

# Du Nano Transistor aux Terabit CMOS



Thomas Skotnicki

ST Fellow  
Director Advanced  
Devices Program

STMicroelectronics  
Crolles, France

J3N, Grenoble

20 Oct. 2008

# OUTLINE

 **INTRODUCTION**

 **MANUFACTURABILITY PROBLEMS**

 **POWER / SPEED PROBLEMS**

 **LOOKING FOR SOLUTIONS – MATERIALS**

 **LOOKING FOR SOLUTIONS – DEVICES**

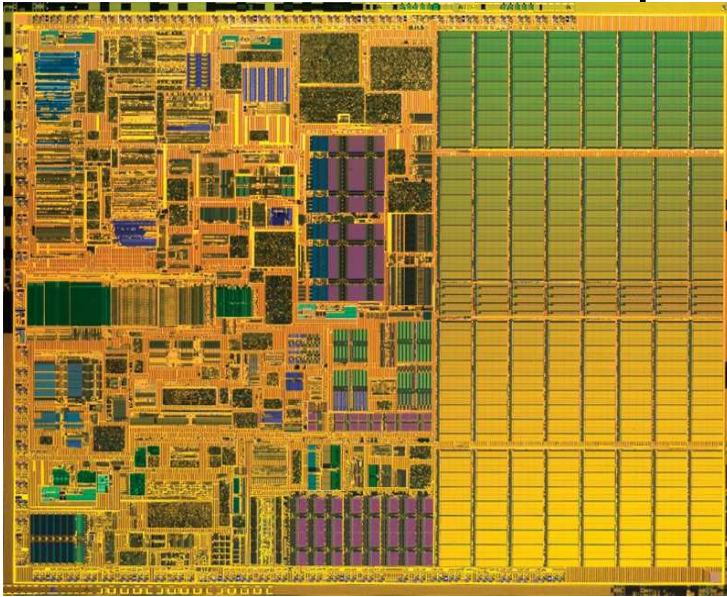
 **LOOKING FOR SOLUTIONS - EMERGING**

 **NEW PARADIGM FOR 22nm and BEYOND**

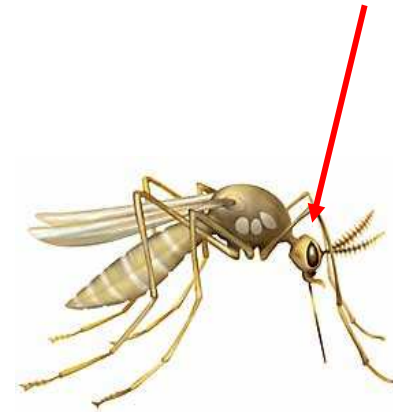
 **LONGER TERM**

# SILICON IS NOT SO BAD....AND THERE IS STILL ROOM FOR IMPROVEMENT

Pentium:  $\sim 100\text{mm}^2 \times 5\mu\text{m}$  active layer =  $0.5\text{mm}^3$



Mosquito's head  $\sim 0.5\text{mm}^3$  ?



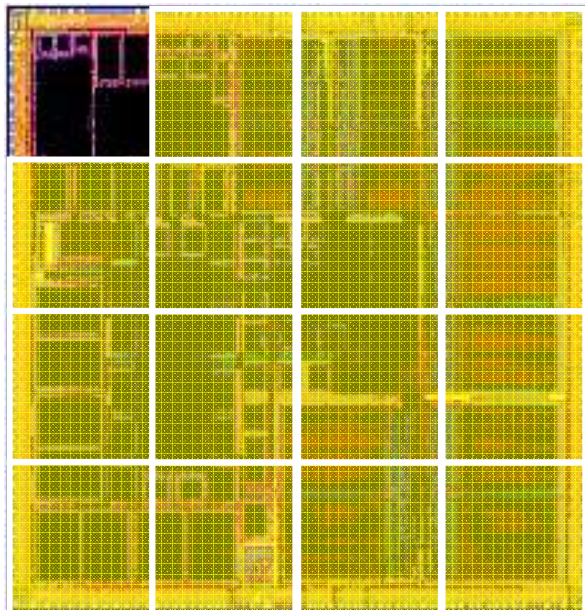
## COMPUTERS ARE CAPABLE OF:

- Winning chess with Gary Kasparov (**FLOPS** :Deeper Blue > 1000xPentium)
- Calculating trajectory to Mars
- Making my slides

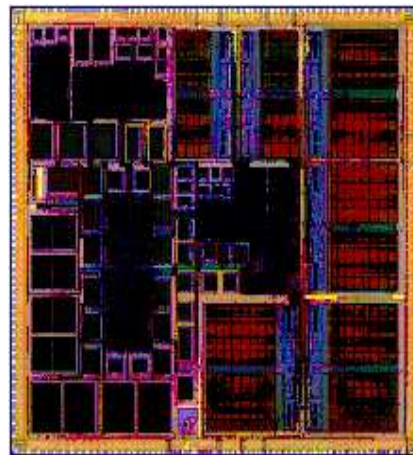
# SILICON IS NOT SO BAD, BUT PACKAGING AND SYSTEM INTEGRATION HAVE TO BE IMPROVED



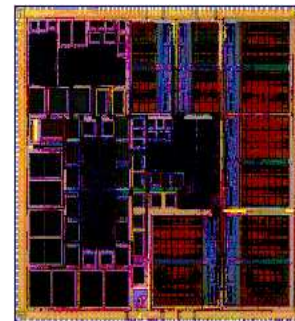
# CMOS Scaling IS NOT ONLY FOR PERFORMANCE...



