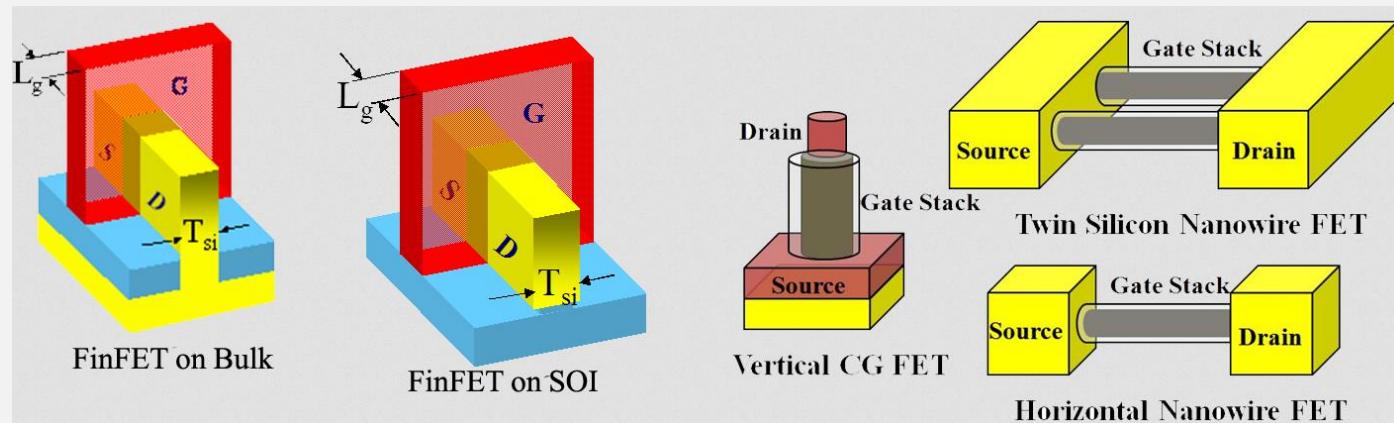
parasitic tunneling currents → high leakage current



? Is there a design window for symmetrical  $V_t$ ,  $I_{on}$ ,  $I_{off}$ , SCE and DIBL ?

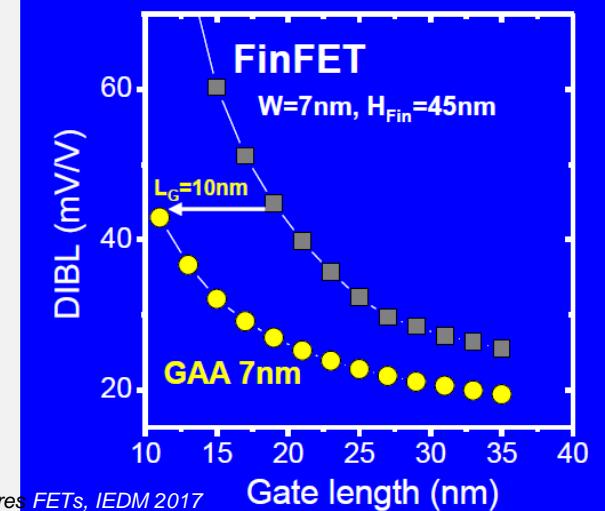
YEAR OF PRODUCTION	2018	2020	2022	2025	2028	2031	2034
	G54M36	G48M30	G45M24	G42M21	G40M16	G40M16T2	G40M16T4
<i>Logic industry "Node Range" Labeling (nm)</i>	"7"	"5"	"3"	"2.1"	"1.5"	"1.0 eq"	"0.7 eq"
<i>IDM-Foundry node labeling</i>	i10-f7	i7-f5	i5-f3	i3-f2.1	i2.1-f1.5	i1.5e-f1.0e	i1.0e-f0.7e
<i>Logic device structure options</i>	FinFET	finFET	finFET LGAA	LGAA	LGAA VGAA	LGAA-3D VGAA	LGAA-3D VGAA
<i>Mainstream device for logic</i>	finFET	finFET	finFET	LGAA	LGAA	LGAA-3D	LGAA-3D

The figure shows the International Roadmap for Devices and Systems (IRDS) 2018 Update. It displays a timeline from 2018 to 2034, showing the progression of FinFET technology. The 2018 update logo is on the left. The timeline shows cross-sectional diagrams of FinFET structures at various nodes, illustrating the increasing complexity and number of gates per fin over time.



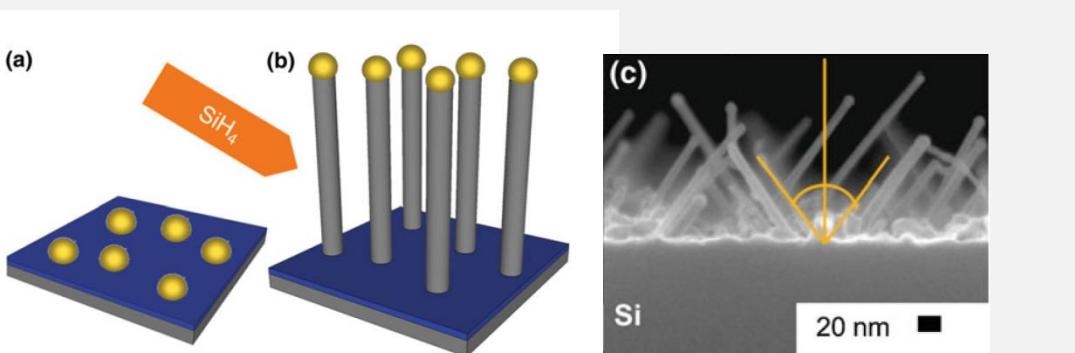
Different FinFET structures which can be modeled by BSIM-CMG since 2012

Source: S.Barraud: Performance and Design considerations for Gate-All-around Stacked-NanoWires FETs, IEDM 2017



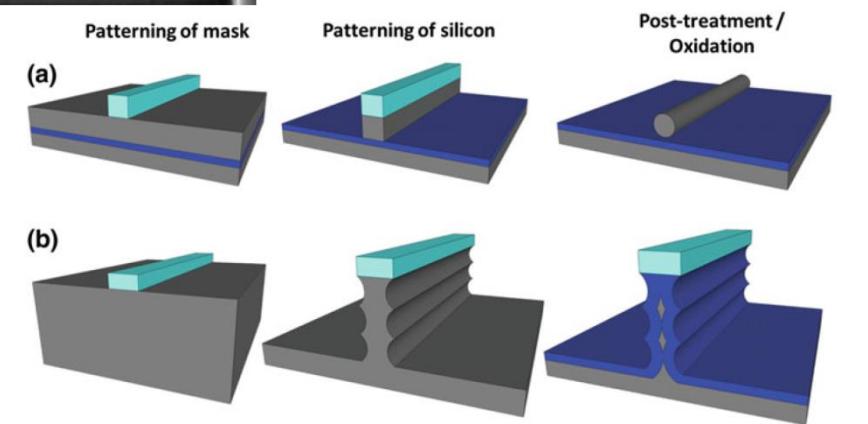
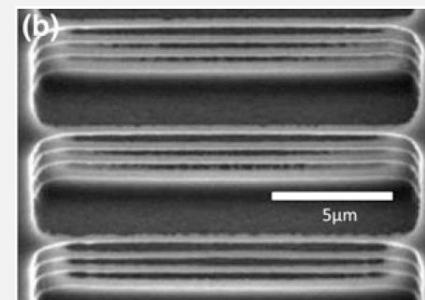
## Bottom-up fabrication

- Material is added to the substrate in a self-organized way.
- Most common is the liquid growth mechanism (VLSI)
  - Material to be grown is delivered in gas phase
  - Gold particles acts as catalyst and collects the silicon
  - Because it leads to lowering of the dissociation energies of the gas
- Advantages
  - Perfect geometry
  - Different types of substrate can be used
  - Complex structures without lithography
- Disadvantages
  - No CMOS process
  - Place of nanowires is not well defined
  - No mass production structuring
- Alternative
  - Gold free processes developed
  - Combine it with Top-down fabrication



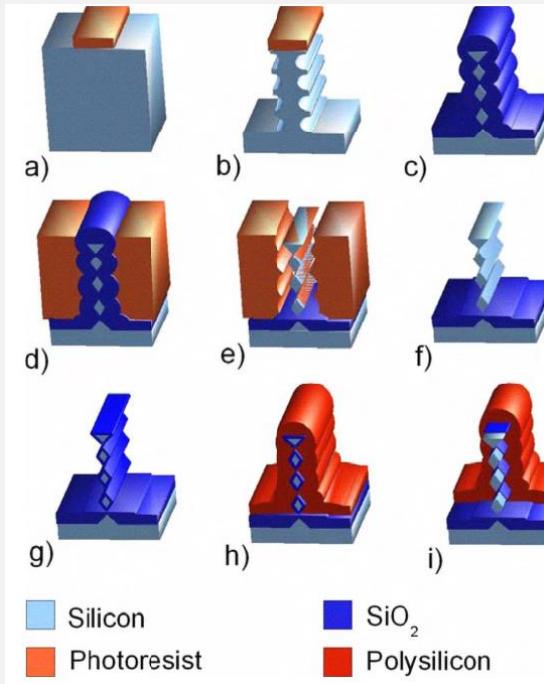
## Top-down fabrication

- Silicon technology can be used
- Advantages
  - Well established process
  - Mass production
  - Parallel nanowires
- Disadvantages
  - Non-perfect geometry
  - Advanced lithography
  - Lateral wires



Source: Anisotropic Nanomaterials, Preparation, Properties, and Applications, Li Q, 2015, Springer Verlag

## Top-down fabrication



- Lithography
- DRIE-etching
- Wet oxidation
- Lithography
- CMP and BHF etching
- Nanowires
- Oxidation to form the gate oxide
- Depositing polysilicon
- Patterning of the gatestack

Fabrication and Characterization of Vertically Stacked Gate-All-Around Si Nanowire FET Arrays

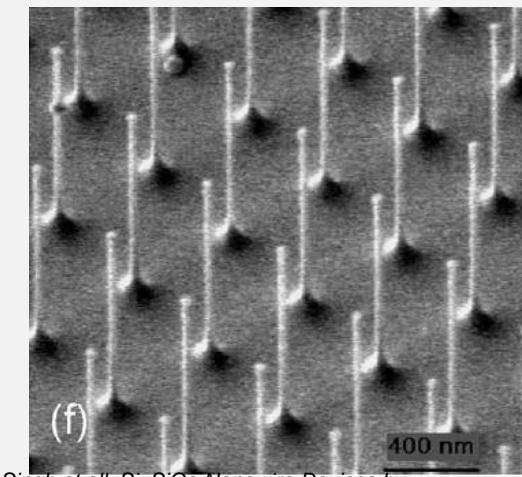
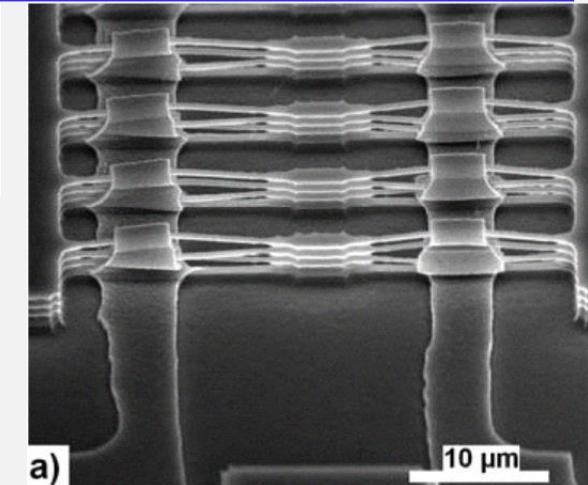
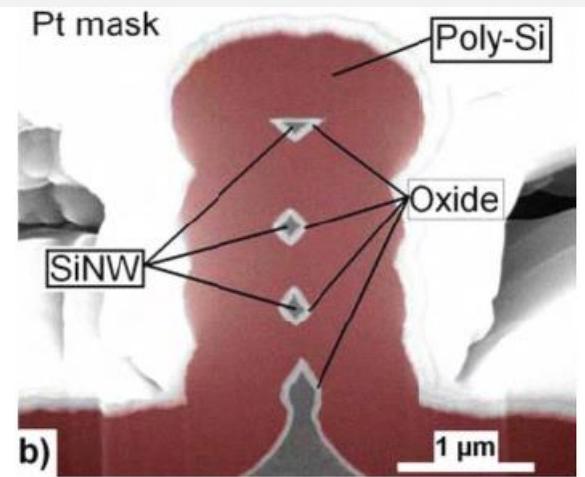
Davide Sacchetto<sup>1</sup>, M. Haykel Ben-Jamaa<sup>1</sup>, Giovanni De Micheli<sup>1</sup> and Yusuf Leblebici<sup>2</sup>

<sup>1</sup>Integrated System Laboratory (LSI), <sup>2</sup>Microelectronic System Laboratory (LMS)

Ecole Polytechnique Fédérale de Lausanne, Lausanne, Switzerland

e-mail: davide.sacchetto@epfl.ch

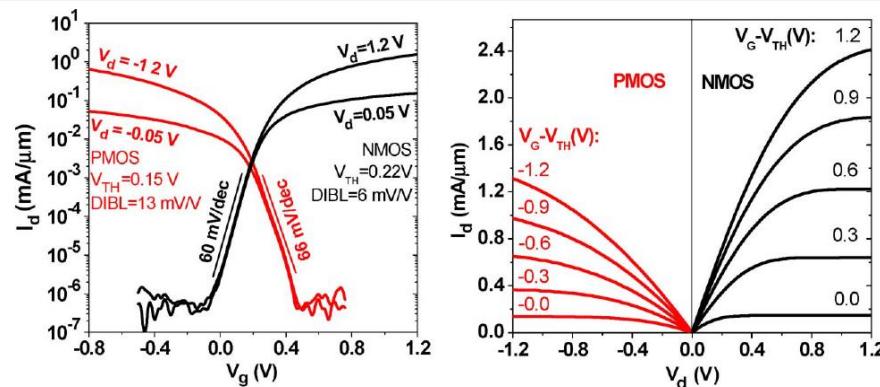
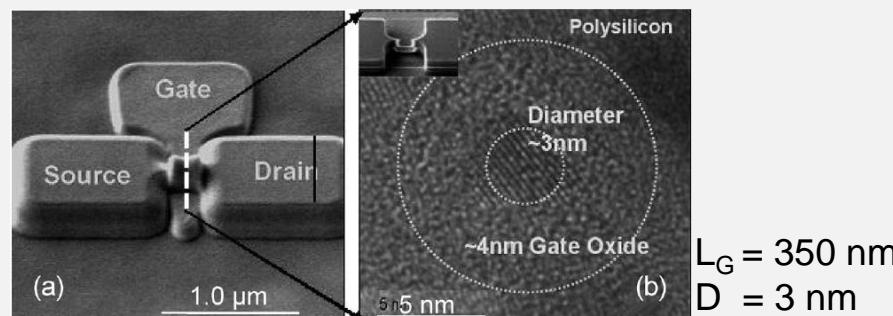
IEEE 2009



Singh et al., Si, SiGe Nanowire Devices by Top-Down Technology and their Applications, IEEE 2008

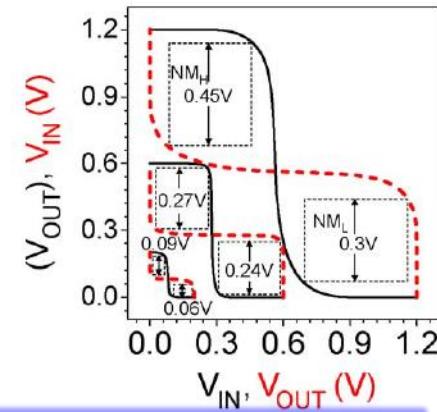
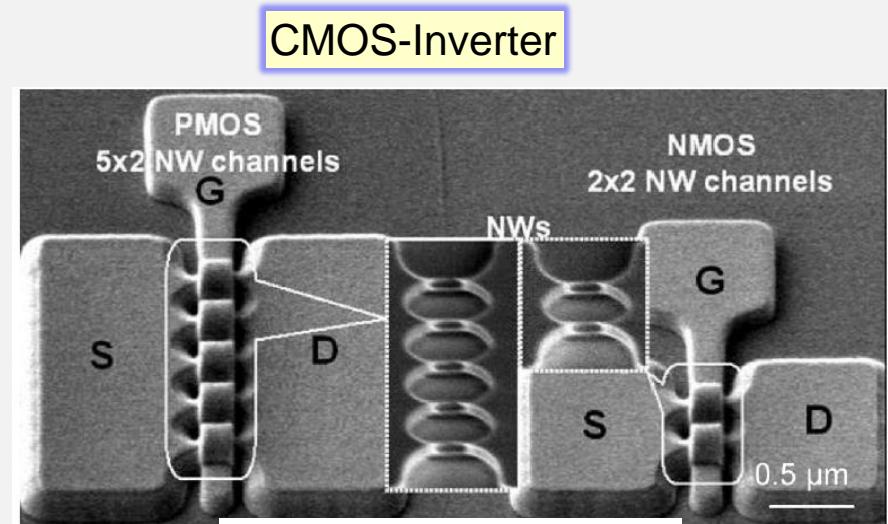
# Si, SiGe Nanowire Devices by Top–Down Technology and Their Applications

Navab Singh, Kavitha D. Buddharaju, S. K. Manhas, A. Agarwal, Subhash C. Rustagi,  
G. Q. Lo, N. Balasubramanian, and Dim-Lee Kwong  
IEEE 2008



► Transistors show good IV-characteristic

► Reducing short-channel effects



Fabrication of CMOS-circuits

► High fluctuation of  $I_{ON}$  and  $V_T$

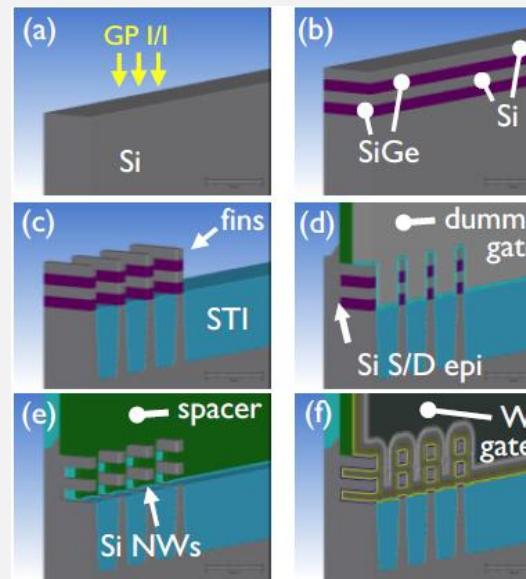
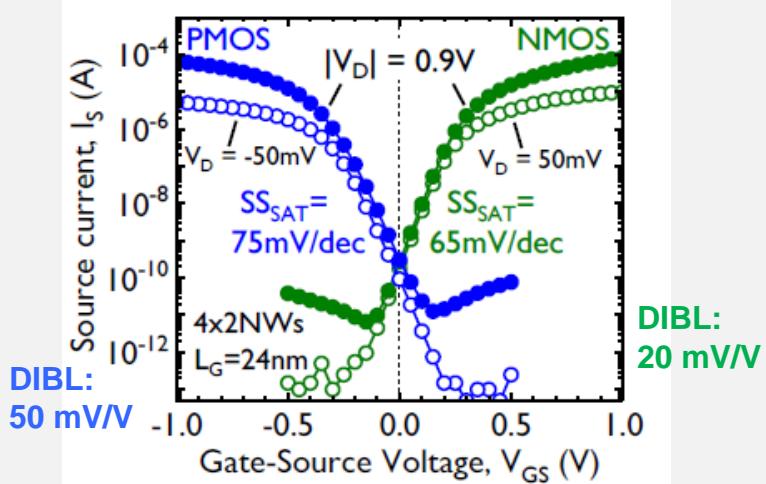
► Tuning of  $V_T$

## Vertically Stacked Gate-All-Around Si Nanowire Transistors: Key Process Optimizations and Ring Oscillator Demonstration

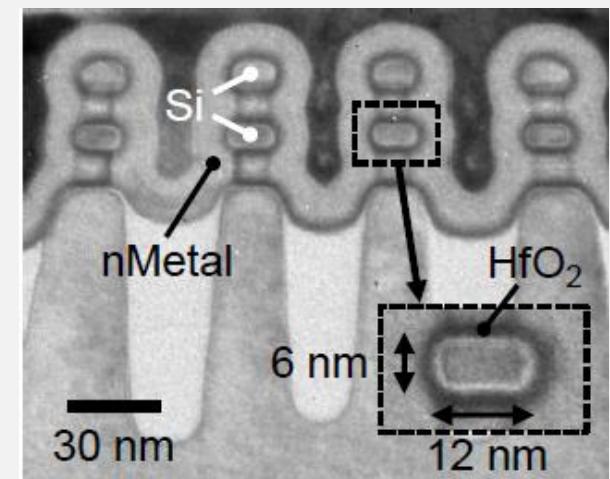
H. Mertens\*, R. Ritzenthaler, V. Pena<sup>1</sup>, G. Santoro<sup>1</sup>, K. Kenis, A. Schulze, E. D. Litta, S. A. Chew, K. Devriendt, T. Chiarella, S. Demuynck, D. Yakimets, D. Jang, A. Spessot, G. Eneman, A. Dangol, P. Lagrain, H. Bender, S. Sun<sup>2</sup>, M. Korolik<sup>2</sup>, D. Kioussis<sup>2</sup>, M. Kim<sup>2</sup>, K.-H. Bu<sup>2</sup>, S. C. Chen<sup>2</sup>, M. Cogorno<sup>2</sup>, J. Devrajan<sup>2</sup>, J. Machillot<sup>1</sup>, N. Yoshida<sup>2</sup>, N. Kim<sup>2</sup>, K. Barla, D. Mocuta, N. Horiguchi

IEDM 2017

<sup>1</sup>Imec, Kapeldreef 75, B-3001 Leuven, Belgium, \*hans.mertens@imec.be  
<sup>2</sup>Applied Materials, Leuven, Belgium, <sup>2</sup>Applied Materials 3050 Bowers Avenue, Santa Clara, CA 95053, USA



- Starting material: Bulk Si wafer
- Zero markers
- CMOS ground plane I/I + anneal
- Si<sub>0.7</sub>Ge<sub>0.3</sub>/Si/Si<sub>0.7</sub>Ge<sub>0.3</sub>/Si epitaxy
- Spacer-defined fin patterning (SADP)
- Low-temperature STI fill (New)
- Dummy gate patterning
- Spacer formation
- In-situ doped emb. S/D epi (New)
- FCVD-based ILD0 fill (New)
- Dummy poly/oxide removal
- Si NW formation by SiGe etch (New)
- Gate dielectric (IL-SiO<sub>2</sub>/HfO<sub>2</sub>) dep.
- Dual work function metal int. (New)
- Fill metal (W) deposition/CMP
- Contact formation + M1 BEOL



► IMEC showed last year working CMOS-GAA transistors on one Wafer

► Dual-workfunction metallization

► Working ring oscillator with 41 inverter stages

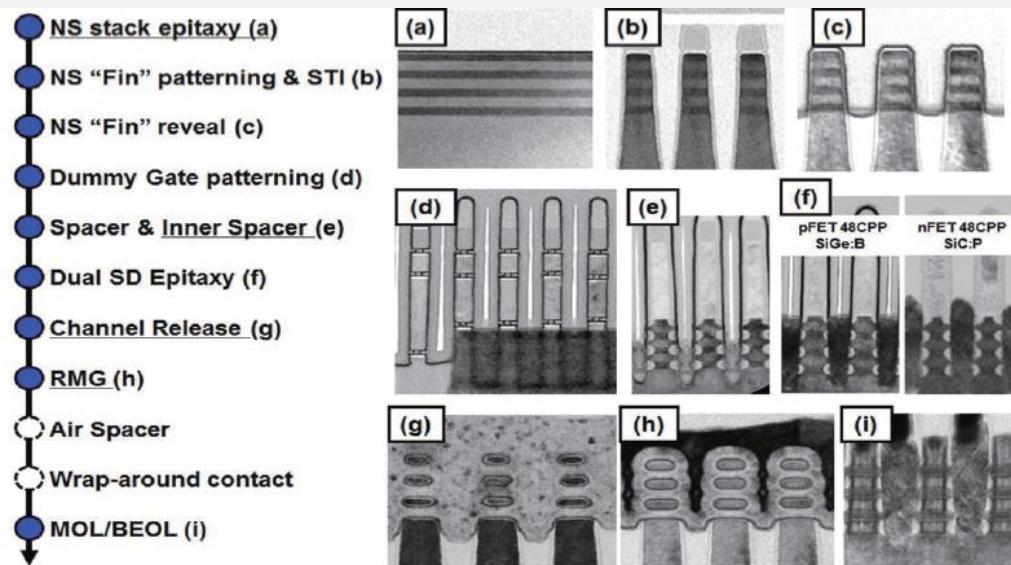
## Stacked Nanosheet Gate-All-Around Transistor to Enable Scaling Beyond FinFET

N. Loubet<sup>1</sup>, T. Hook<sup>1</sup>, P. Montanini<sup>1</sup>, C.-W. Yeung<sup>1</sup>, S. Kanakasabapathy<sup>1</sup>, M. Guillorn<sup>1</sup>, T. Yamashita<sup>1</sup>, J. Zhang<sup>1</sup>, X. Miao<sup>1</sup>, J. Wang<sup>1</sup>, A. Young<sup>1</sup>, R. Chao<sup>1</sup>, M. Kang<sup>2</sup>, Z. Liu<sup>1</sup>, S. Fan<sup>1</sup>, B. Hamieh<sup>1</sup>, S. Sieg<sup>1</sup>, Y. Mignot<sup>1</sup>, W. Xu<sup>1</sup>, S.-C. Seo<sup>1</sup>, J. Yoo<sup>2</sup>, S. Mochizuki<sup>1</sup>, M. Sankarapandian<sup>1</sup>, O. Kwon<sup>2</sup>, A. Carr<sup>1</sup>, A. Greene<sup>1</sup>, Y. Park<sup>2</sup>, J. Frougier<sup>3</sup>, R. Galatage<sup>3</sup>, R. Bao<sup>1</sup>, J. Shearer<sup>1</sup>, R. Conti<sup>1</sup>, H. Song<sup>2</sup>, D. Lee<sup>2</sup>, D. Kong<sup>1</sup>, Y. Xu<sup>1</sup>, A. Arceo<sup>1</sup>, Z. Bi<sup>1</sup>, P. Xu<sup>1</sup>, R. Muthinti<sup>1</sup>, J. Li<sup>1</sup>, R. Wong<sup>1</sup>, D. Brown<sup>3</sup>, P. Oldiges<sup>1</sup>, R. Robison<sup>1</sup>, J. Arnold<sup>1</sup>, N. Felix<sup>1</sup>, S. Skordas<sup>1</sup>, J. Gaudiello<sup>1</sup>, T. Standaert<sup>1</sup>, H. Jagannathan<sup>1</sup>, D. Corliss<sup>1</sup>, M.-H. Na<sup>1</sup>, A. Knorr<sup>3</sup>, T. Wu<sup>1</sup>, D. Gupta<sup>1</sup>, S. Lian<sup>2</sup>, R. Divakaruni<sup>1</sup>, T. Gow<sup>1</sup>, C. Labelle<sup>3</sup>, S. Lee<sup>2</sup>, V. Paruchuri<sup>1</sup>, H. Bu<sup>1</sup>, and M. Khare<sup>1</sup>

IBM Research, 257 Fuller Road, Albany, NY 12203

<sup>1</sup>IBM, <sup>2</sup>Samsung Electronics, <sup>3</sup>GLOBALFOUNDRIES, email: njloubet@us.ibm.com

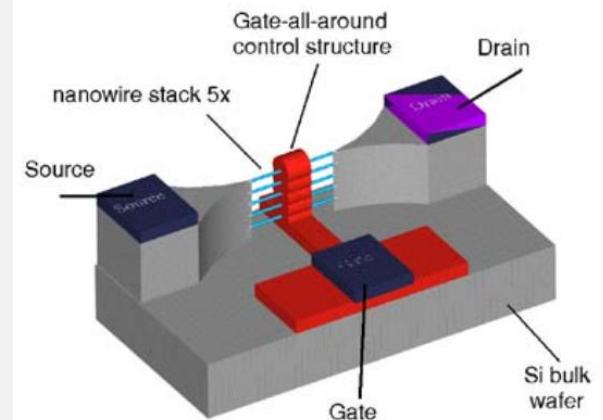
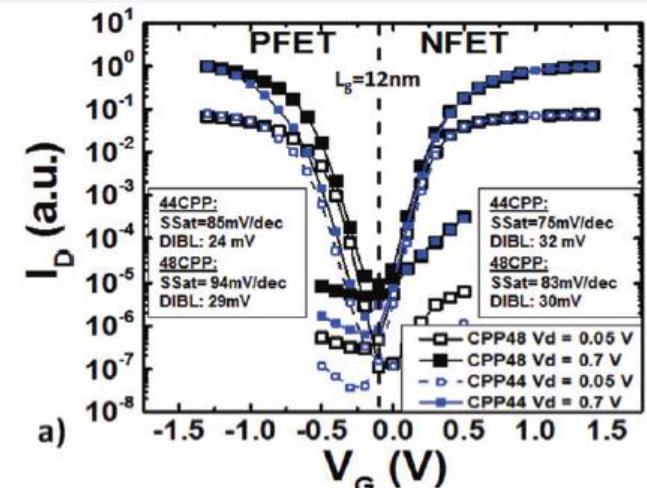
2017 Symposium on VLSI Technology Digest of Technical Papers



First presentation of GAA-FETs from these companies

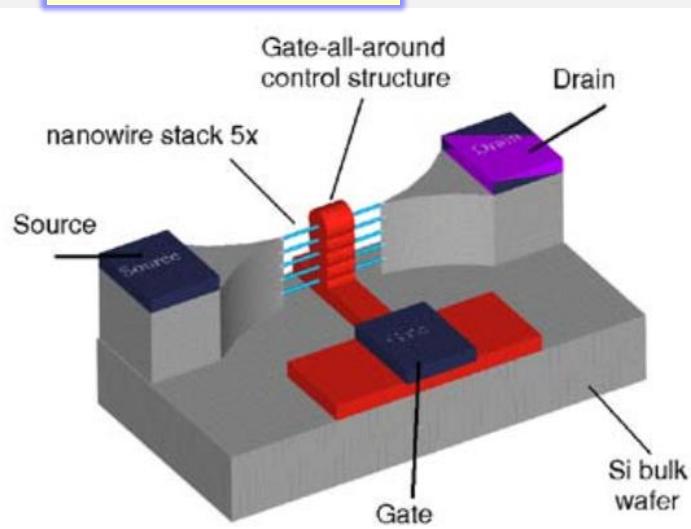
Dual-workfunction metallization

Maybe paths the way for sub 7 nm



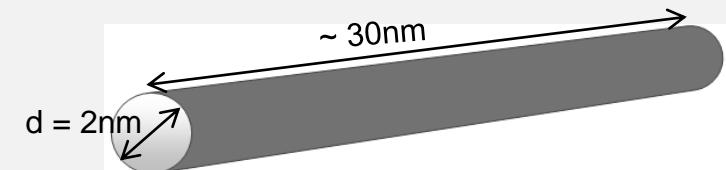
<http://semiengineering.com/uncertainty-grows-for-5nm-3nm/>

Picture correct ?



<http://semiengineering.com/uncertainty-grows-for-5nm-3nm/>

According to this picture and the roadmap (6.11):  
Gate length ~ 10nm  
Length of wire ~ 30nm  
Diameter of wire ~ 1-2nm



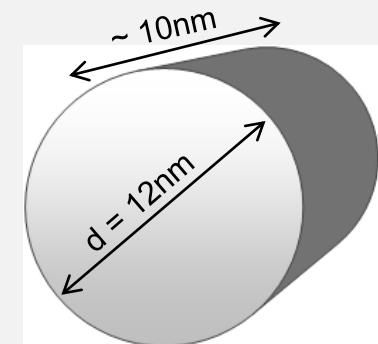
1-dim wire

With  $R = 25\text{k}\Omega$  (normally channel  $R_{\text{channel}} = 1\text{k}\Omega$ ) and  $C \sim 1.24\text{aF/nm}$  (see 6.11)  
 $\Rightarrow RC = 0.025\text{psec}$  see chapter 8 (single electron device)  
 $RC(\text{today}) \sim 1\text{psec}$



Reality

- Values of companies show we have no 1-dim wire
- Diameter of wire is almost the same size as length
- Reality looks different than on the top left picture

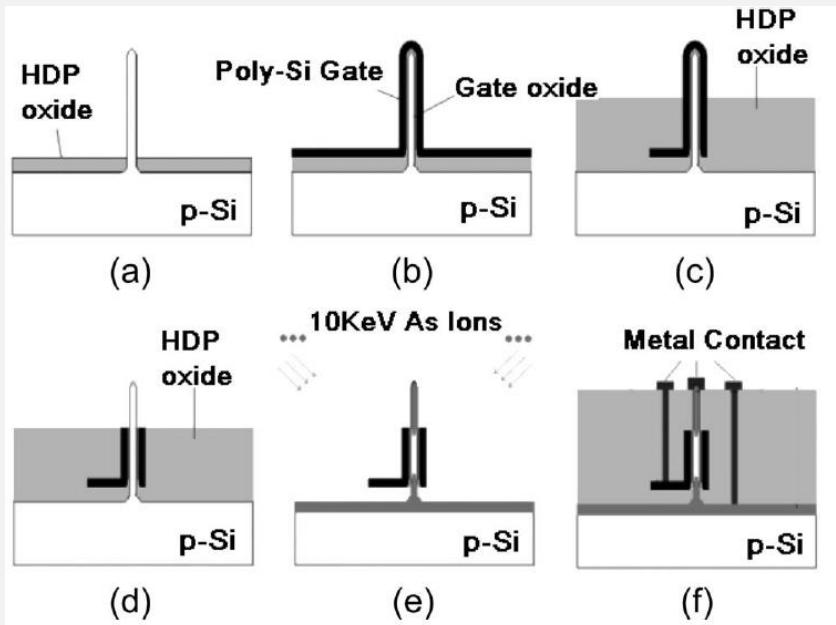


Reality we have no 1-dim nanowire

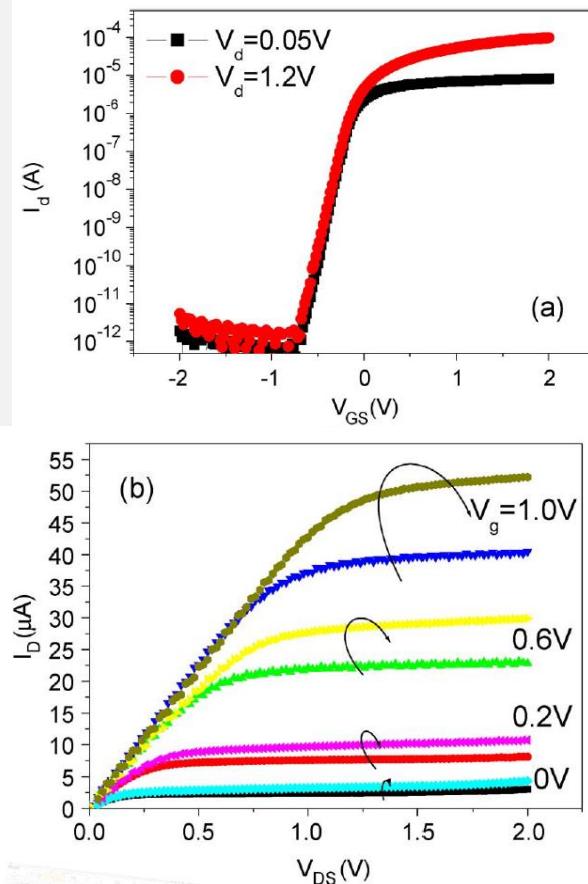
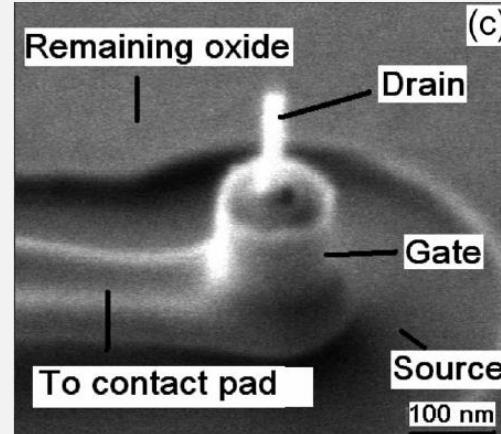
# Vertical Silicon-Nanowire Formation and Gate-All-Around MOSFET

IEEE 2008

B. Yang, K. D. Buddharaju, S. H. G. Teo, N. Singh, G. Q. Lo, and D. L. Kwong



$L_G = 150 \text{ nm}$   
 $D = 25 \text{ nm}$



Nanowires with diameters from 20 to 450 nm

Subthreshold Slope 75 mV/dec, DIBL 10~50 mV/V

CMOS fabrication technology was used

First Demonstration of Vertical Ge<sub>0.92</sub>Sn<sub>0.08</sub>/Ge and Ge GAA Nanowire pMOSFETs with Low SS of 66 mV/dec and Small DIBL of 35 mV/V

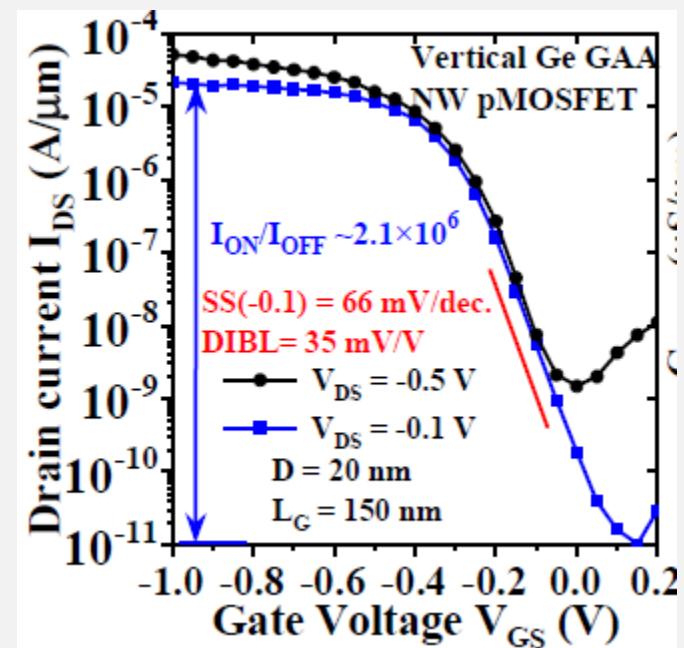
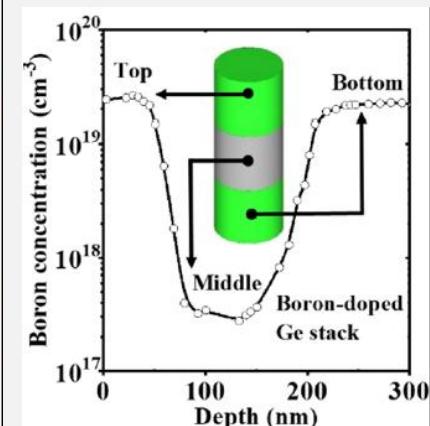
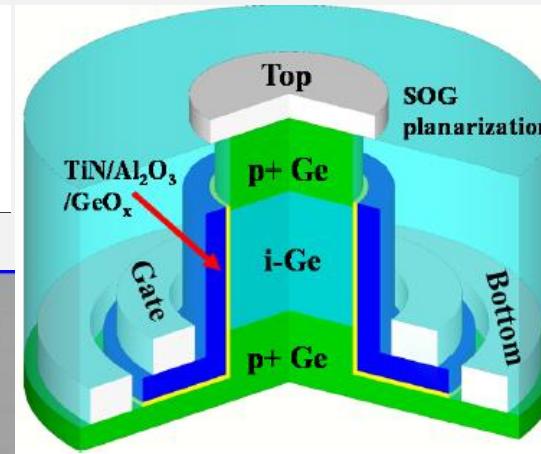
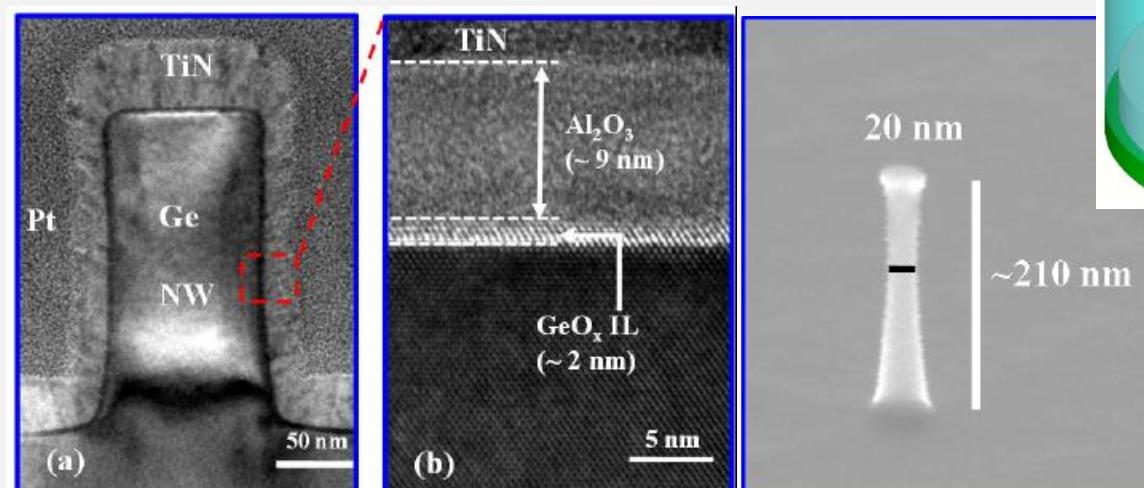
Mingshan Liu<sup>1</sup>, Stefan Scholz<sup>2</sup>, Konstantin Mertens<sup>1</sup>, Jin Hee Bae<sup>1</sup>, Jean-Michel Hartmann<sup>3</sup>, Joachim Knoch<sup>2</sup>, Dan Buca<sup>1</sup>, and Qing-Tai Zhao<sup>1</sup>

<sup>1</sup>Peter-Grünberg-Institute (PGI 9), Forschungszentrum Jülich, 52428, Germany, \*Email: m.liu@fz-juelich.de

IEEE 2019

<sup>2</sup>School of Electrical Engineering, RWTH Aachen University, 52074, Aachen, Germany.

<sup>3</sup>CEA, LETI, MINATEC Campus and University of Grenoble Alpes, 38054 Grenoble, France.