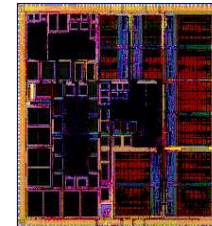
CMOS120



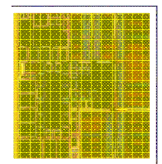
CMOS090



CMOS065



CMOS045



CMOS032



# OUTLINE

 **INTRODUCTION**

 **MANUFACTURABILITY PROBLEMS**

 **POWER / SPEED PROBLEMS**

 **LOOKING FOR SOLUTIONS – MATERIALS**

 **LOOKING FOR SOLUTIONS – DEVICES**

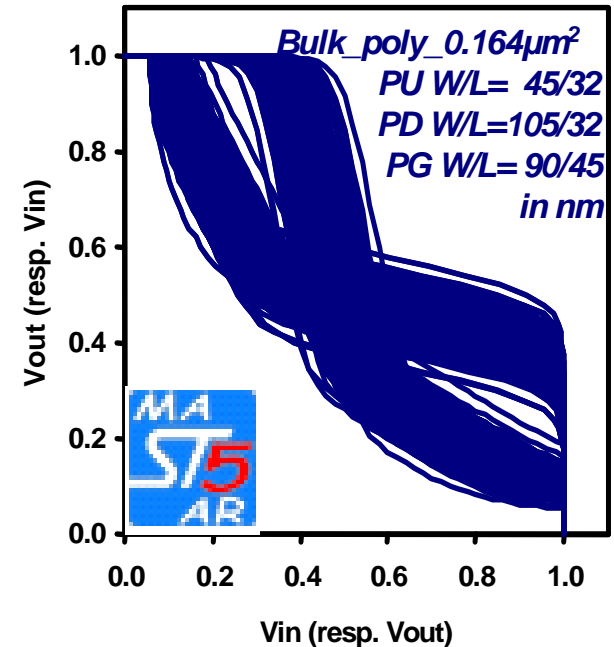
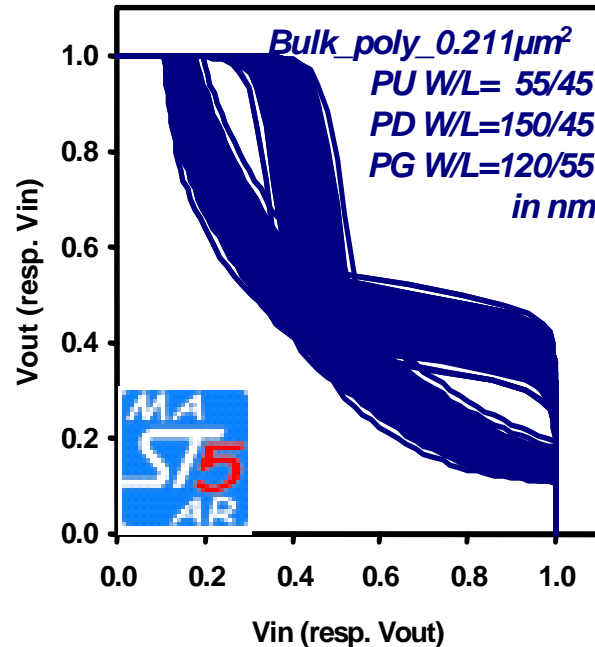
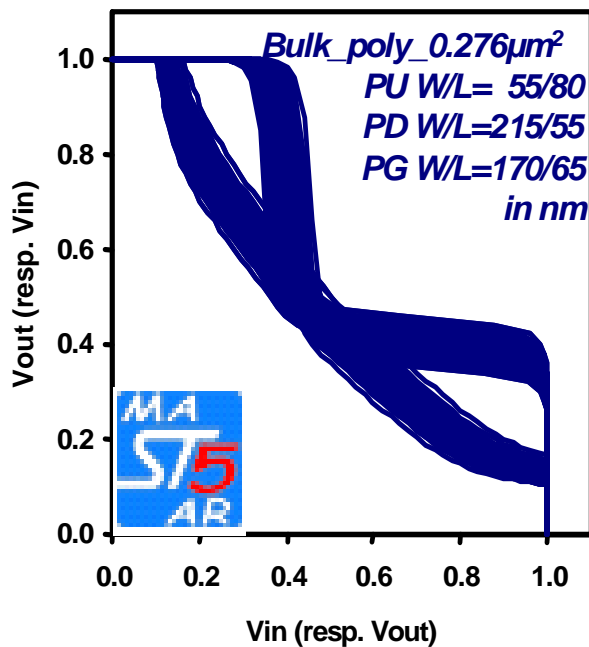
 **LOOKING FOR SOLUTIONS – EMERGING**