► Vertical Ge GAA p-channel transistor

► Low drive on-current

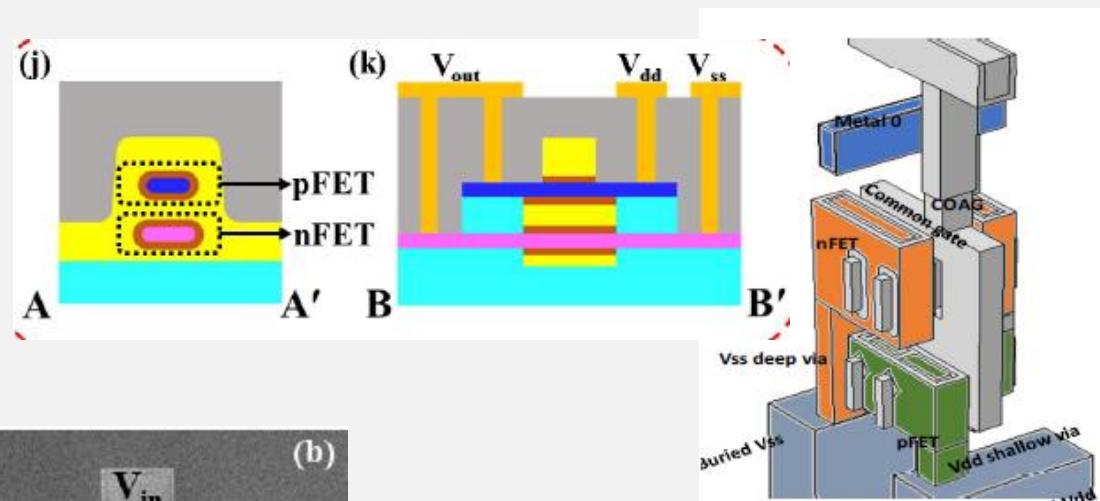
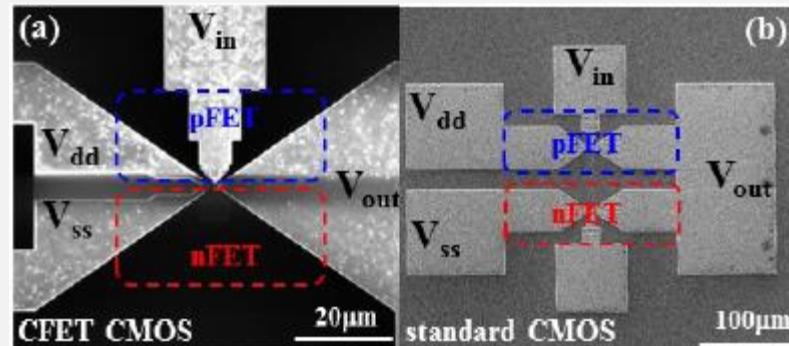
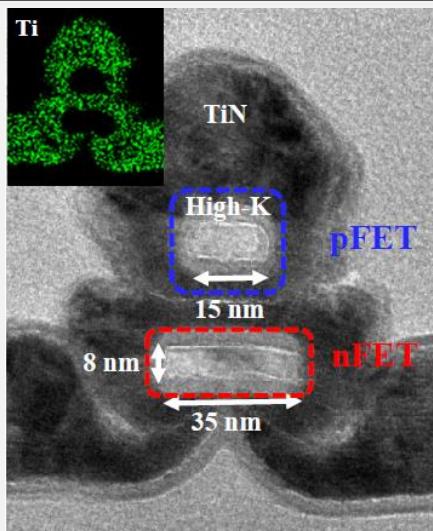
► Processes can be used for Si

► Still problems with Ge

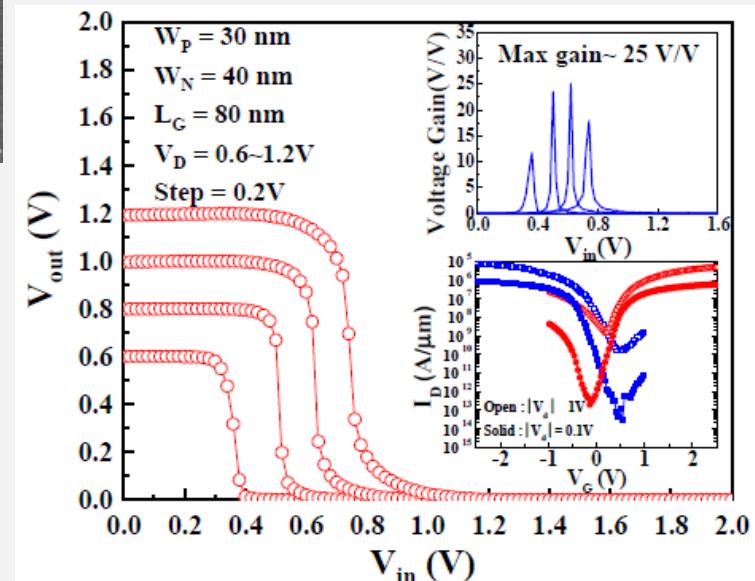
## First Demonstration of CMOS Inverter and 6T-SRAM Based on GAA CFETs Structure for 3D-IC Applications IEEE 2019

S.-W. Chang<sup>1,2</sup>, P.-J. Sung<sup>1,3,\*\*</sup>, T.-Y. Chu<sup>2</sup>, D. D. Lu<sup>2</sup>, C. -J. Wang<sup>1</sup>, N.-C. Lin<sup>1</sup>, C.-J. Su<sup>1</sup>, S.-H. Lo<sup>2</sup>, H.-F. Huang<sup>2</sup>, J.-H. Li<sup>2</sup>, M.-K. Huang<sup>2</sup>, Y.-C. Huang<sup>3</sup>, S.-T. Huang<sup>4</sup>, H.-C. Wang<sup>4</sup>, Y.-J. Huang<sup>4</sup>, J.-Y. Wang<sup>4</sup>, L.-W. Yu<sup>4</sup>, Y.-F. Huang<sup>5</sup>, F.-K. Hsueh<sup>1</sup>, C.-T. Wu<sup>1</sup>, W. C.-Y. Ma<sup>4</sup>, K.-H. Kao<sup>4</sup>, Y.-J. Lee<sup>1,\*</sup>, C.-L. Lin<sup>3</sup>, R.W. Chuang<sup>2</sup>, K.-P. Huang<sup>6</sup>, S. Samukawa<sup>7</sup>, Y. Li<sup>8</sup>, W.-H. Lee<sup>2</sup>, T.-S. Chao<sup>3</sup>, G.-W. Huang<sup>1</sup>, W.-F. Wu<sup>1</sup>, J.-Y. Li<sup>9</sup>, J.-M. Shieh<sup>1</sup>, W. -K. Yeh<sup>1</sup>, Y.-H. Wang<sup>2,10</sup>

<sup>1</sup>Taiwan Semiconductor Research Institute, Hsinchu, Taiwan; <sup>2</sup>Dept. of Electrical Engineering, National Cheng Kung University, Tainan, Taiwan; <sup>3</sup>Dept. of Electrophysics, National Chiao Tung University, Hsinchu, Taiwan; <sup>4</sup>Dept. of Electrical Engineering, National Sun Yat-sen University, Kaohsiung, Taiwan; <sup>5</sup>Dept. of Electronic Engineering, Feng Chia University, Taichung, Taiwan; <sup>6</sup>Mechanical and Systems Research Laboratories, Industrial Technology Research Institute, Hsinchu, Taiwan; <sup>7</sup>Institute of Fluid Science, Tohoku University, Sendai, Japan; <sup>8</sup>Dept. of Electrical and Computer Engineering, National Chiao Tung University, Hsinchu, Taiwan; <sup>9</sup>Dept. of Electrical Engineering and Graduate Institute of Electronics Engineering, National Taiwan University, Taipei, Taiwan; <sup>10</sup>National Applied Research Laboratories, Taipei, Taiwan; Tel: +886-3-5726100-7793, Fax: +886-3-5722715, \*Email: yilee@narlabs.org.tw, \*\*Email: pj sung@narlabs.org.tw



Source: J. Ryckaert: Enabling "Sub-5nm CMOS Technology Scaling Thinner and Taller!", IEEE 2019

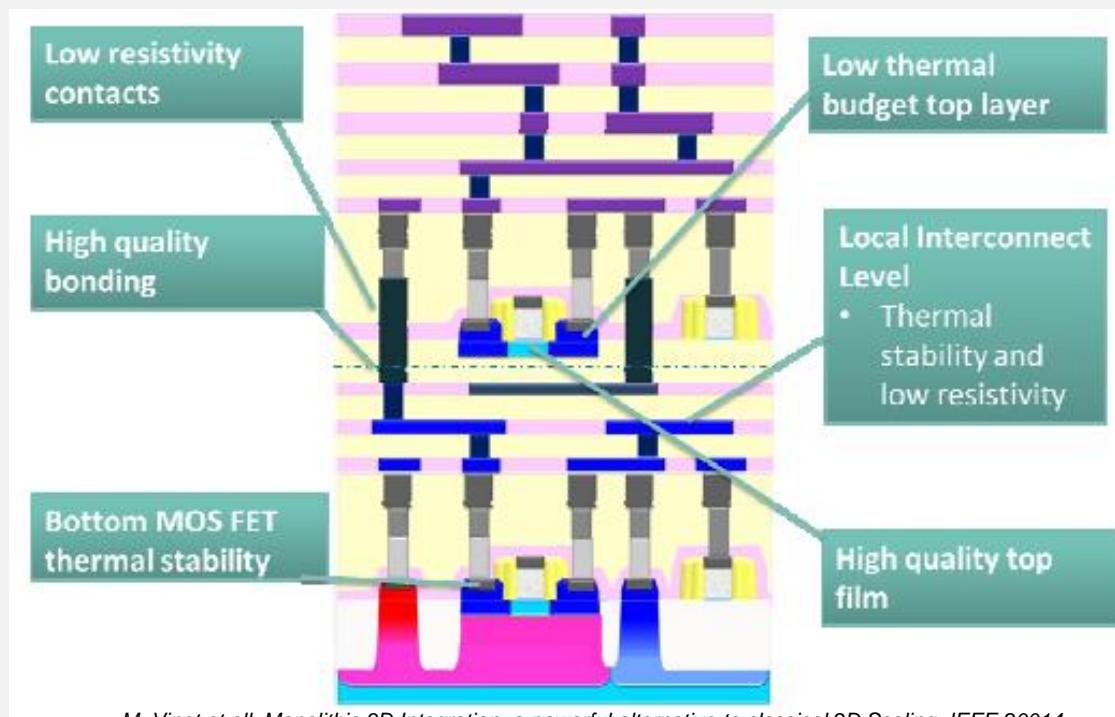


► First presentation of GAA-CFET Inverter

► Factor 2 smaller footprint

► Paths the way for sub 1nm node, but still work to be done

STMicroelectronics/IBM/Leti/Qualcomm (2008- now)

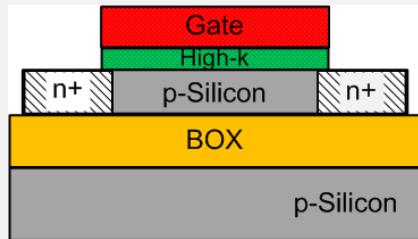


- Sequentially processing transistors on top of each other
- Ultra-high 3D contact density
- Challenge is that the top layers are processed with a low thermal budget to preserve any damage to the bottom transistor
- Process temperature for the top transistor may not be higher than 600°C

## Fabrication

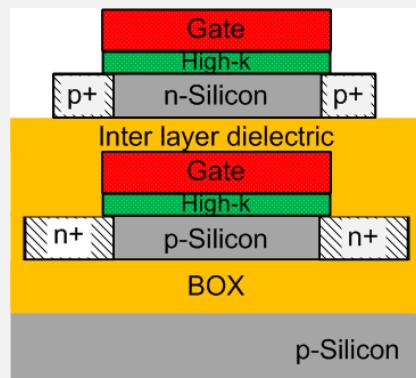
P. Batude et all, Advances in 3D CMOS Sequential Integration, IEEE 2009

### 1. Bottom MOSFET



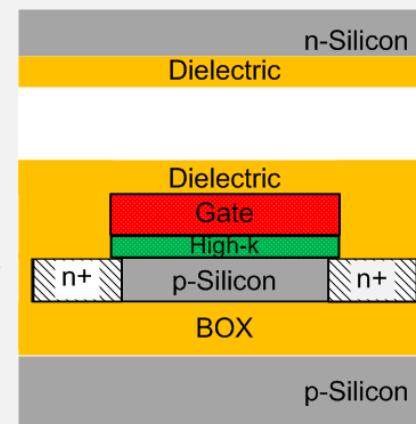
- FDSOI-MOSFET
- High-k/Metal-Gate

### 4. Top MOSFET



- FDSOI-MOSFET
- Temperature Budget max. 650°C
- High-k/Metal-Gate

### Top Wafer

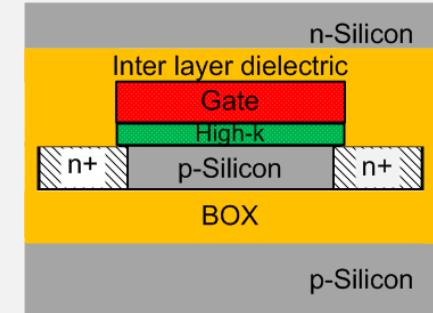


- Dielectric deposited
- Planarization with CMP

### 2. Inter-layer-dielectric

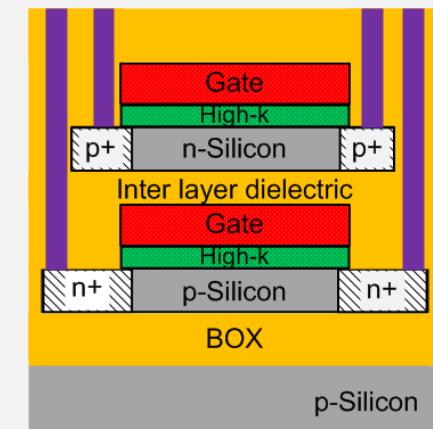
### Substrate Wafer

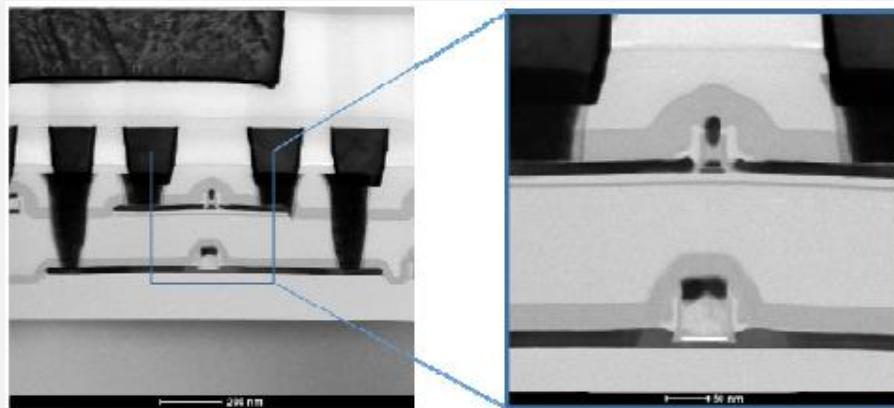
### 3. Bonding



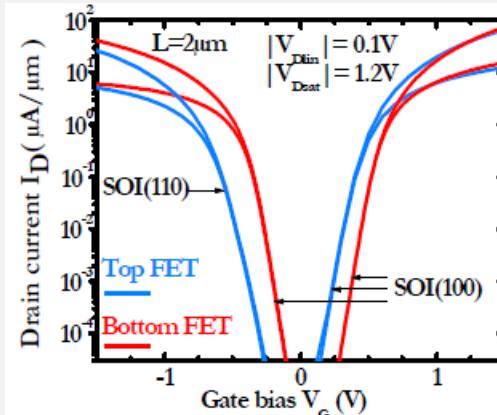
- Low temperature bonding
- Initial substrate removal

### 5. Contact process



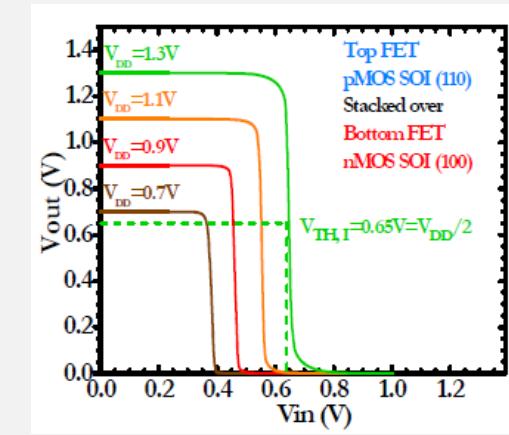


M. Vinet et all, Opportunities brought by sequential 3D CoolCubeTM integration, IEEE 20016



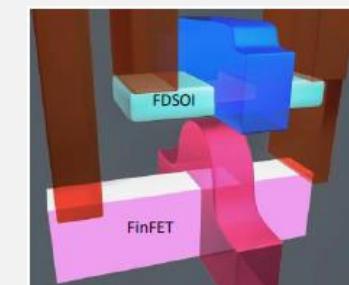
P. Batude et all, Advances in 3D CMOS Sequential Integration, IEEE 2009

- No degradation of the bottom characteristics due to top integration
- Shift of  $V_T$  due to the increase of the gate-metal work function through the annealing
- CMOS-Inverters and SRAM-Cells are working



P. Batude et all, Advances in 3D CMOS Sequential Integration, IEEE 2009

- ▶ Different channel materials for high mobility were integrated
- ▶ Heat difficulties in the bottom part
- ▶ Implementing FinFETs started



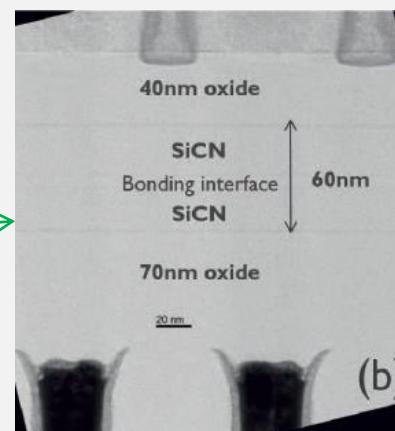
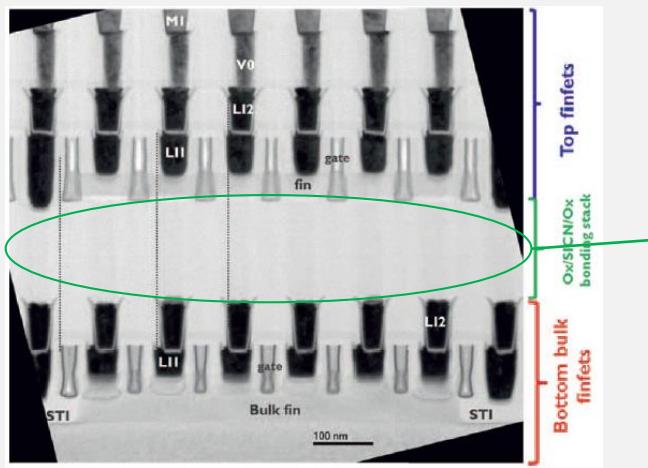
IBM 2009

First Demonstration of 3D stacked Finfets at a 45nm fin pitch and 110nm gate pitch technology on 300mm wafers.

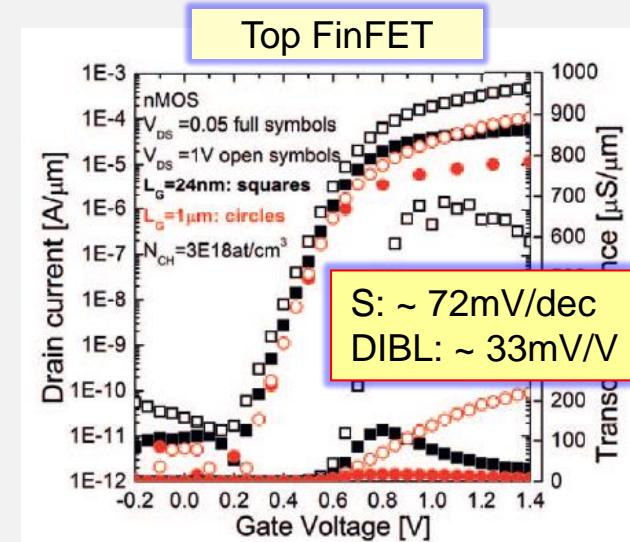
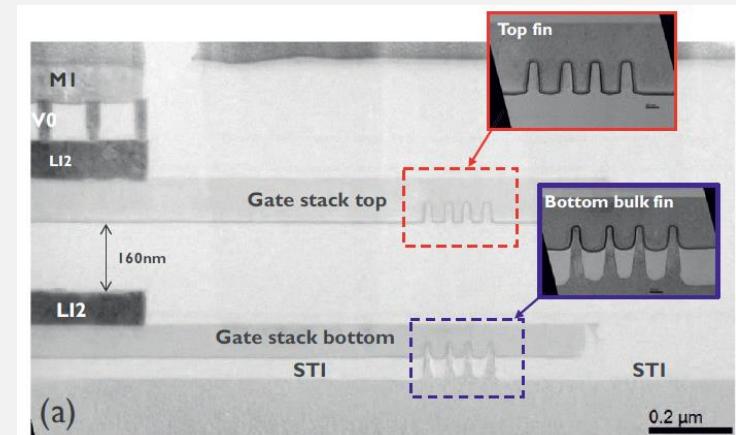
A. Vandooren, J. Franco, Z. Wu<sup>1</sup>, B. Parvais<sup>2</sup>, W. Li, L. Witters, A. Walke, L. Peng, V. Deshpande, N. Rassoul, G. Hellings, G. Jamieson, F. Inoue, K. Devriendt, L. Teugels, N. Heylen, E. Vecchio, T. Zheng, E. Rosseel, W. Vanherle, A. Hikavyy, G. Mannaert, B. T. Chan, R. Ritzenthaler, J. Mitard, L. Ragnarsson, N. Waldron, V. De Heyn, S. Demuyck, J. Boemmel, D. Mocuta, J. Ryckaert and N. Collaert.

IMEC, Kapeldreef 75, 3001 Leuven, Belgium email: [anne.vandooren@imec.be](mailto:anne.vandooren@imec.be)  
<sup>1</sup> also with KULEuven, Leuven, Belgium, <sup>2</sup> also with VUB, Brussels, Belgium

IEEE 2018



- Top FinFET shows good IV-characteristic
- No information about the bottom FinFET, no circuits
- First demonstration of stacked FinFETs through monolithic 3D-Integration



## 300mm Heterogeneous 3D Integration of Record Performance Layer Transfer Germanium PMOS with Silicon NMOS for Low Power High Performance Logic Applications

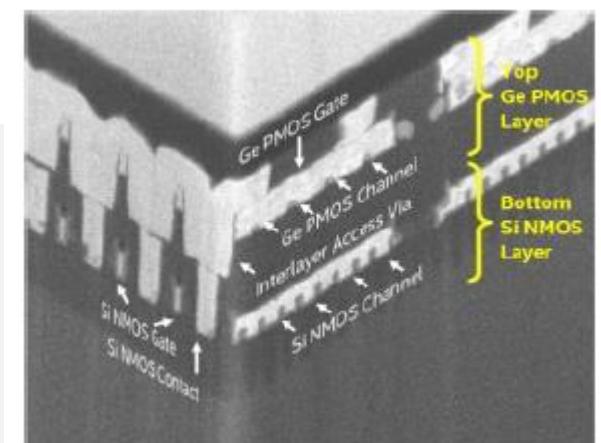
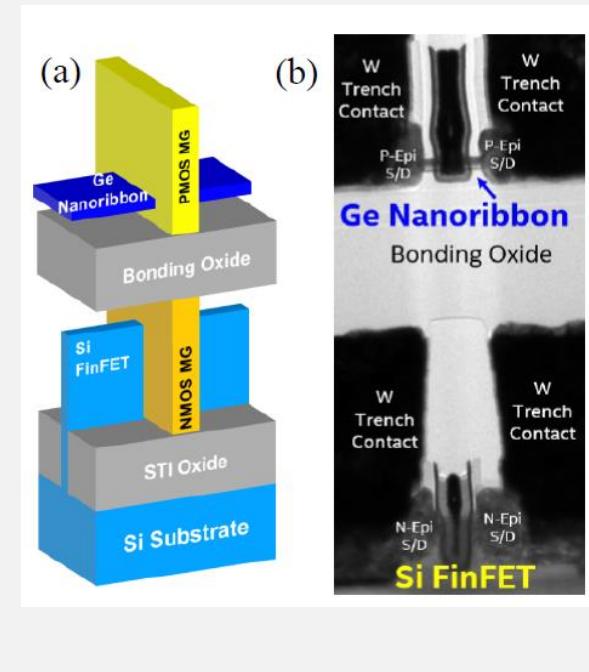
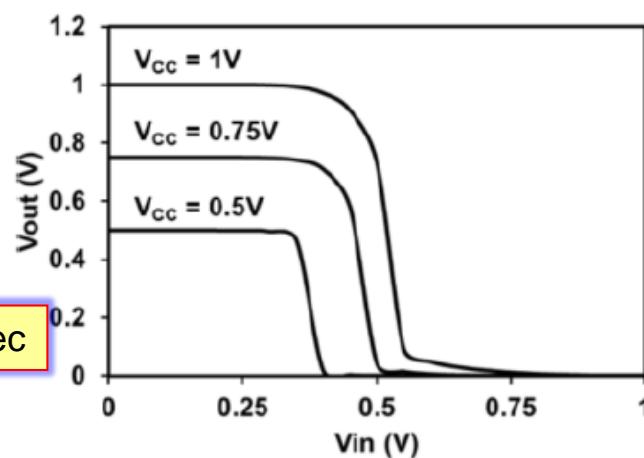
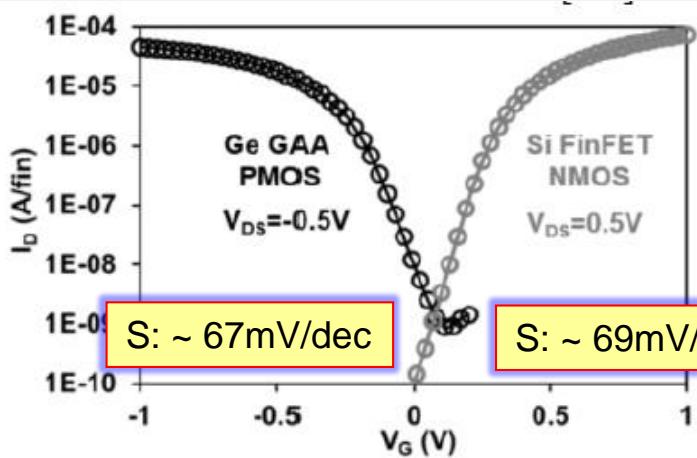
W. Rachmady, A. Agrawal, S.H. Sung, G. Dewey, S. Chouksey, B. Chu-Kung, G. Elbaz, P. Fischer, C.Y. Huang, K. Jun, B. Krist, M. Metz, T. Michaelos, B. Mueller, A.A. Oni, R. Paul, A. Phan, P. Sears, T. Talukdar, J. Torres,

R. Turkot, L. Wong, H. J. Yoo, and J. Kavalieros

Components Research, Technology Development, Intel Corporation, Hillsboro, OR 97124, USA

Contact: E-mail [willy.rachmady@intel.com](mailto:willy.rachmady@intel.com)

IEEE 2019



► Ge GAA shows good Subthreshold Slope,  $\text{GeO}_2$  ??