 **NEW PARADIGM FOR 32nm and BEYOND**

 **LONGER TERM**

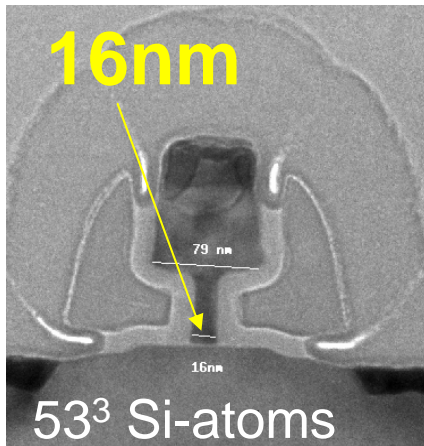
# MANUFACTURABILITY :

## SMALL IS DIFFICULT

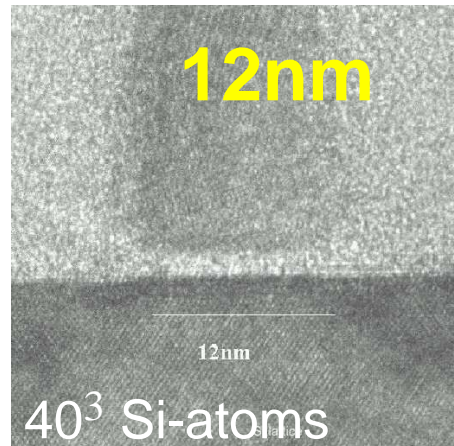


# TWO SOURCES OF FLUCTUATIONS:

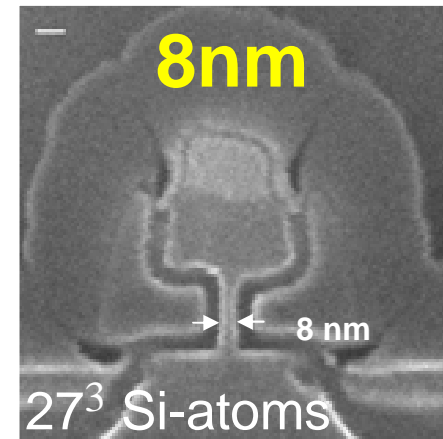
## 1) DISCRETENESS OF MATTER (CROLLES EXPERIMENTAL MOSFET SAMPLES):



&  $3.5^3$  B-dopants  
( $1e19cm^{-3}$ ) under gate



&  $2.6^3$  B-dopants  
( $1e19cm^{-3}$ ) under gate



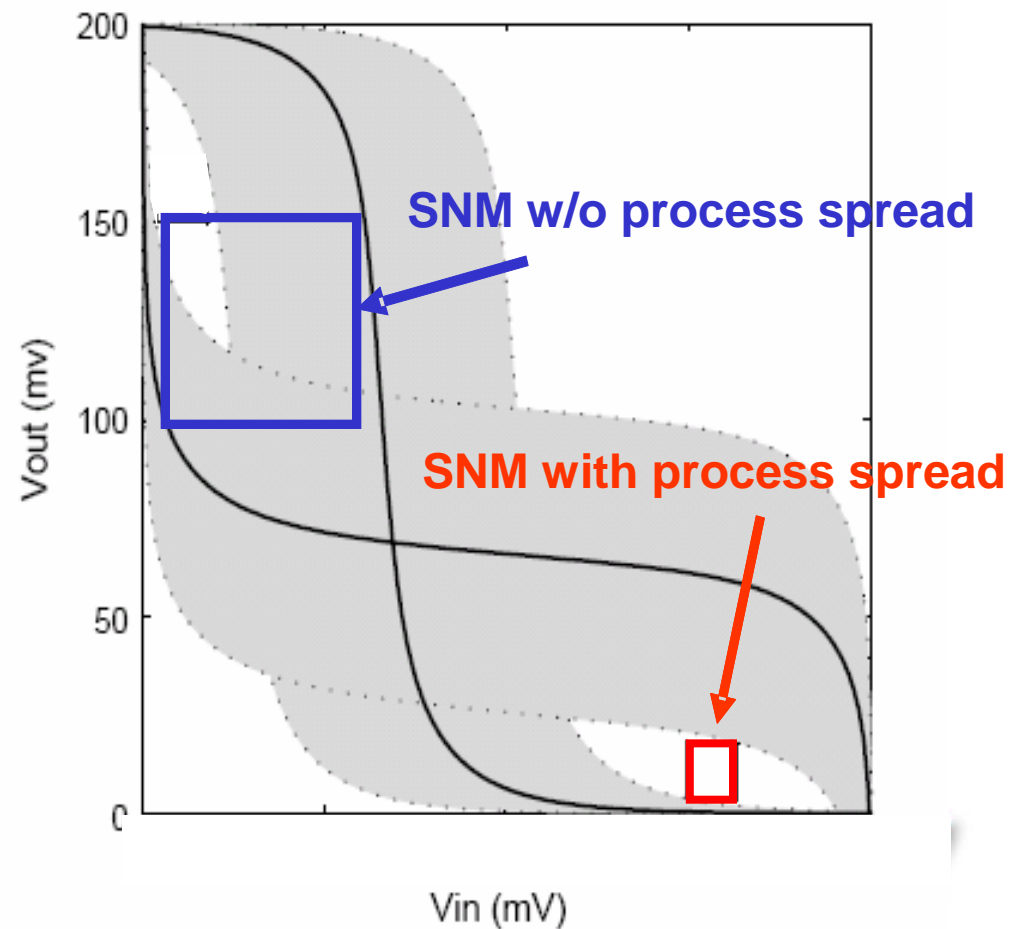
&  $1.7^3$  B-dopants  
( $1e19cm^{-3}$ ) under gate

Active volume under gate is supposed for simplicity to be equal  $Lg^3$

# FLUCTUATIONS – IMPACT ON SRAM

$$\delta V_{th} = \frac{\sqrt[4]{2\varepsilon_s \varepsilon_0 q^3 N_c \Phi_d}}{\varepsilon_{ox} \varepsilon_0} T_{ox} \frac{1}{\sqrt{LW}} \equiv A_{vt} \frac{1}{\sqrt{LW}}$$

Ref.: Microelectronics Journal 36  
(2005) 789–800  
« Standby supply voltage minimization for deep  
sub-micron SRAM »,  
Huifang Qin\*, Yu Cao, Dejan Markovic, Andrei  
Vladimirescu, Jan Rabaey



# OUTLINE

 **INTRODUCTION**

 **MANUFACTURABILITY PROBLEMS**

 **POWER / SPEED PROBLEMS**

 **LOOKING FOR SOLUTIONS – MATERIALS**

 **LOOKING FOR SOLUTIONS – DEVICES**

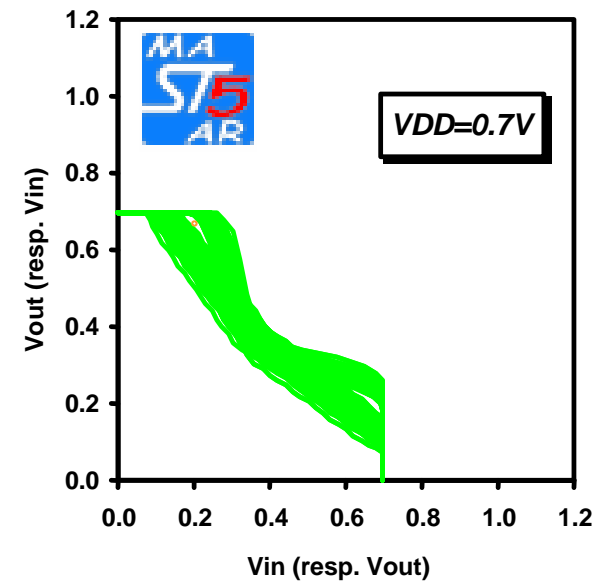
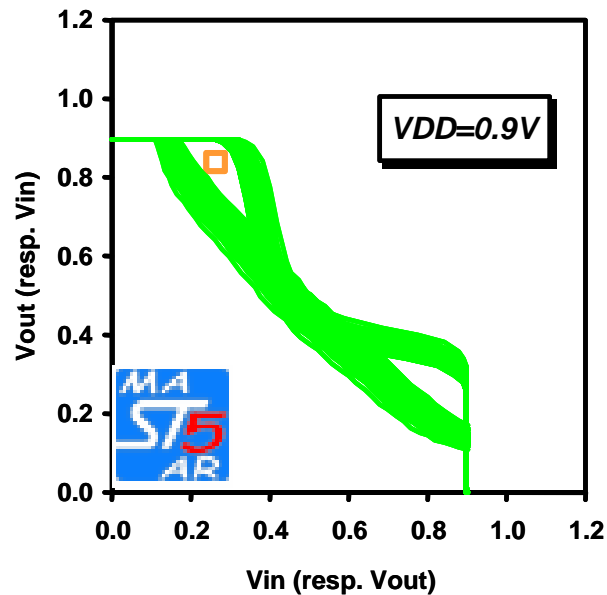
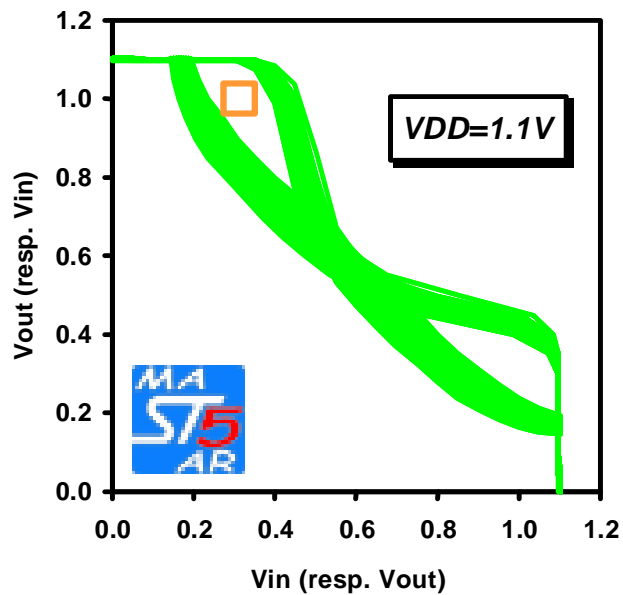
 **LOOKING FOR SOLUTIONS – EMERGING**