► No degradation of bottom FinFET due to low temperatures

► Working but not fully symmetrical Inverter

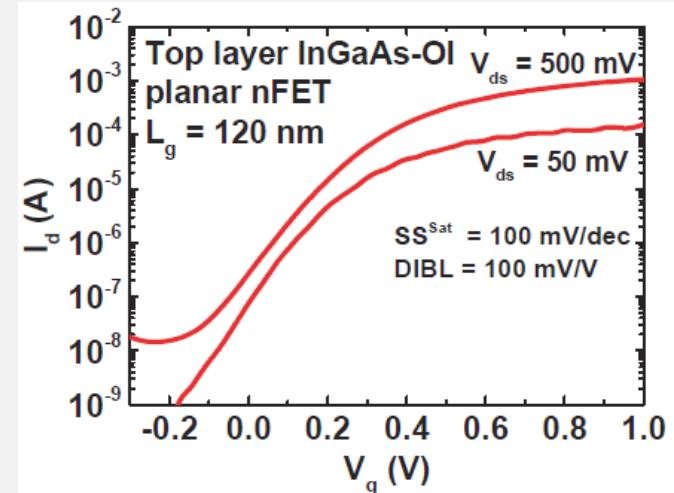
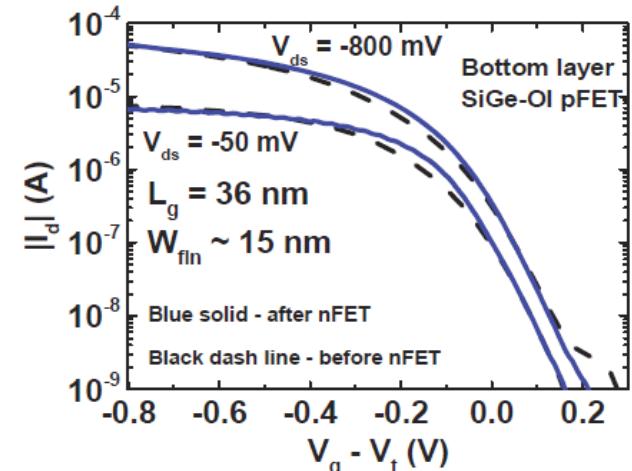
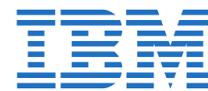
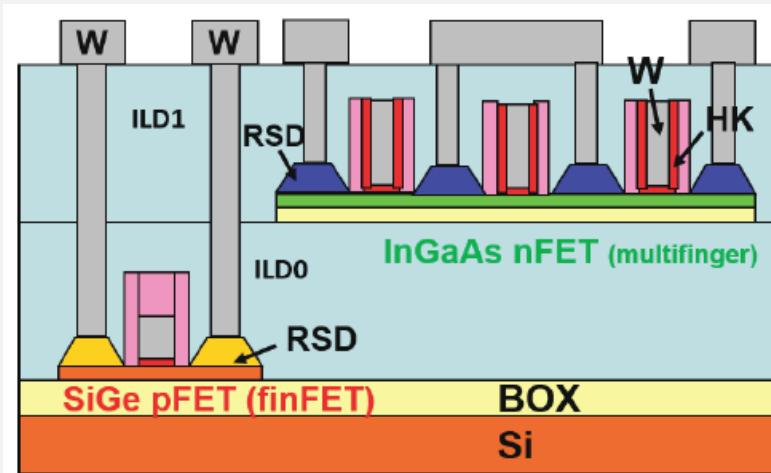
## First RF Characterization of InGaAs Replacement Metal Gate (RMG) nFETs on SiGe-OI FinFETs Fabricated by 3D Monolithic Integration

IEEE 2016

V. Deshpande, V. Djara, E. O'Connor, D. Caimi, M. Sousa, L. Czornomaz and J. Fompeyrine  
IBM Zurich Research Laboratory,  
Säumerstrasse 4, CH-8803 Rüschlikon,  
Switzerland

P. Hashemi, K. Balakrishnan

IBM T.J.Watson Research Center,  
1101 Kitchawan Rd., Route 134 Yorktown Heights, NY,  
USA

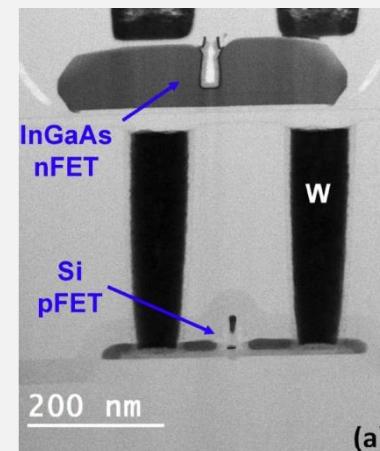
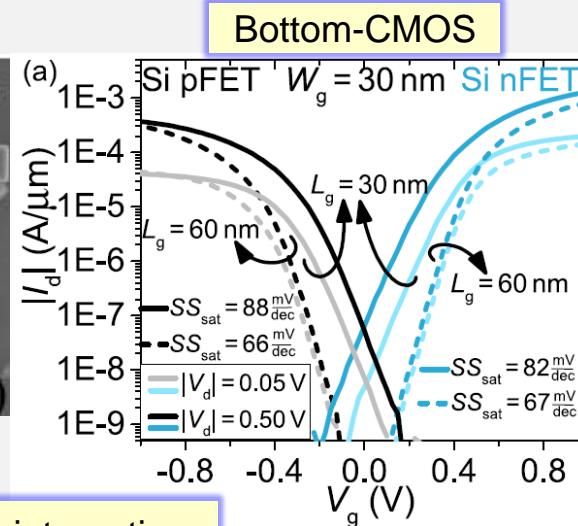
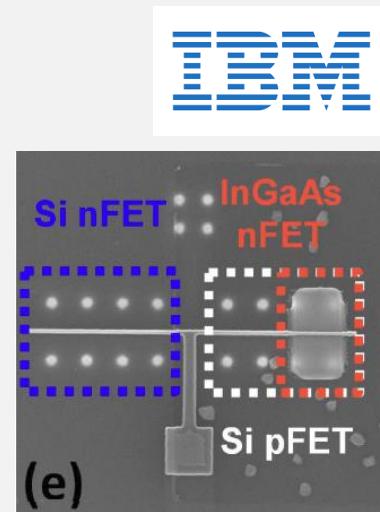
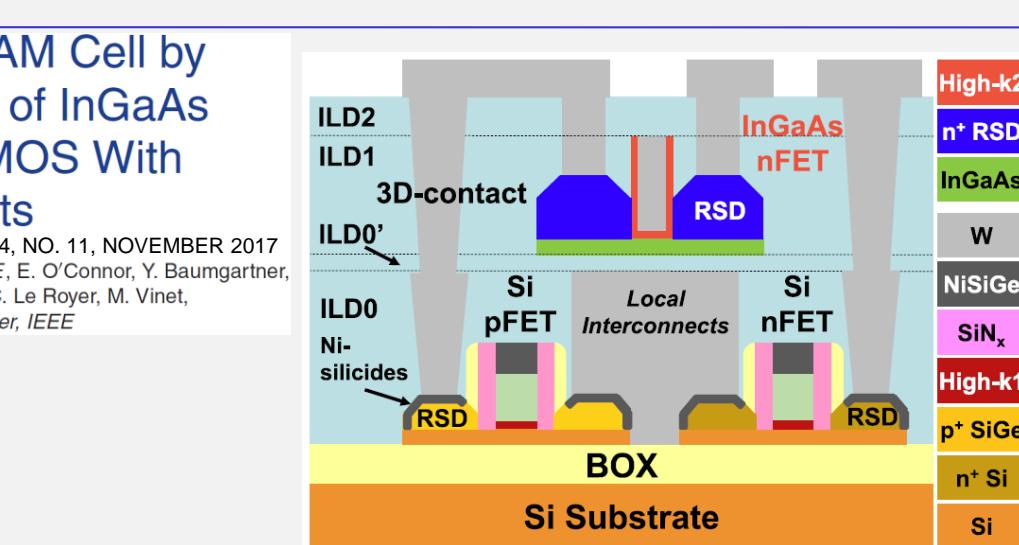
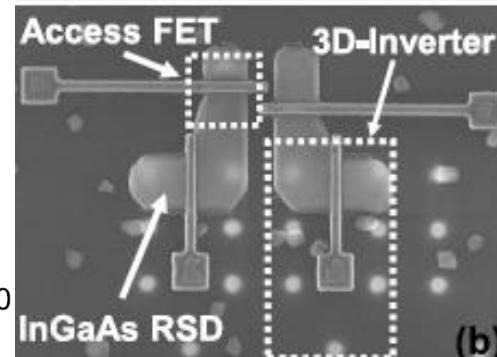
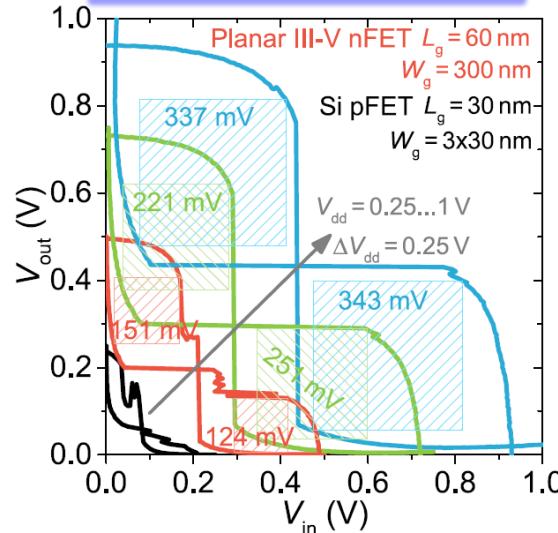


- ▶ Implementation of a pFinFET in the bottom layer
- ▶ Implementation of a InGaAs nFET in the top layer
- ▶ No degradation of the bottom characteristics due to top integration
- ▶ RF devices required for high-frequency analog or mixed signal circuits integrated on top of MOSFETs

## Demonstration of 3-D SRAM Cell by 3-D Monolithic Integration of InGaAs n-FinFETs on FDSOI CMOS With Interlayer Contacts

IEEE TRANSACTIONS ON ELECTRON DEVICES, VOL. 64, NO. 11, NOVEMBER 2017  
Veeresh Deshpande<sup>1</sup>, Member, IEEE, H. Hahn<sup>2</sup>, Member, IEEE, E. O'Connor, Y. Baumgartner,  
D. Caimi, M. Sousa, H. Boutry, J. Widiez, L. Brévard, C. Le Royer, M. Vinet,  
J. Pompeyrine, and L. Czornomaz, Member, IEEE

### SRAM Top-InGaAs Bottom-p-FET



No degradation of the bottom characteristics due to top integration

First demonstration of 3-DM SRAM cell with InGaAs nFET and FDSOI pFET

Gate-All-Around-FETs will come at the 5 nm/3 nm technology node (2020)

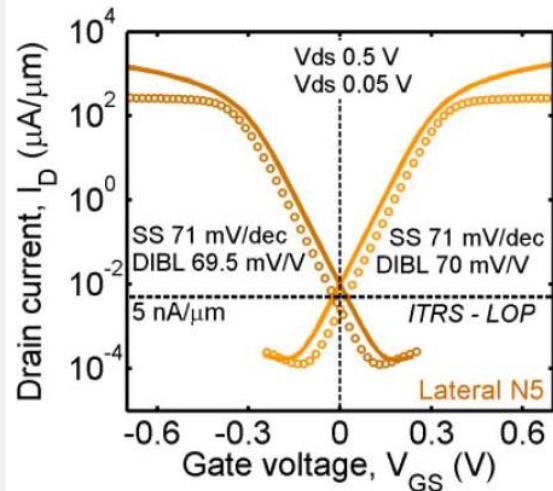
## Lateral versus Vertical Gate-all-around FETs for Beyond 7nm Technologies

D. Yakimets<sup>1,2</sup>, T. Huynh Bao<sup>1,3</sup>, M. Garcia Bardon<sup>1</sup>, M. Dehan<sup>1</sup>, N. Collaert<sup>1</sup>, A. Mercha<sup>1</sup>, Z. Tokei<sup>1</sup>, A. Thean<sup>1</sup>, D. Verkest<sup>1,3</sup>, and K. De Meyer<sup>1,2</sup>

IEEE 2014

<sup>1</sup> Imec, Kapeldreef 75, B-3001 Leuven, Belgium;

<sup>2</sup> Katholieke Universiteit Leuven, Dept. ESAT, Belgium; <sup>3</sup> Vrije Universiteit Brussel, Dept. ETRO, Belgium  
Tel: +32 16 28 31 67, Fax: +32 16 28 15 15, Email: [yakimets@imec.be](mailto:yakimets@imec.be)



- Simulation of Lateral GAA for 5 nm technology node
- IV-characteristic shows working transistor

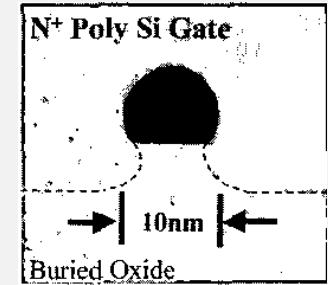
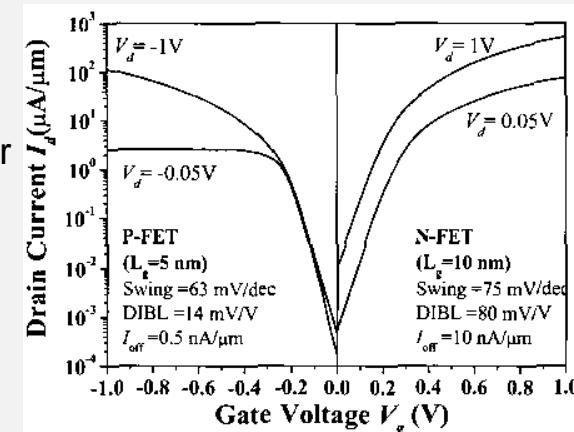
- In 2004 TSMC showed a working Nanowire-FET for 5 nm node
- No GAA

Fu-Liang Yang, Di-Hong Lee, Hou-Yu Chen, Chang-Yun Chang, Sheng-Da Liu, Cheng-Chuan Huang, Tang-Xuan Chung, Hung-Wei Chen, Chien-Chao Huang, Yi-Hsuan Liu, Chung-Cheng Wu, Chi-Chun Chen, Shih-Chang Chen, Ying-Tsung Chen, Ying-Ho Chen, Chih-Jian Chen, Bor-Wen Chan, Peng-Fu Hsu, Jyu-Horng Shieh, Han-Jan Tao, Yee-Chia Yeo, Yiming Li\*, Jian-Wen Lee\*, Pu Chen\*, Mong-Song Liang, and Chenming Hu

Taiwan Semiconductor Manufacturing Company, No. 8, Li-Hsin Rd. 6, Hsinchu Science Park, Hsin-Chu, Taiwan

\*National Nano Device Laboratories and National Chiao Tung University, Hsin-Chu, Taiwan  
Phone: 886-3-6665152, Email: [flyang@tsmc.com](mailto:flyang@tsmc.com)

IEE 2004

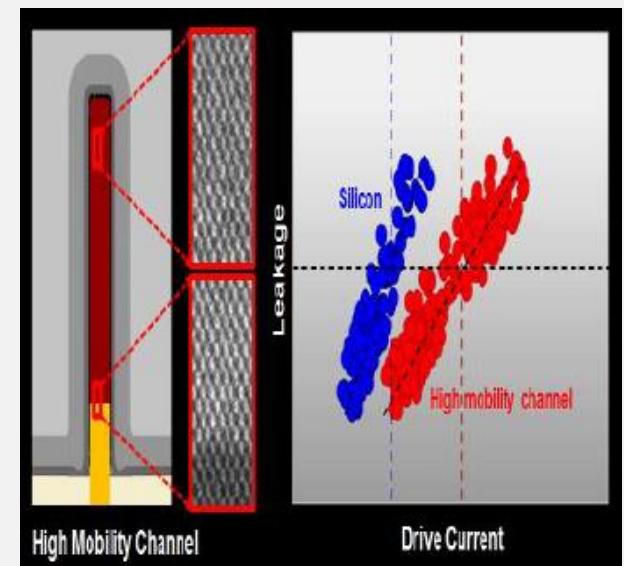
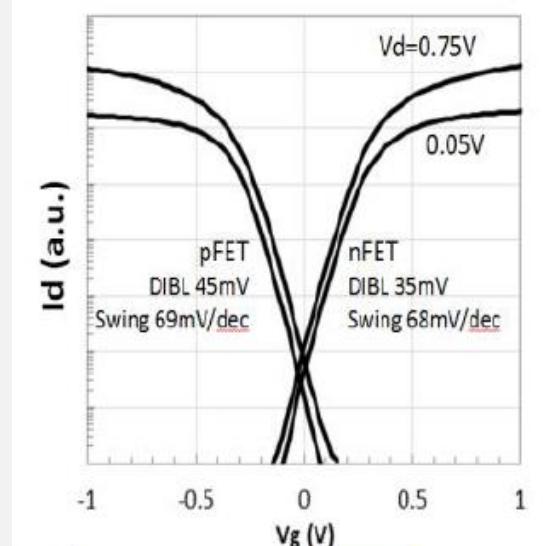
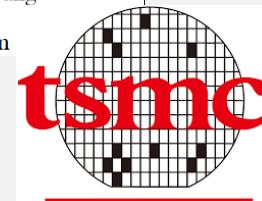
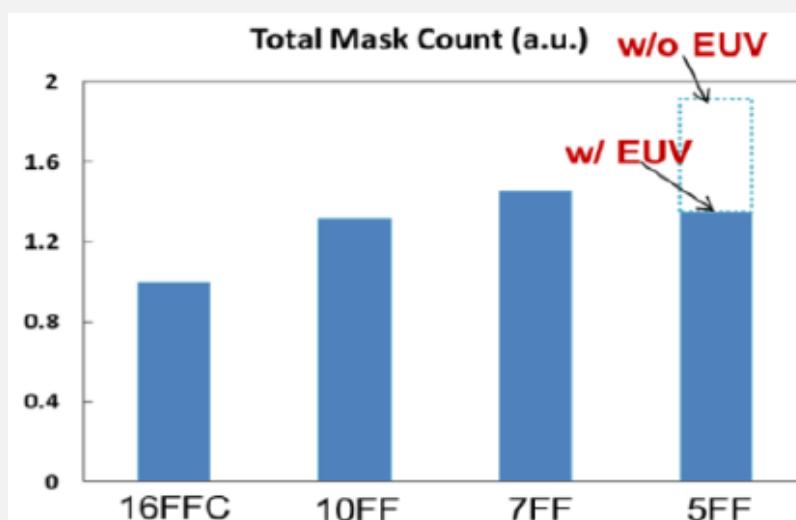


- ▶ Fabrication of Gate-All-Around FETs and design costs are expensive
- ▶ M3D needs implementation of GAA-FET for the future
- ▶ TSMC announced that in the 5 and 3 nm node they will use a FinFET

**5nm CMOS Production Technology Platform featuring full-fledged EUV, and High Mobility Channel FinFETs with densest  $0.021\mu\text{m}^2$  SRAM cells for Mobile SoC and High Performance Computing Applications**

Geoffrey Yeap, S.S. Lin, Y.M. Chen, H.L. Shang, P.W. Wang, H.C. Lin, Y.C. Peng, J.Y. Sheu, M. Wang, X. Chen, B.R. Yang, C.P. Lin, F.C. Yang, Y.K. Leung, D.W. Lin, C.P. Chen, K.F. Yu, D.H. Chen, C.Y. Chang, H.K. Chen, P. Hung, C.S. Hou, Y.K. Cheng, J. Chang, L. Yuan, C.K. Lin, C.C. Chen, Y.C. Yeo, M.H. Tsai, H.T. Lin, C.O. Chui, K.B. Huang, W. Chang, H.J. Lin, K.W. Chen, R. Chen, S.H. Sun, Q. Fu, H.T. Yang, H.T. Chiang, C.C. Yeh, T.L. Lee, C.H. Wang, S.L. Shue, C.W. Wu, R. Lu, W.R. Lin, J. Wu, F. Lai, Y.H. Wu, B.Z. Tien, Y.C. Huang, L.C. Lu, Jun He, Y. Ku, J. Lin, M. Cao, T.S. Chang, S.M. Jang

IEEE 2019 Taiwan Semiconductor Manufacturing Company, Hsinchu, Taiwan, R.O.C. Email: gyeap@tsmc.com



► High mobility channel, SiGe?