 **NEW PARADIGM FOR 32nm and BEYOND**

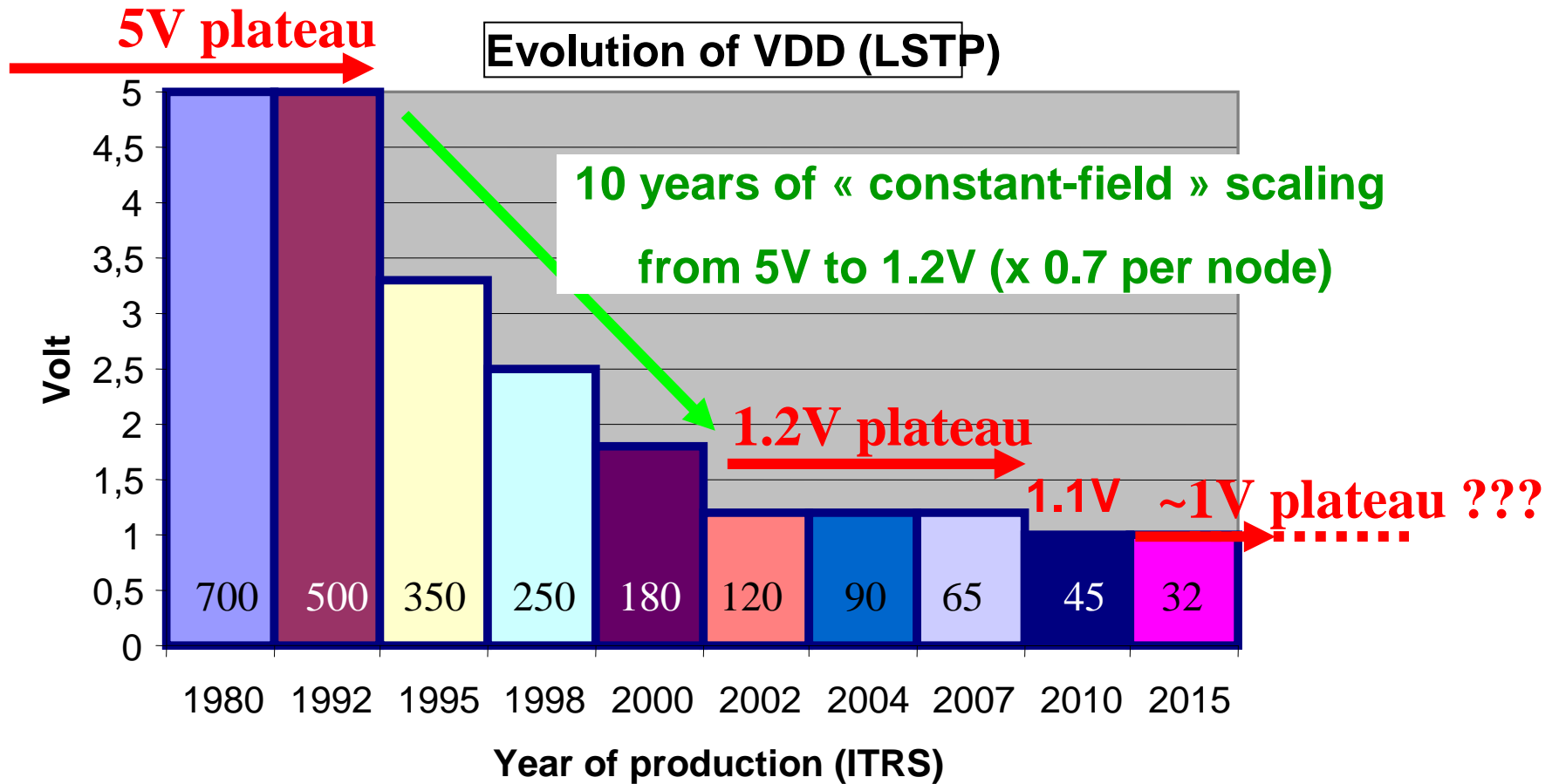
 **LONGER TERM**

# POWER

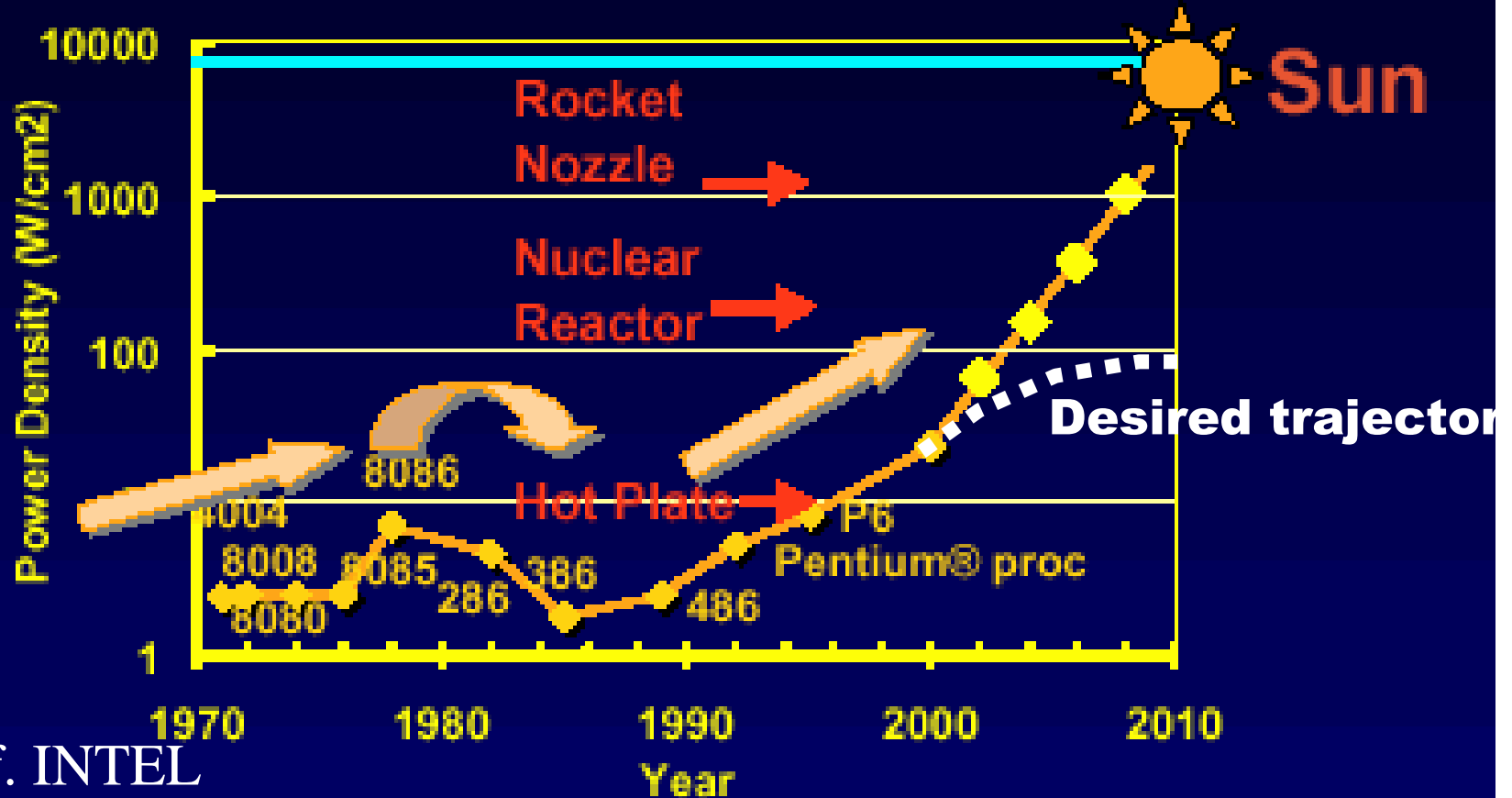
## LP IS DIFFICULT



# FLUCTUATIONS ARE BEHIND THE VDD CLUMSY SCALING AND THUS BEHIND THE « POWER CRISIS »



# Power density will increase



Power density too high to keep junctions at low temp

# OUTLINE

 **INTRODUCTION**

 **MANUFACTURABILITY PROBLEMS**

 **POWER / SPEED PROBLEMS**

 **LOOKING FOR SOLUTIONS – MATERIALS**

 **LOOKING FOR SOLUTIONS – DEVICES**

 **LOOKING FOR SOLUTIONS – EMERGING**

 **NEW PARADIGM FOR 32nm and BEYOND**

 **LONGER TERM**

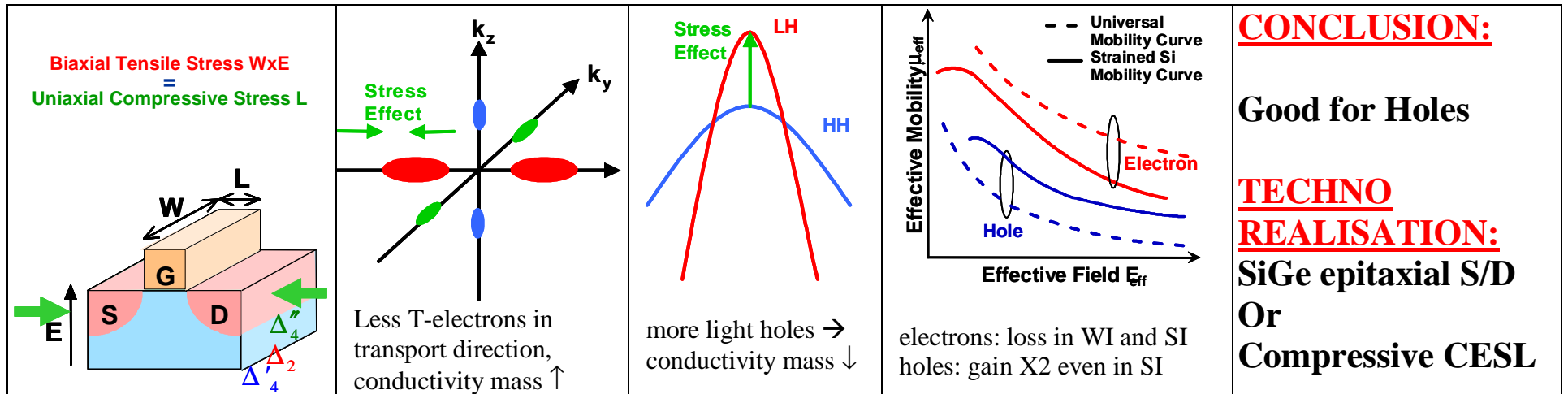