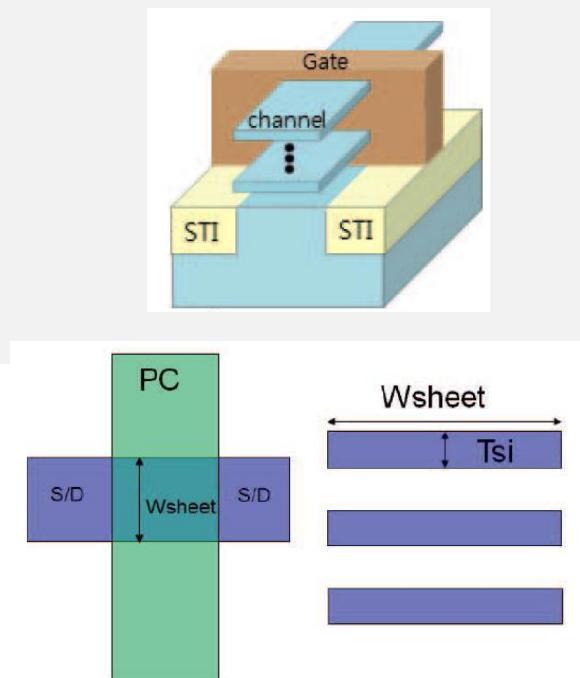
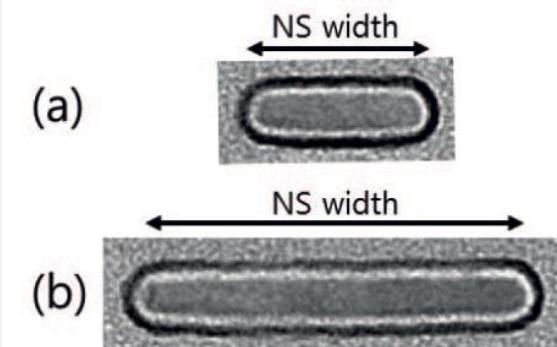
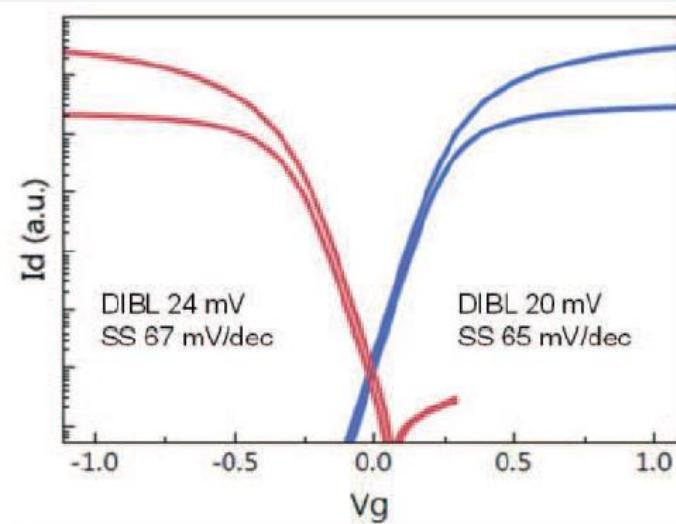
► SRAM cell size  $0.021\mu\text{m}^2$

► Mass production should start in 1H 2020

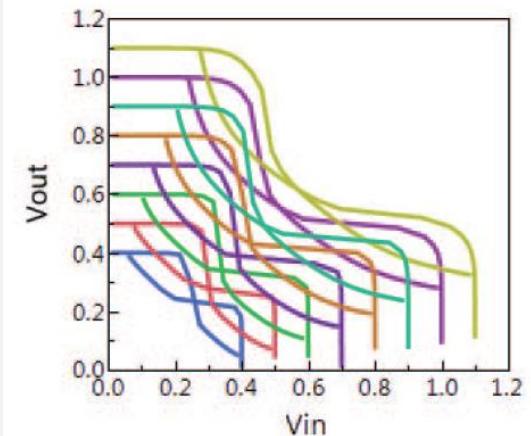
### 3nm GAA Technology featuring Multi-Bridge-Channel FET for Low Power and High Performance Applications

Geumjong Bae, D.-I. Bae, M.Kang, S.M.Hwang, S.S.Kim, B.Seo, T.Y.Kwon, T.J.Lee, C.Moon, Y.M.Choi, K.Oikawa, S.Masuoka, K.Y.Chun, S.H.Park, H.J.Shin, J.C.Kim, K.K.Bhuwalka, D.H.Kim, W.J. Kim, J.Yoo, H.Y.Jeon, M.S.Yang, S.-J.Chung, D.Kim, B.H.Ham, K.J.Park, W.D.Kim, S.H.Park, G.Song, Y.H.Kim, M.S.Kang, K.H.Hwang, C.-H.Park, J.-H.Lee, D.-W. Kim, S-M.Jung, H.K.Kang

Samsung R&D Center, Samsung Electronics, Hwasung-City, Gyeonggi-Do, Republic of Korea,  
email: [goldenbell.bae@samsung.com](mailto:goldenbell.bae@samsung.com)



Source: C.W. Yeung et al., Channel Geometry Impact and Narrow Sheet Effect of Stacked Nanosheet, IEEE 2018



- No details about channel material, channel length
- Dual-workfunction metallization
- Mass production should start in 2022

## 6.1 Overview

Feature size, pitch, node

## 6.2 State of the Art MOSFETs

Electrical and geometrical parameters

## 6.3 Mobility Enhanced MOSFETs

Strain, SiGe, Liners, Directed Channels, Gate-Last

## 6.4 High-k Metal-Gate MOSFETs

Poly-Depletion, FuSi

## 6.5 Vertical MOSFETs

## 6.6 Halo MOSFETs

## 6.7 SOI-MOSFETs

## 6.8 Multi-Gate MOSFETs

Fully-depleted MOSFETs, FinFETs, Nanowire FETs

## 6.9 Advanced Memory Devices

End of Planar technology ?

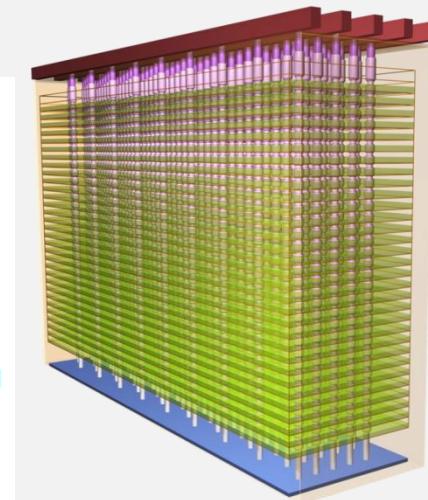
3d devices possible ?

### Intel promises 10TB+ SSDs thanks to 3D Vertical NAND flash memory

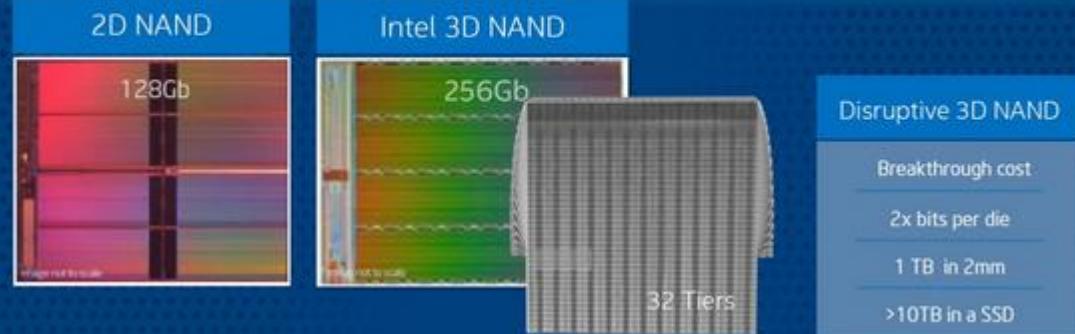
November 21st, 2014 at 5:54 pm -Author Anton Shilov

At present solid-state drives with extreme capacities are very expensive and even the best of them cannot match high-capacity hard disk drives for nearline storage applications. However, thanks to the evolution of NAND flash memory in general, and 3D vertical NAND (3D V-NAND) in particular, the situation may change soon and SSDs with 10TB or higher capacities will become reality.

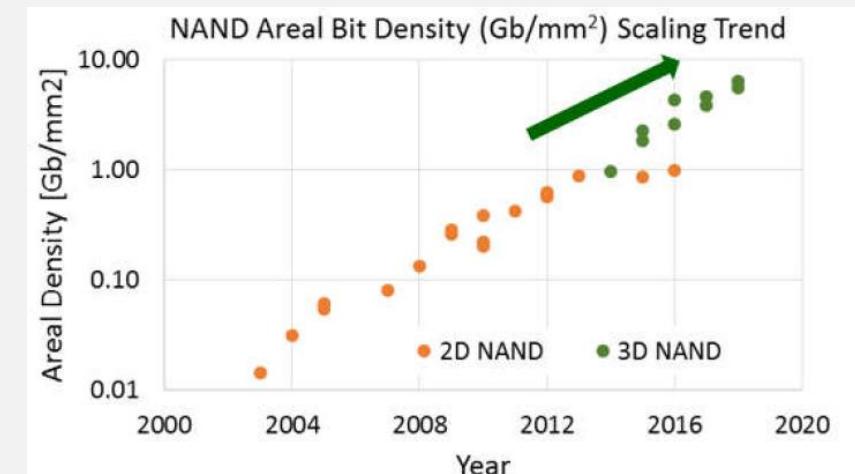
Intel Corp. revealed at its Investor Meeting 2014 event this week that in the second half of 2015 its joint venture with Micron Technology – Intel Micron Flash Technologies (IMFT) – will start mass production of 3D vertical NAND flash memory chips with up to 256Gb (multi-level cell, 2-bit-per cell) or 384Gb (triple-level cell, 3-bit-per cell) capacity. 3D V-NAND flash memory chips will feature 32-layer vertically stacked cell arrays that are “interconnected” using four billion through silicon vias (TSVs).



### NAND Leadership



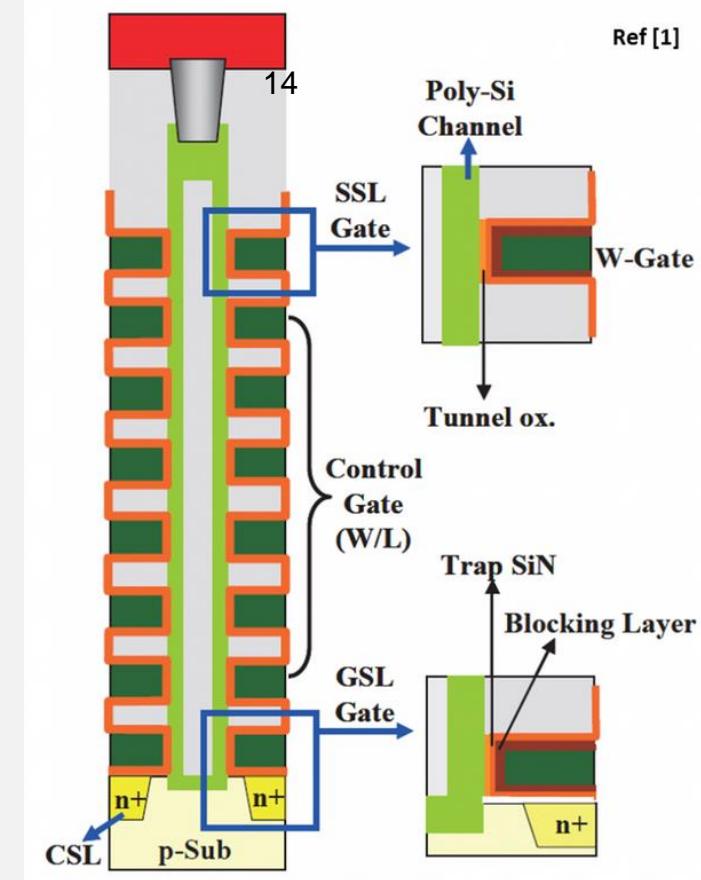
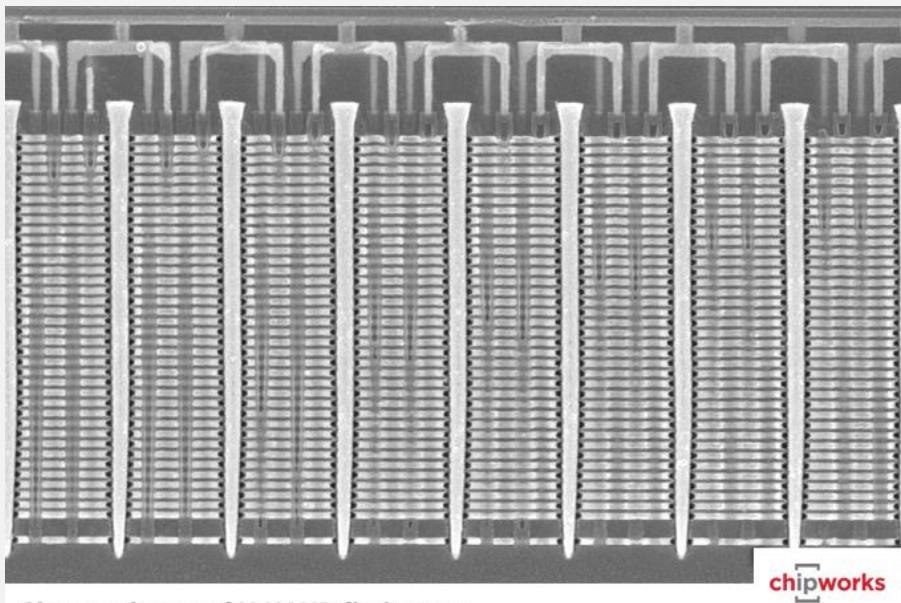
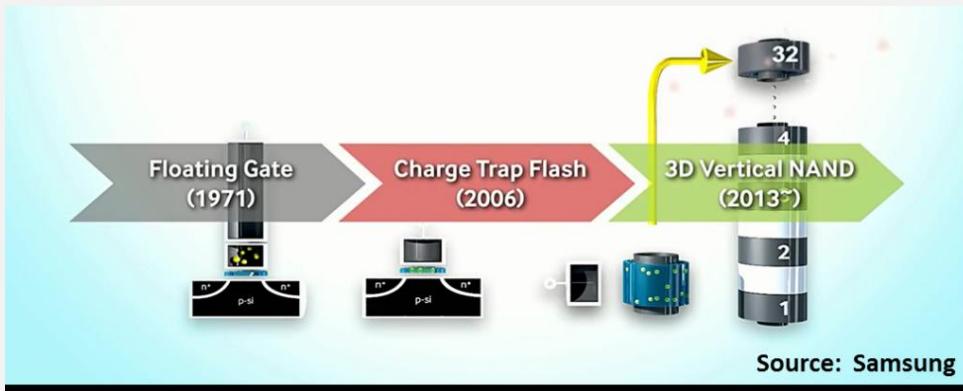
<https://www.micron.com/products/solid-state-storage/3d-nand-client-ssds>



K.Parat and A. Goda ,Scaling Trends in NAND Flash, IEEE2018

August 5, 2014

## The Second Shoe Drops - Samsung V-NAND Flash

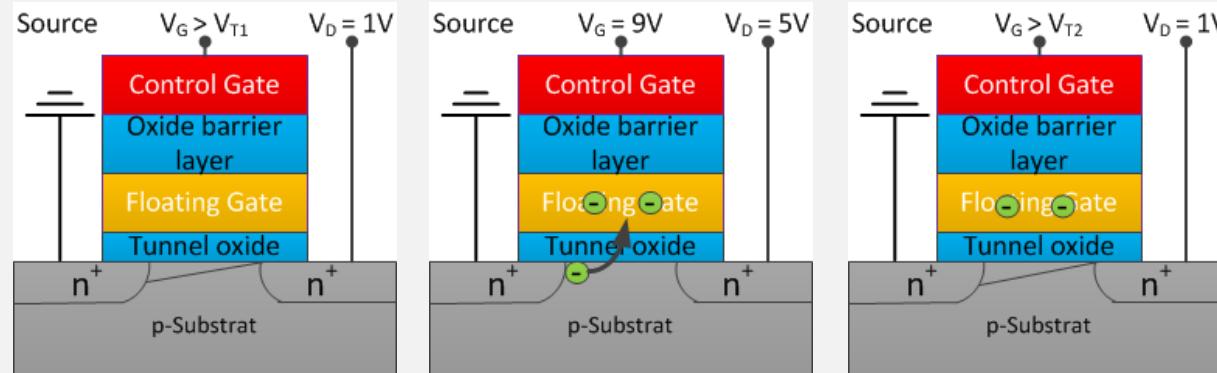


Planar technology is overruled by 3D-Technology

To understand the 3D-NAND, we look first at the 2D-NAND

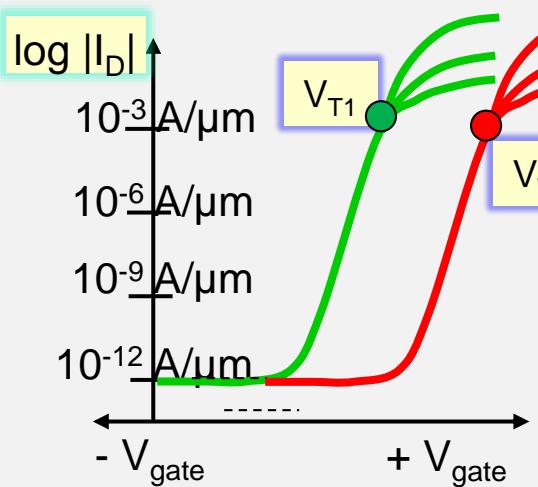
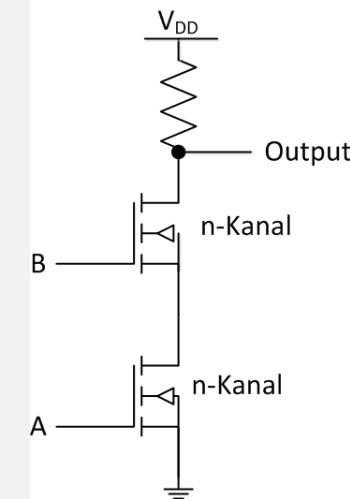
The Flash-NAND Memory consists out of Floating Gate MOSFETs which are linked together in a NAND-circuit.

## Floating Gate MOSFET



## NAND-circuit

Input		Output
A	B	
0	0	1
1	0	1
0	1	1
1	1	0

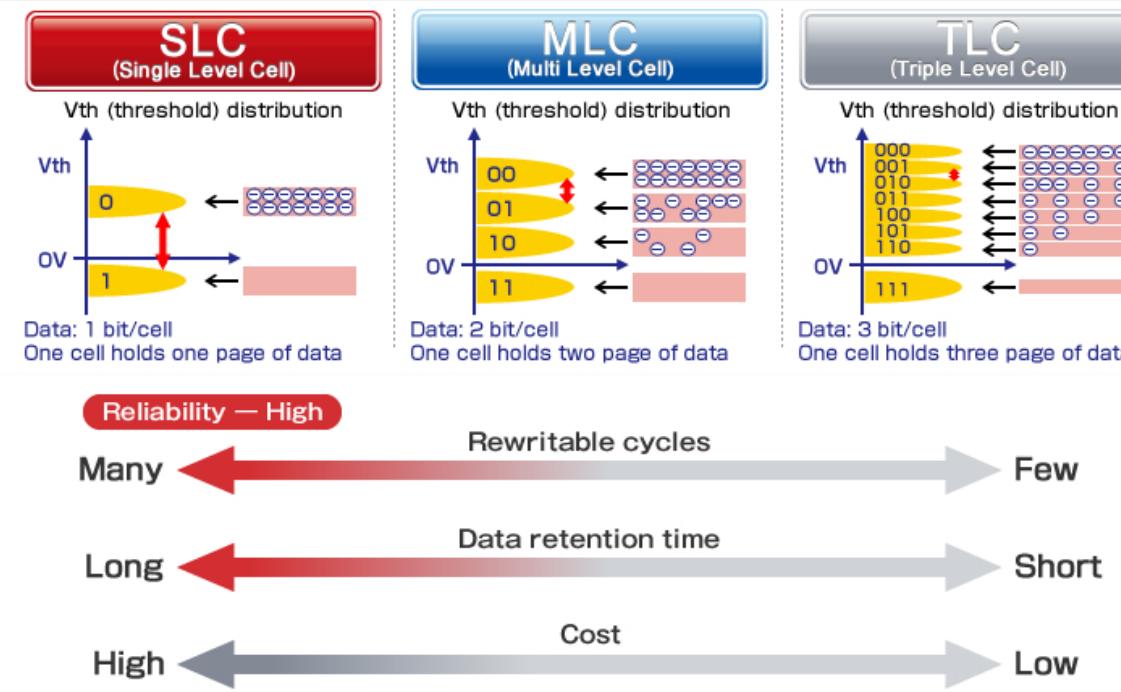


To store 1 the electrons have to tunnel into the Floating Gate

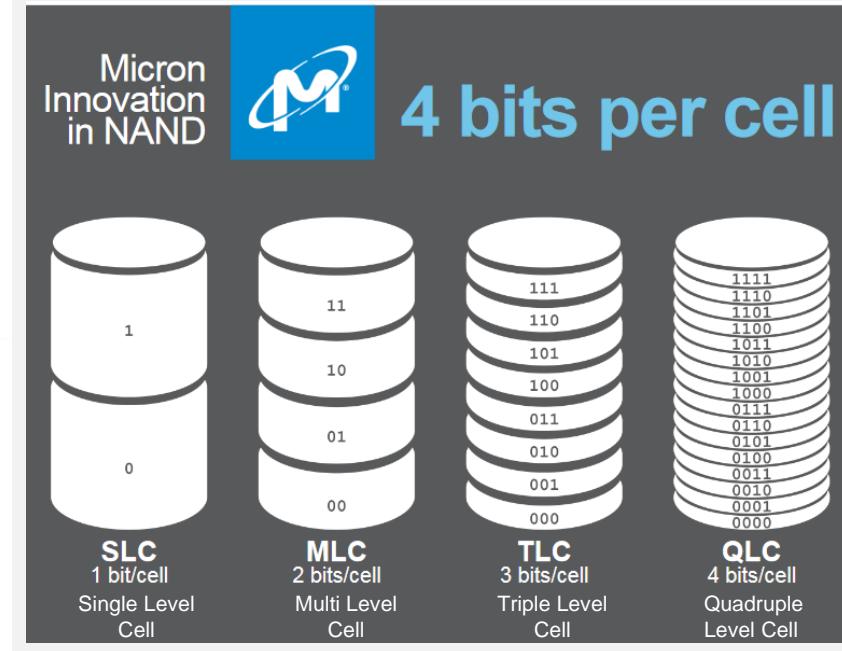
Stored electrons in the Floating Gate lead to a higher Threshold Voltage  $V_{T2}$

Only when both transistors are on, the output in the NAND-circuit changes

In the literature there is also sometimes written about multilevel cells.



[http://panasonic.net/avc/sdcard/industrial\\_sd/img/lineup\\_img01.jpg](http://panasonic.net/avc/sdcard/industrial_sd/img/lineup_img01.jpg)



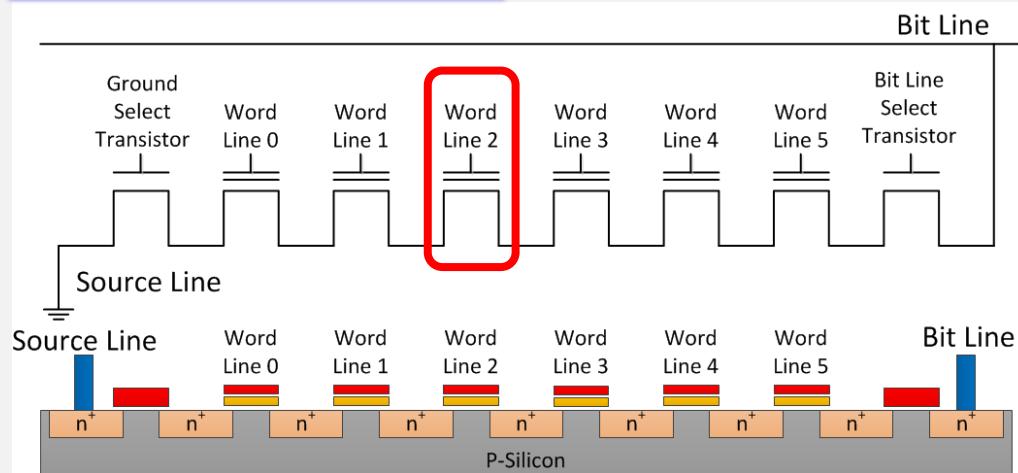
All three/four cells are still one Floating Gate transistor

MLC can save 2 bits in one single Floating Gate, TLC can save 3 bits in one Floating Gate, QLC can save 4 bits

Cells more sensitive to physical degradation and has lower endurance

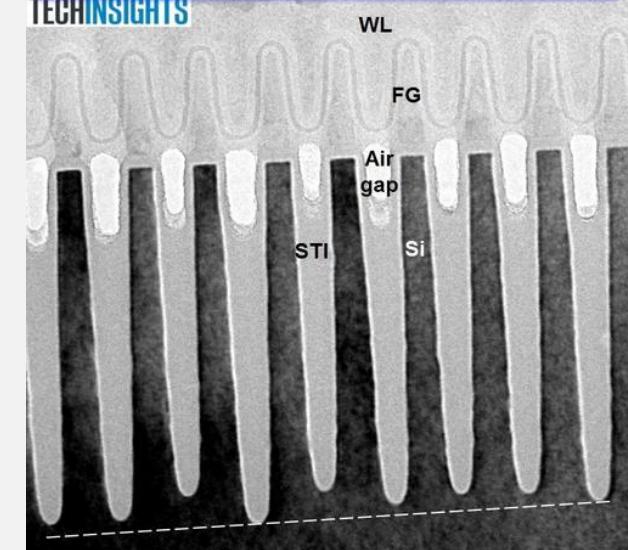
How does now the memory look like?

## Working principle with SLC

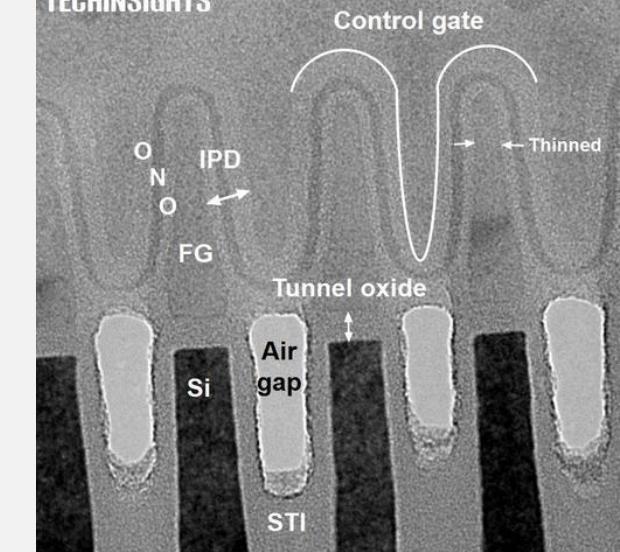


2015 15/16 nm 2D-NAND Hynix

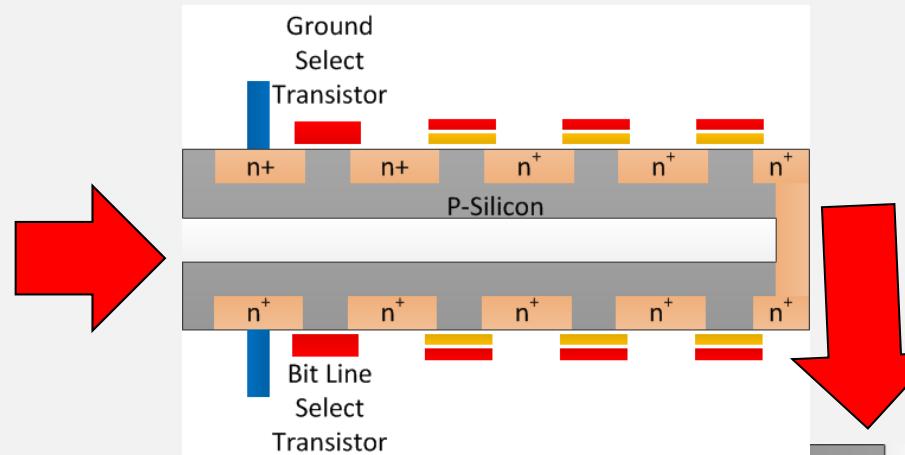
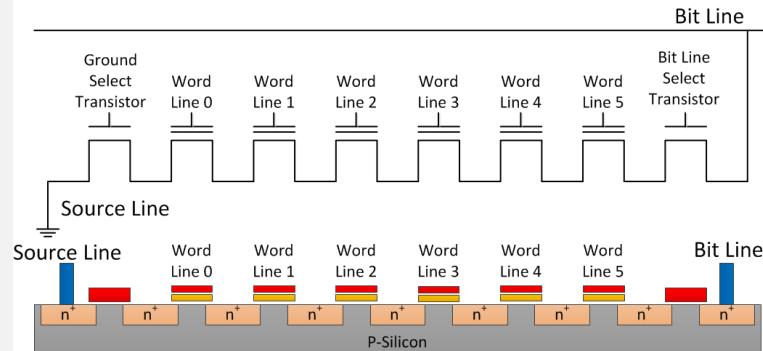
TECHINSIGHTS



TECHINSIGHTS

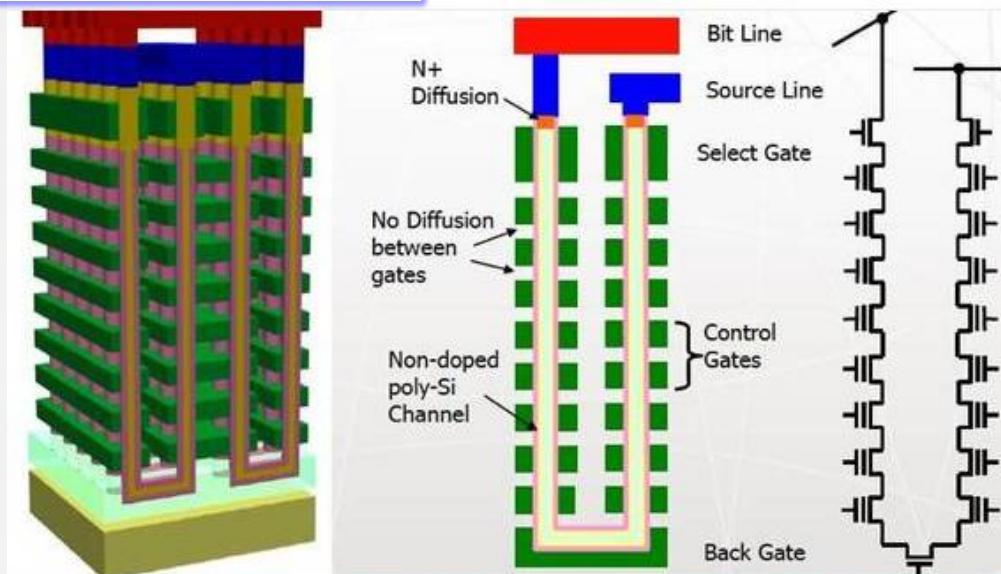


2D-NAND Memories are used till today

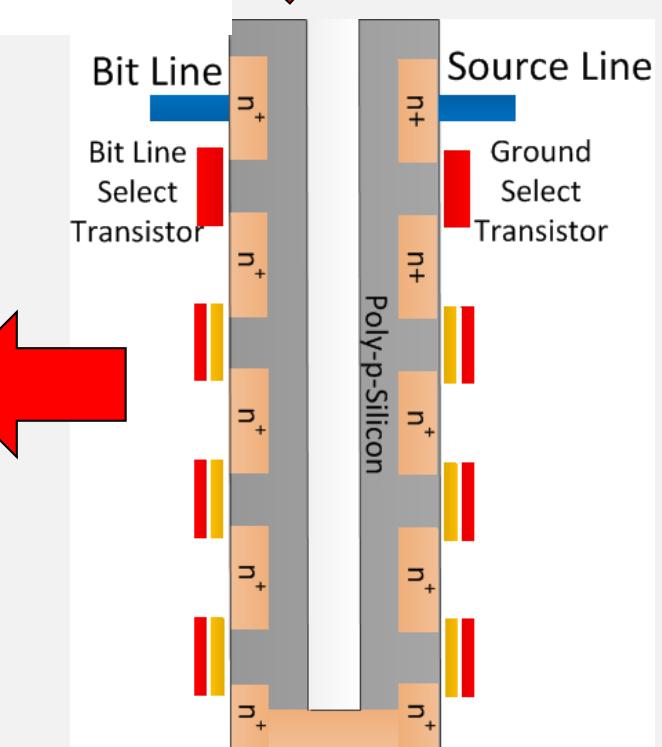


Due to stacking,  
change from single  
crystalline substrate  
to poly crystalline  
devices

3D-NAND by Toshiba/San Disk

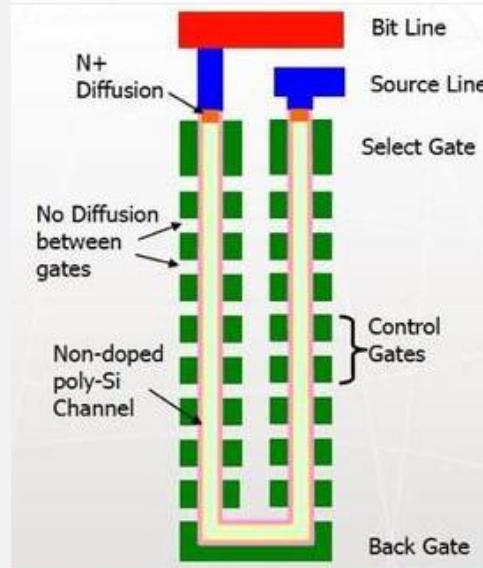


[http://www.theregister.co.uk/2013/07/23/sandisk\\_takes\\_the\\_bics\\_route\\_to\\_3d/](http://www.theregister.co.uk/2013/07/23/sandisk_takes_the_bics_route_to_3d/)

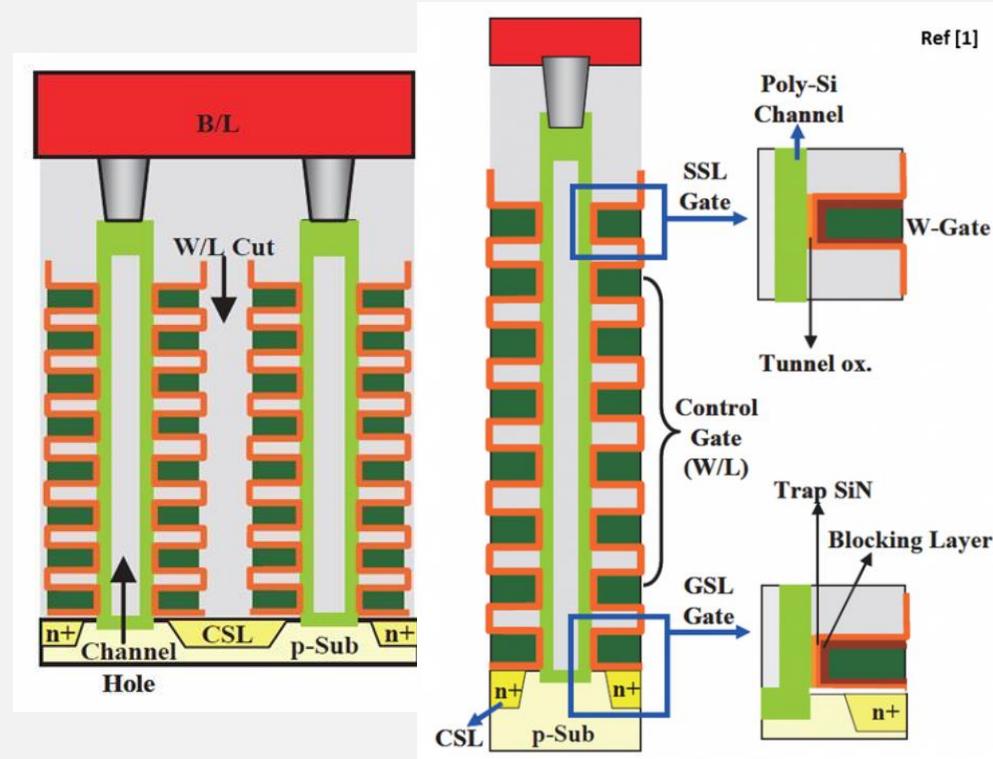


der Bundeswehr

3D-NAND by Toshiba/SanDisk



3D-NAND by Samsung



3D-structure can vary from company to company

Samsung is using  $\text{Si}_3\text{N}_4$  as trapping layer instead of polysilicon

Same working principle as the 2D-NAND

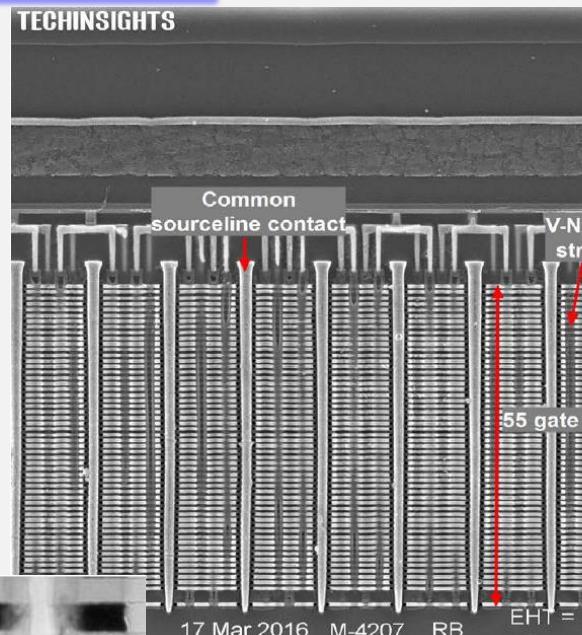
In production since 2014

GSL: Ground Select Line  
SSL: Source Select Line  
CSL: Connected Source Line  
WL: Word Line

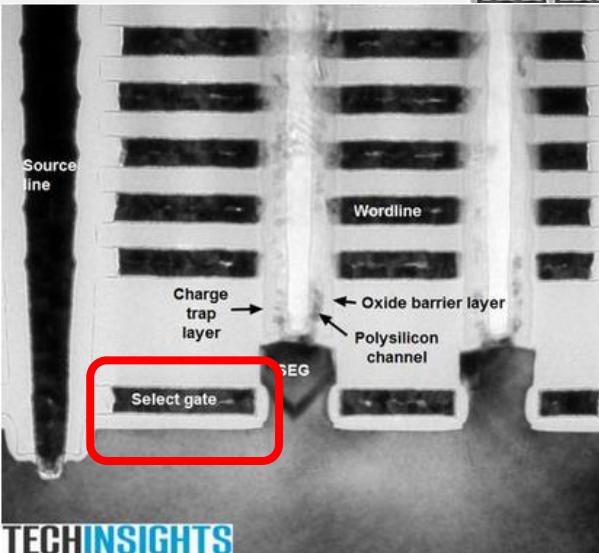
## 2016 Samsung SSD with 3D-V-NAND

Samsung SSD 960 EVO M.2

Data Sheet Rev.1.0 (October 2016)



	Micron 16 nm	Hynix 16 nm	Samsung 16 nm	Samsung 48L V-NAND	Samsung 64L 2018	Toshiba 128L 2019
Year	2014	2014	2015	2016	20	19
Process node (nm)	16	16	16	21	512	512
Capacity (Gb)	128	64	64	256		
Die area (mm <sup>2</sup> )	176	93	86.4	99		
Density (Mb/mm <sup>2</sup> )	730	690	740	2,600	3400	7800



3D-NAND (21 nm node) has a 3.5 times higher memory density than 2D (16 nm node)

Charge trap layers are used to store electrons

Current in the polysilicon channel is decreasing  
No long term scaling

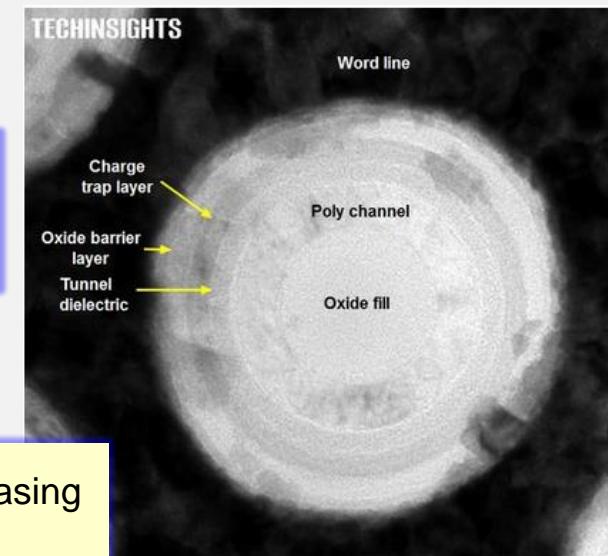
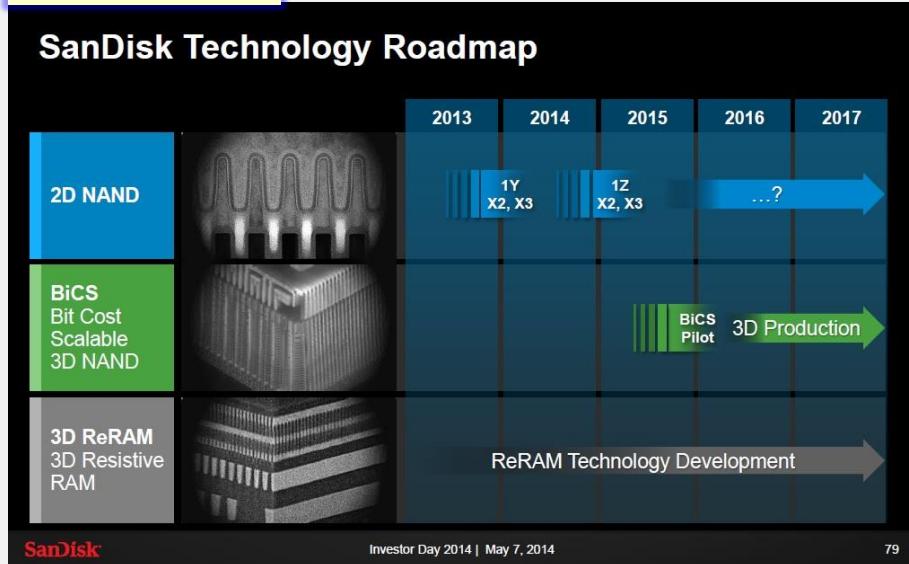
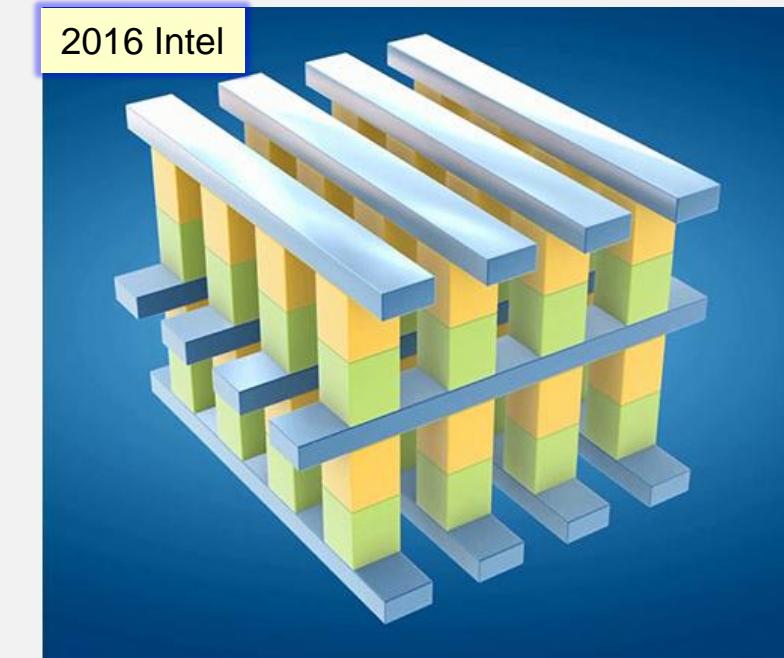


Figure 8: TEM Plan View of V-NAND String (Source: TechInsights)