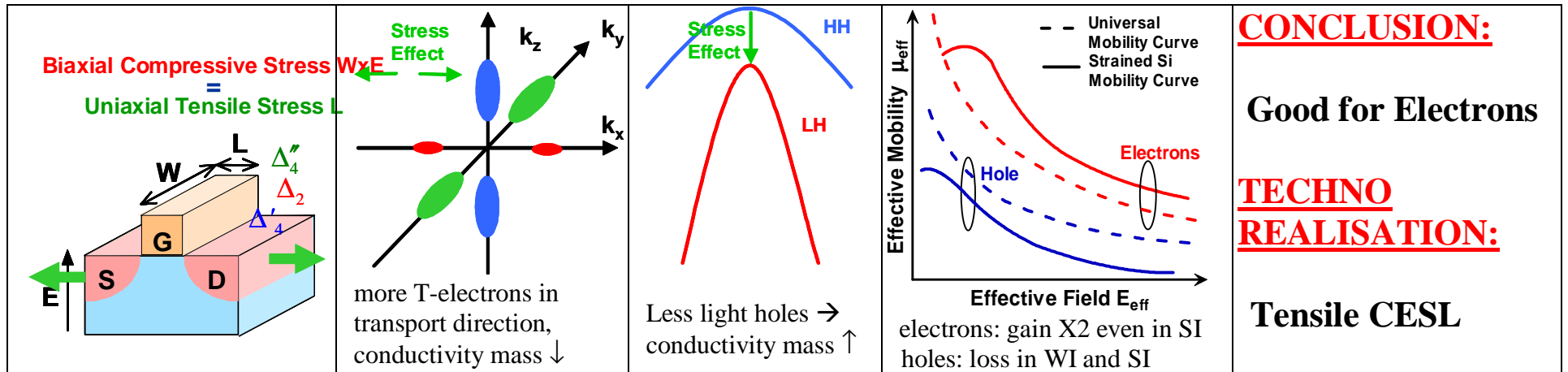
# Materials, Materials, Materials:

IA	IIA	IIIB	IVB	VB	VIB	VII B	VIII	Ib	IIb	IIIA	IVA	VA	VIA	VIIA	VIIIA		
H															He		
Li	Be									B	C	N	O	F	Ne		
Na	Mg									Al	Si	P	S	Cl	Ar		
K	Ca	Sc	Ti	V	Cr	Mn	Fe	Co	Ni	Cu	Zn	Ga	Ge	As	Se	Br	Kr
Rb	Sr	Y	Zr														
Cs	Ba	La	Hf														
Fr	Ra	Ac															