2014 SanDisk



2016 Intel



3D XPoint™ Technology

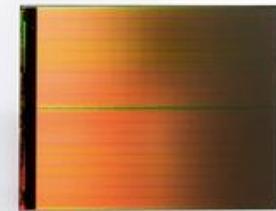
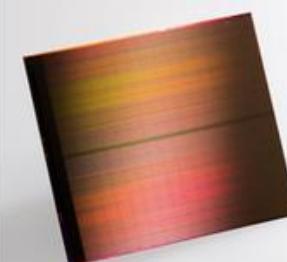


► NAND is coming to an end because of reliability and degrading performance

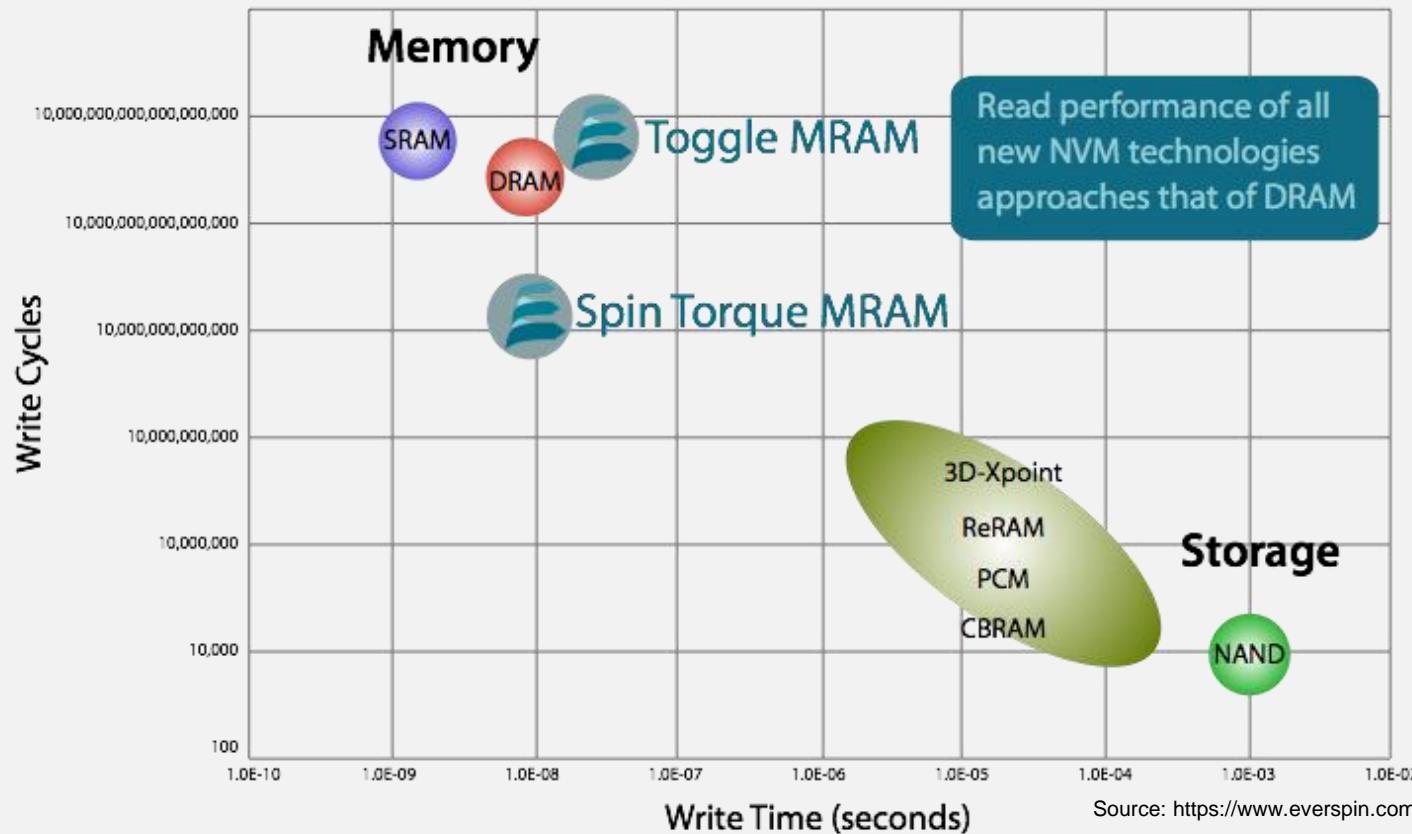
► All big companies see the RERAM/RRAM (Resistive RAM) as an alternative

► Intel/Micron announced 2015/2016 their 3D-Xpoint Technology

128 Gbit memory

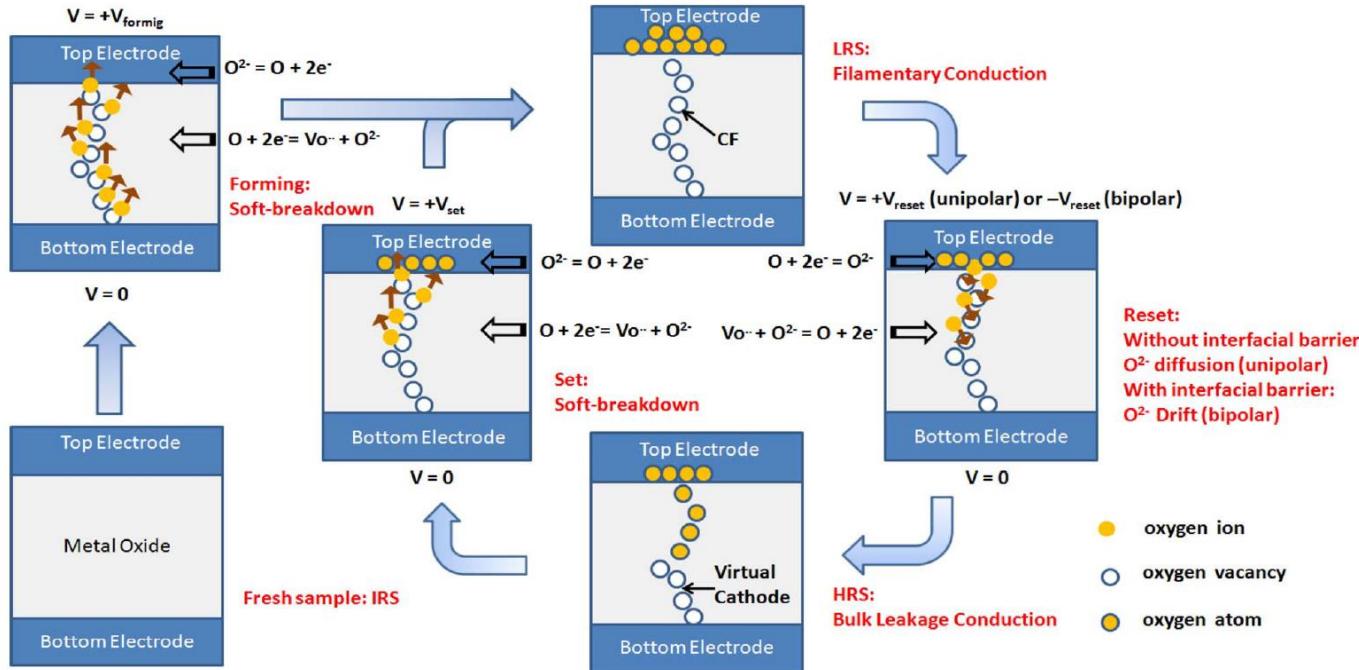


[http://www.eetimes.com/document.asp?doc\\_id=1327289&page\\_number=2](http://www.eetimes.com/document.asp?doc_id=1327289&page_number=2)



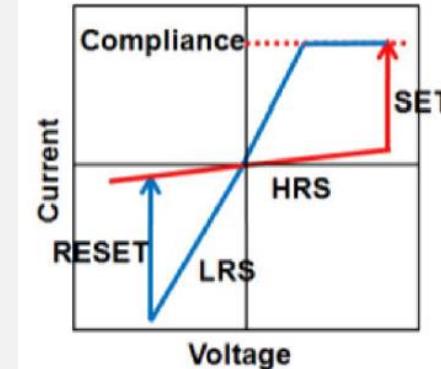
- ▶ ST-MRAM seems to be a good candidate to replace DRAM/SRAM, because it is non-volatile and scalable
- ▶ RRAM/PCM seems a good candidate for Storage

## RRAM mechanism

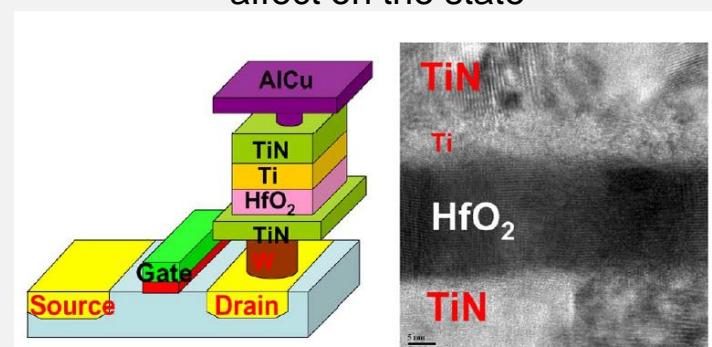


H.-S. Philip Wong et. all, Metal–Oxide RRAM, IEEE 2012

- Device structure is metal insulator (i.e.  $\text{HfO}_2$ ) metal
- Formation process:
  - High electric field: oxygen atoms drift to the anode
  - Electrode/oxide interface behaves like an oxygen reservoir
  - Localized deficiency of oxygen leads to a conductive filament
  - LRS state



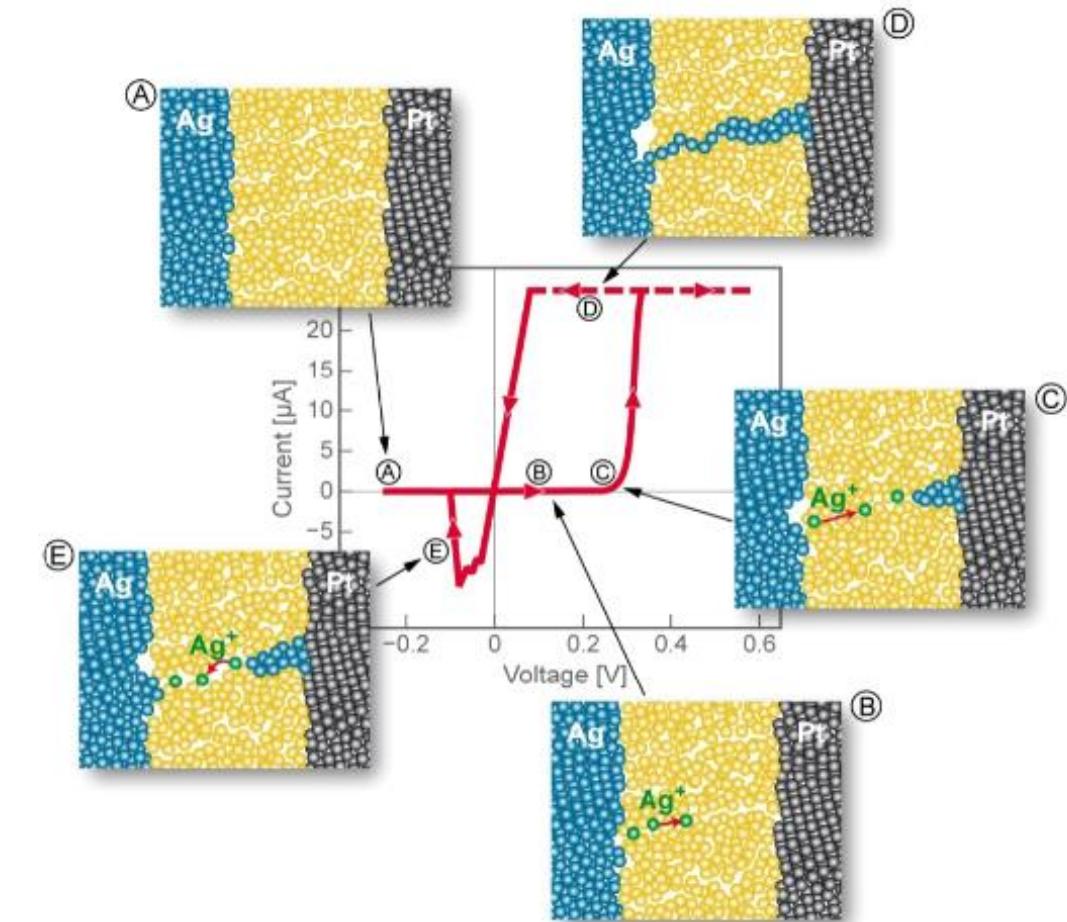
- Two states:
  - HRS high-resistance state
  - LRS low-resistance state
- Set process:  
Switching from HRS to LRS
- Reset process:  
Switching from LRS to HRS
- Read process:  
Small voltage which has no affect on the state



H.-S. Philip Wong et. all, Metal–Oxide RRAM, IEEE 2012

## Conductive Bridging RAM (CBRAM)

- Device structure is metal insulator(i.e.  $\text{SiO}_2$ ) metal
    - One electrode is an electrochemically active electrode (Ag, Cu, Ni)
    - Second electrode is an electrochemically inert counter electrode (Pt, W)
    - Solid electrolyte ( $\text{SiO}_2$ ,  $\text{TiO}_2$  and more)
- A) High-resistance state (OFF-state)
  - No electrodeposit on the inert electrode
- B) SET-process
  - Positive voltage on the Ag electrode
  - $\text{Ag}^+$  ions drift through the thin film to the Pt-electrode
  - ElectrocrySTALLization of Ag on the Pt
- C) ElectrocrySTALLization leads to the formation of a metal filament
- D) When the filament made a contact to the Ag, the cell is switched on
- E) RESET-process
  - Cell stays on till a negative voltage is applied
  - Filament is broken



Valov et. all, Electrochemical metallization memories-fundamentals, applications, prospects IOP 2011



For a new cell an electroforming process has to be made

## Phase Change Memory (PCM)

- Phase change between amorphous state and crystalline state caused by Joule heating
- Change in electrical resistance
- Material: chalcogenide materials like GeSbTe

### A) High-resistance state (OFF-state)

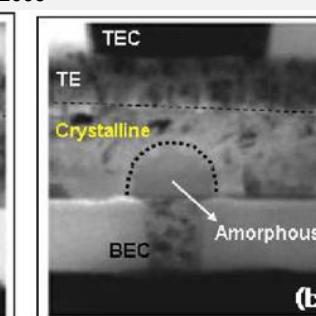
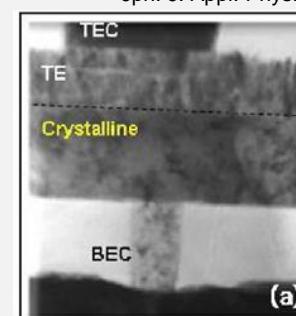
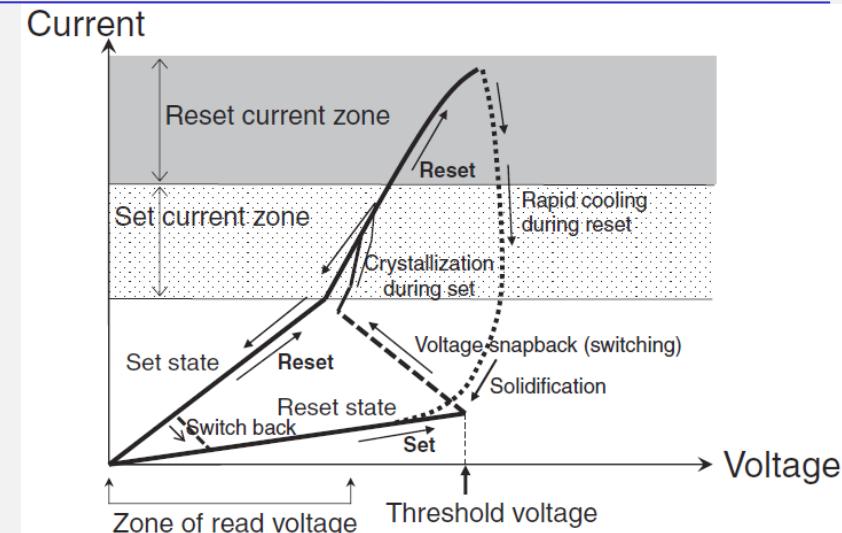
- Material is amorphous and has a high resistance

### B) Set state

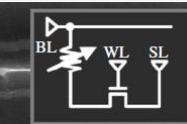
- Applied voltage is higher than the threshold voltage (lower amplitude with long pulse)
- Electrons are injected into the material
- Electrons cause impact ionization and carrier multiplication by colliding with defects
- Joule heating occurs
- Material reaches a temperature between 300°C and 400°C
- Material changes to crystalline state
- Lower resistance, higher current

### C) Reset state

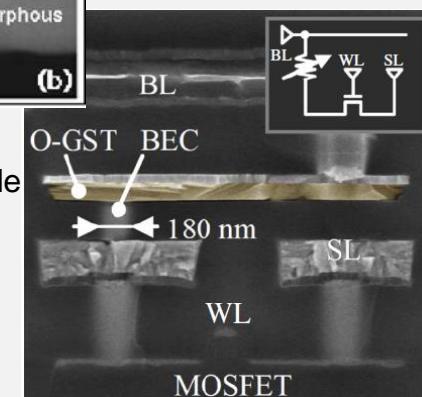
- Voltage is raised (high amplitude with short pulse)
- Current gets higher
- Joule heating makes the crystal melt
- Material becomes amorphous through lowering the voltage rapidly (600°C)



D. Han et all, Recent Advances in High Density Phase Change Memory (PRAM), IEEE2007

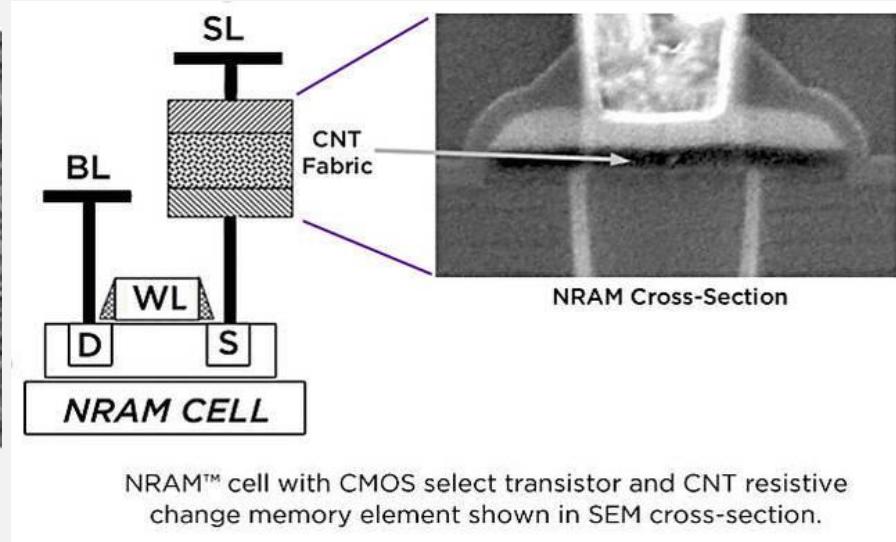
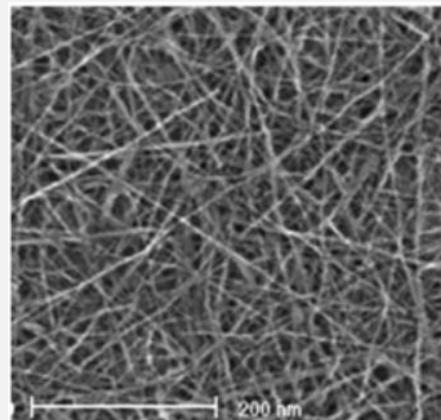


GST = GeSbTe  
BEC = Bottom Electrode  
BL = Bit Line  
WL = Word Line  
SL = Source Line



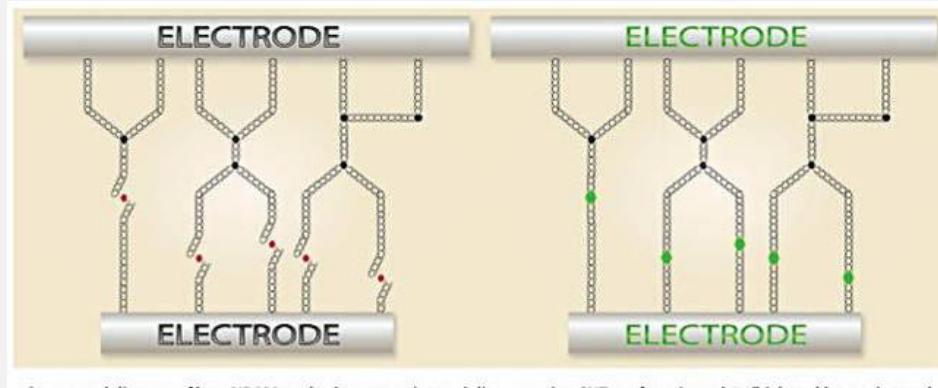
Chalcogenide materials not used in CMOS until now

### Carbon Nanotube Random Access Memory (NRAM)



NRAM™ cell with CMOS select transistor and CNT resistive change memory element shown in SEM cross-section.

[https://www.theregister.co.uk/2016/08/31/nram\\_dev\\_nantero\\_signs\\_fujitsu/](https://www.theregister.co.uk/2016/08/31/nram_dev_nantero_signs_fujitsu/), 2017

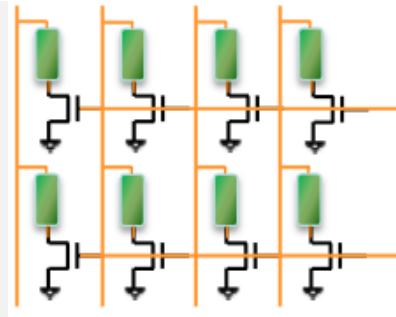
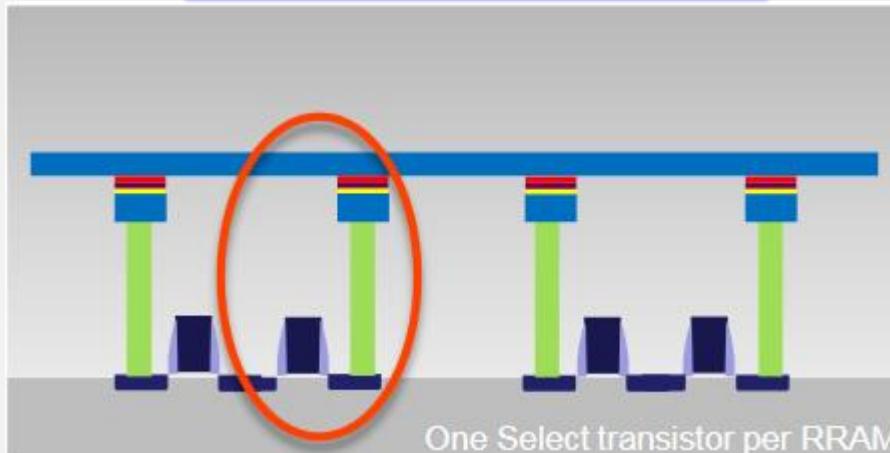


Conceptual diagram of how NRAM works, by connecting and disconnecting CNTs to form 0s and 1s (high and low resistance).



Till today yield and reliability problems

1 Transistor with one RRAM cell

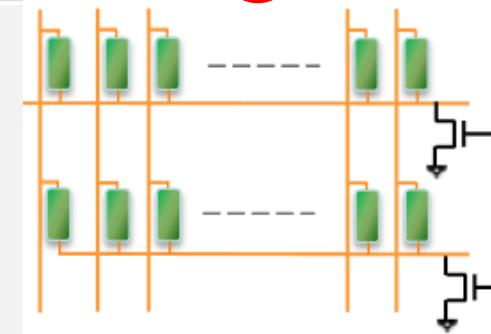
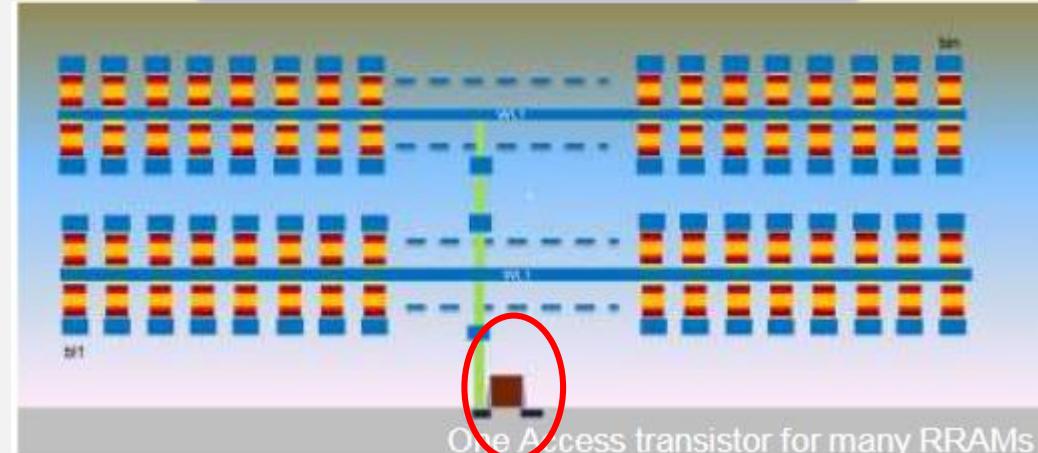


Low latency

Cell size dominated by transistor

Embedded memory applications

1 Transistor with several RRAM cells



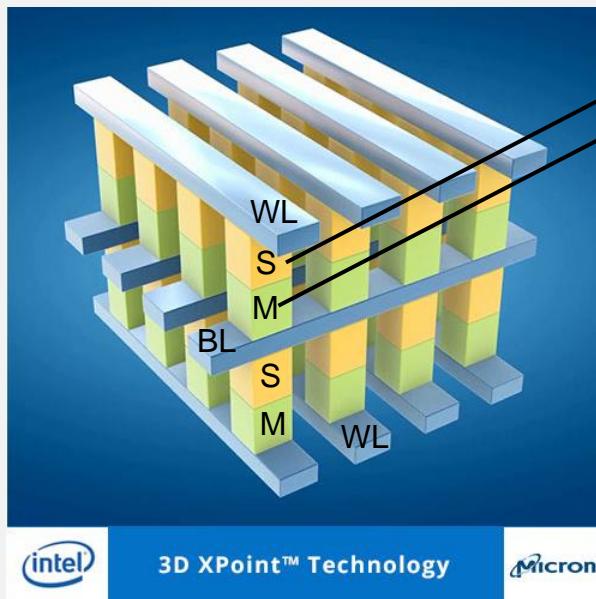
Cost effective, high area density

Implementing a crossbar scheme

Storage devices, SSDs

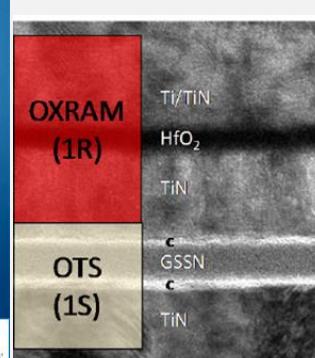
Source: <http://www.crossbar-inc.com/assets/resource/presentation/FMS2014-Slides-Breakthrough-3D-Crossbar-ReRAM-Technology.pdf>

## Working principle of Crossbar Scheme



Cell consists of:

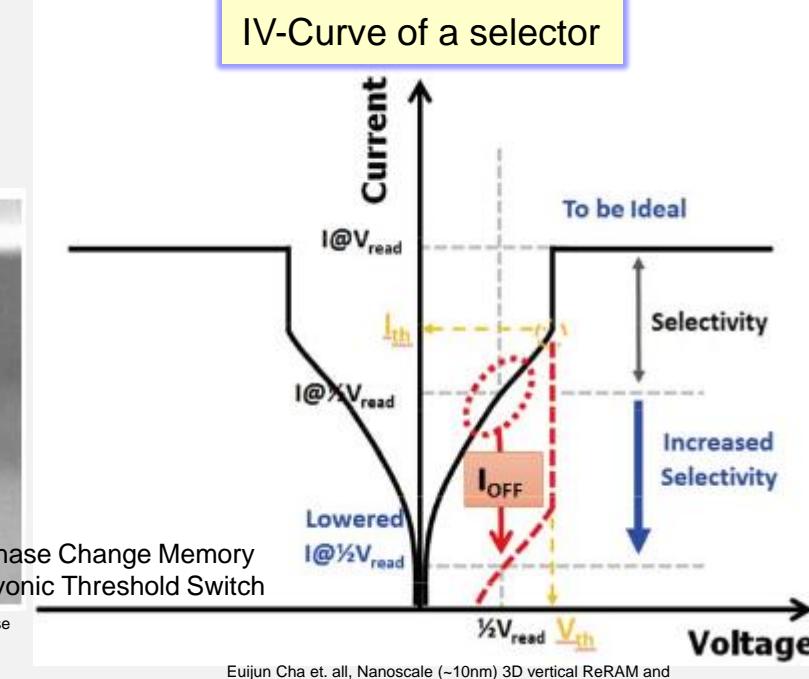
- 1. Selector (Switch element)
- 2. Memory (RRAM, PCM)



D. Alfaro Robayo et. all, Reliability and Variability of 1S1R OxRAM-OTS for High Density Crossbar Integration, IEEE 2019



D. Kau et. all, A stackable cross point phase change memory, IEEE 2009



### Characteristics of a selector

- High On-state current
- Low Off-state current
- Process compatibility with 3D-stacking
- Bidirectional operation
- Voltage compatibility with memory element



### Selectors

- Polysilicon diodes
- Ovonic threshold switching materials (OTS)
- Metal-oxide (Metal-Insulator-Metal) Schottky barrier

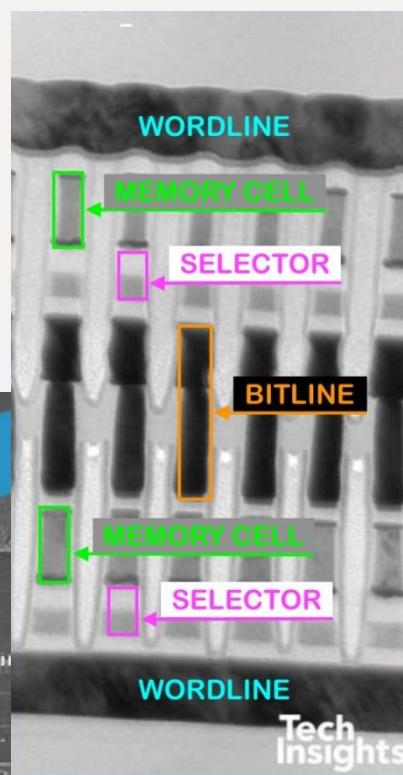
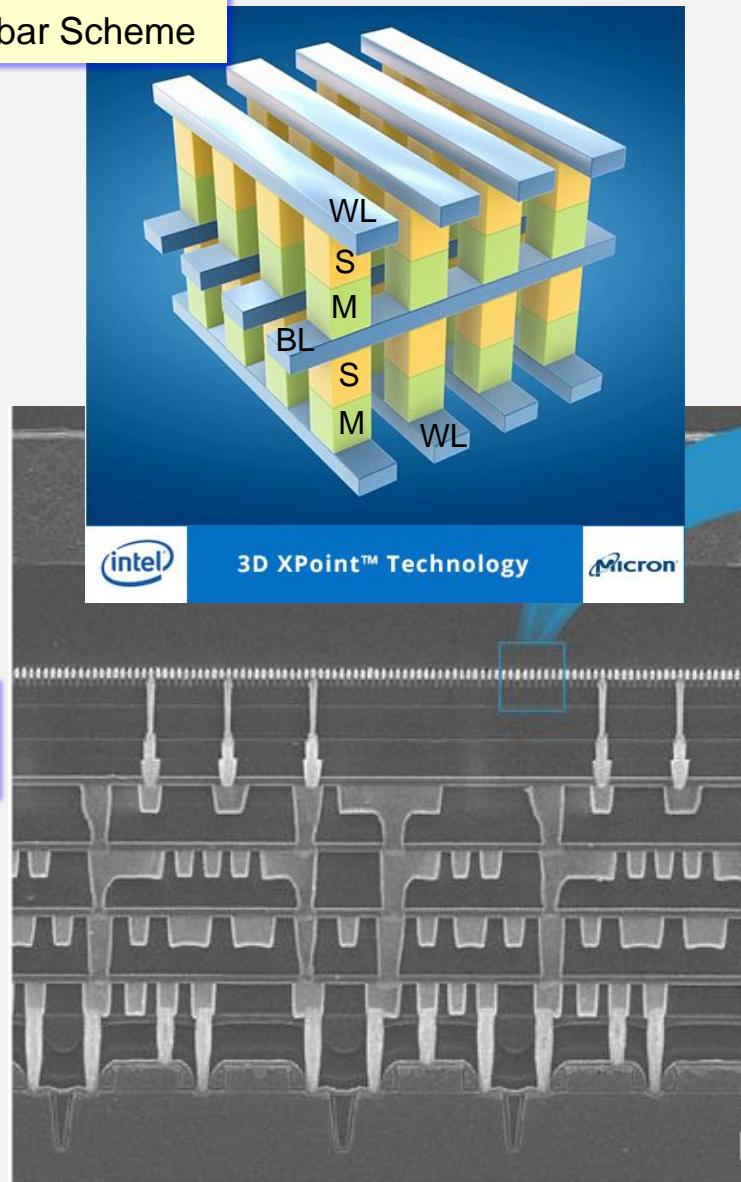


For each RRAM you need a different selector



OTS is used by Intel in their 3D-Xpoint Technology

INTEL Optane is a Phase Change Memory with Crossbar Scheme



Source:  
<http://semimd.com/blog/2017/10/18/xpoint-nvm-array-process-engineering/>

PCM should be a factor 100 faster than NAND  
(see page 6.156)

Memory density is lower than of 3D-NAND

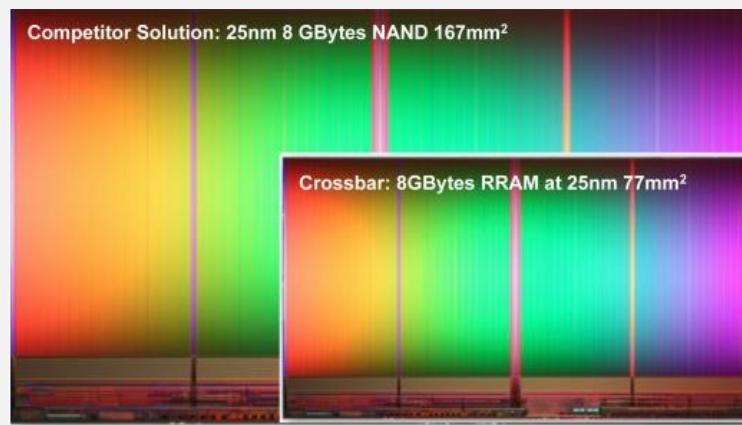
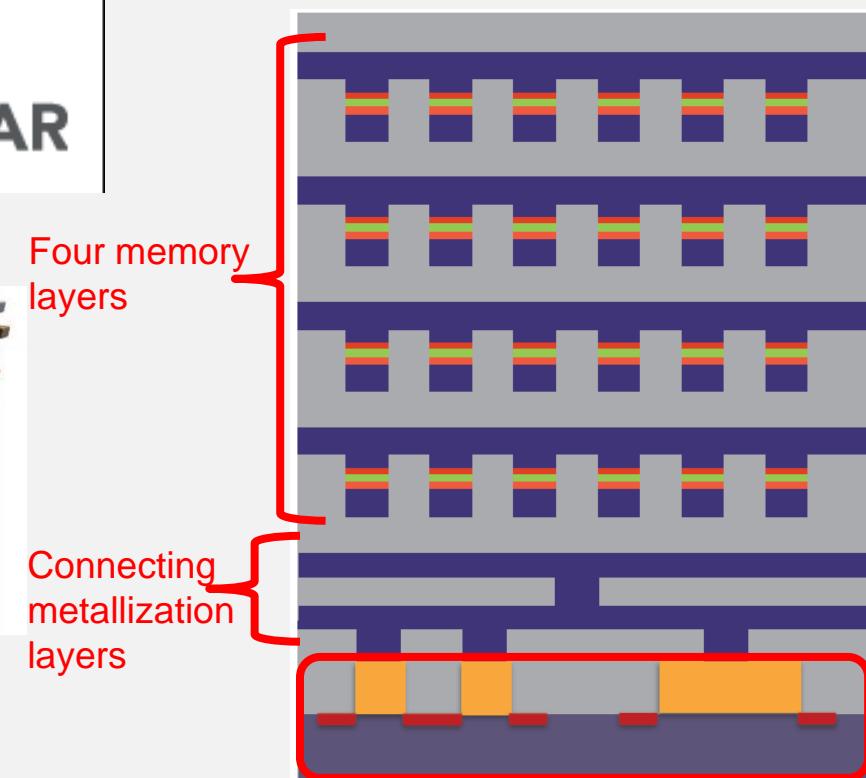
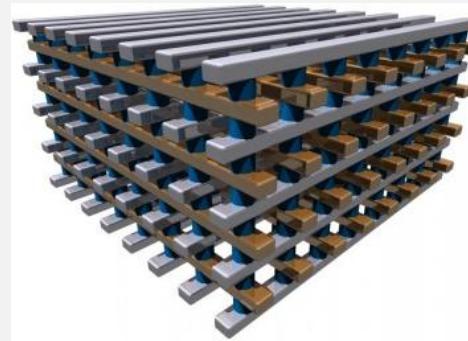
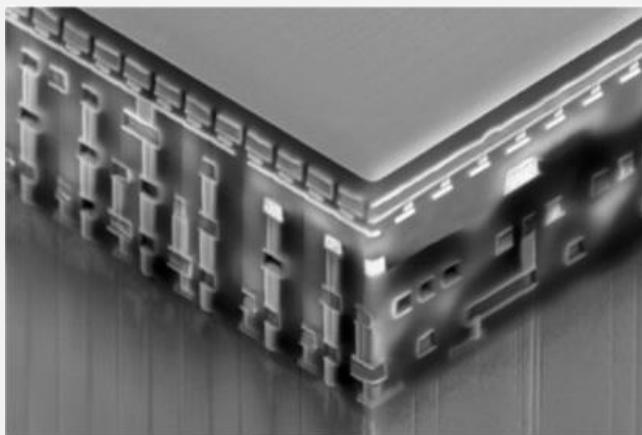
10 times more expensive than SSD with  
3D-NAND

Challenges: High current (1mA) for reset and  
self heating

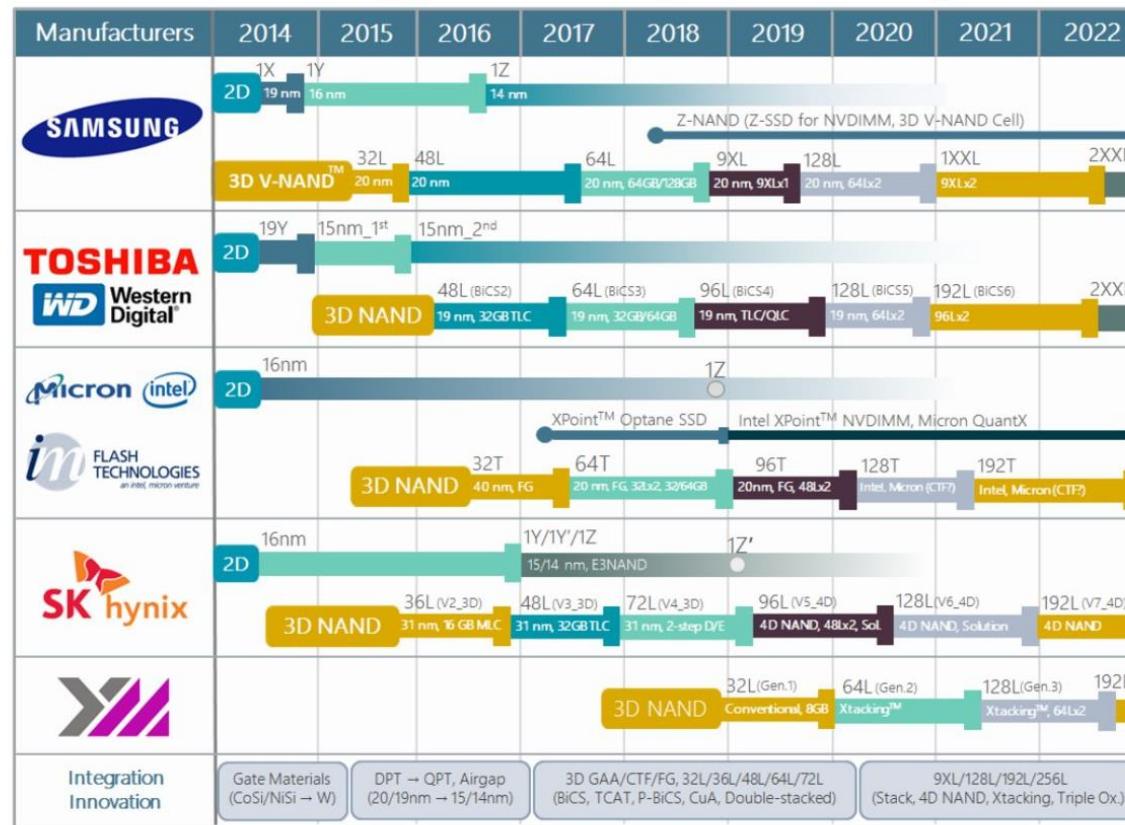
### Crossbar Resistive Memory: The Future Technology for NAND Flash

By Hagop Nazarian, Vice President of Engineering and Co-Founder

<https://www.crossbar-inc.com/technology/reram-overview/>



- ▶ CMOS compatible process and fast switching
- ▶ RRAM is expected to replace NAND at 10nm technology node
- ▶ PCM is using new materials for CMOS technology



Tech  
Insights

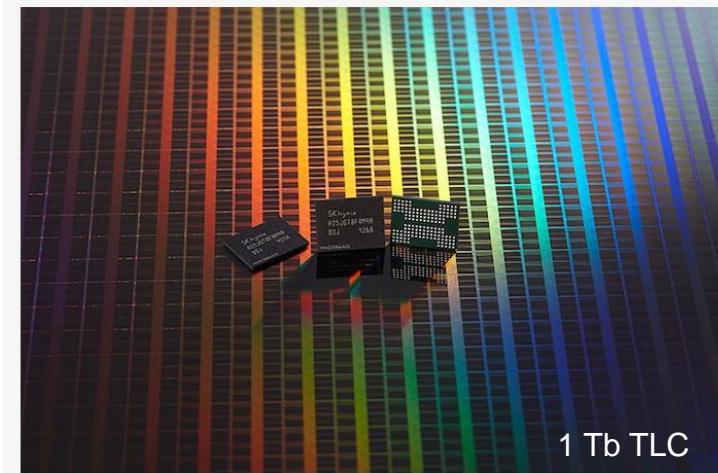
## SK Hynix Sampling 128-layer 3D NAND SSDs

by Billy Tallis & Anton Shilov on November 21, 2019 10:00 AM EST

Posted in [SSDs](#) | [Storage](#) | [3D NAND](#) | [SK Hynix](#) | [3D TLC](#) | [4D NAND](#) | [128-layer](#)

9  
Comments

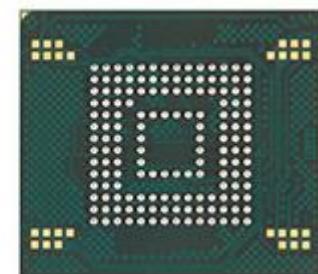
+ Add A Comment



1 Tb TLC

SK Hynix announced this week that they have started sampling products based on their 128-layer 3D NAND flash memory, which will soon start showing up in end-user devices. A year ago, they launched their 96-layer 5th generation 3D NAND, but low prices prompted them to cut output and their 4th generation 72L 3D NAND is still their main flash product. SK Hynix announced back in June that their 128L 3D NAND had moved from development to mass production, and now it has been incorporated into SSDs and UFS modules that are sampling to major customers.

<https://www.anandtech.com/show/15136/sk-hynix-sampling-128layer-3d-nand-ssds>



2D-Flash Memory is used in mobile applications

3D-Flash Memory, current in the polysilicon channel is decreasing

New channel materials instead of the polysilicon channel

Most promising candidate is the RRAM

**It looks like CMOS will  
work at least the next 15  
years**

## 6.1 Overview

Feature size, pitch, node

## 6.2 State of the Art MOSFETs

Electrical and geometrical parameters

## 6.3 Mobility Enhanced MOSFETs

Strain, SiGe, Liners, Directed Channels, Gate-Last

## 6.4 High-k Metal-Gate MOSFETs

Poly-Depletion, FuSi

## 6.5 Vertical MOSFETs

## 6.6 Halo MOSFETs

## 6.7 SOI-MOSFETs

## 6.8 Multi-Gate MOSFETs

Fully-depleted MOSFETs, FinFETs, Nanowire FETs

## 6.9 Advanced Memory Devices

End  
of  
Chapter 6