IA	IIA	IIIB	IVB	VB	VIB	VII B	VIII	Ib	IIb	IIIA	IVA	VA	VIA	VIIA	VIIIA		
H															He		
Li	Be									B	C	N	O	F	Ne		
Na	Mg																
K	Ca	Sc	Ti	V	Cr	Mn	Fe	Co	Ni	Cu	Zn	Ga	Ge	As	Se	Br	Kr
Rb	Sr	Y	Zr	Nb	Mo	Tc	Ru	Rh	Pd	Ag	Cd	In	Sn	Sb	Te	I	Xe
Cs	Ba	La	Hf	Ta	W	Re	Os	Ir	Pt	Au	Hg	Tl	Pb	Bi	Po	At	Rn
Fr	Ra	Ac															
			Ce	Pr	Nd	Pm	Sm	Eu	Gd	Tb	Dy	Ho	Er	Tm	Yb	Lu	
			Th	Pa	U	Np	Pu	Am	Cm	Bk	Cf	Es	Fm	Md	No	Lr	

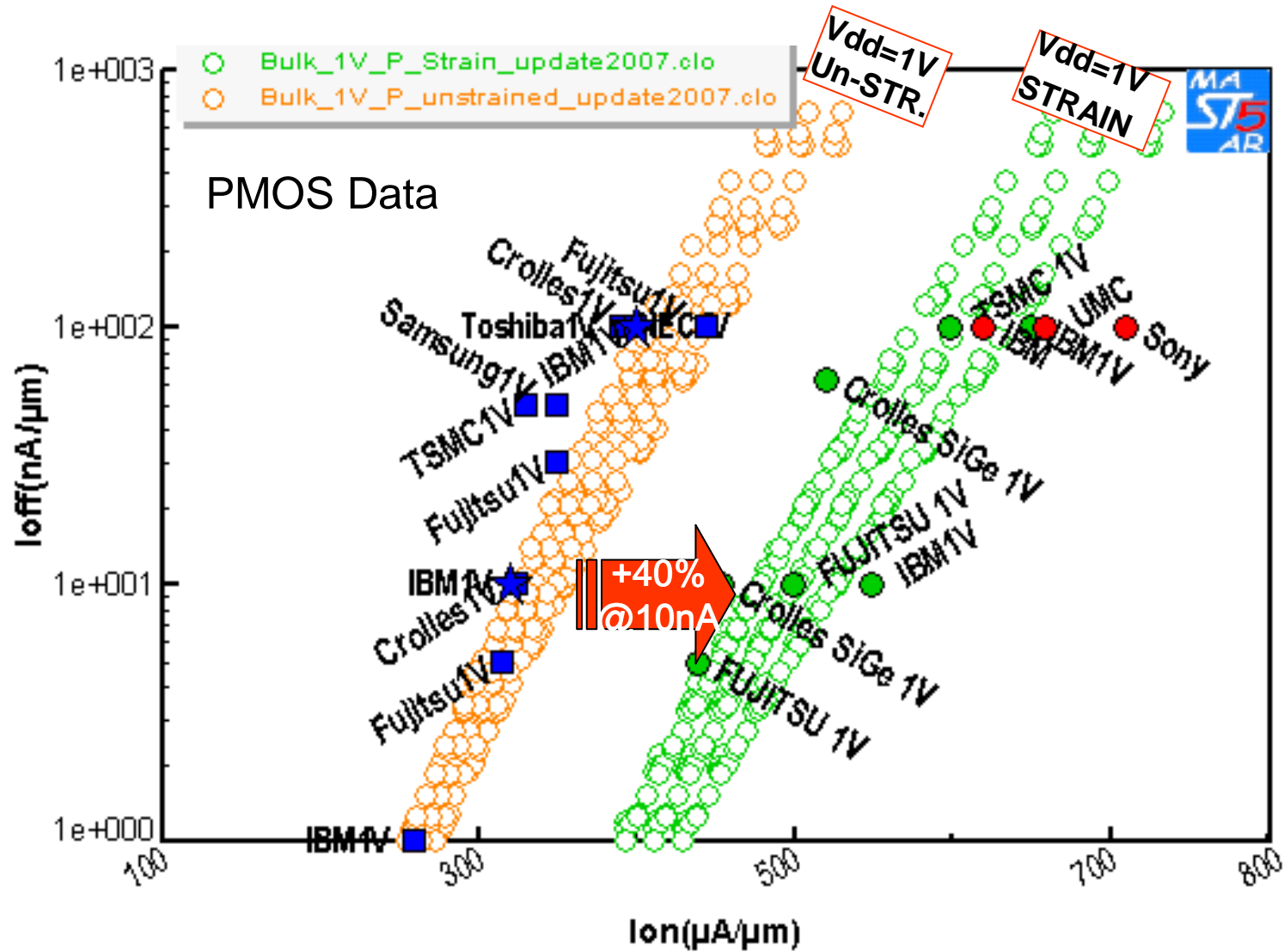
IA	IIA	IIIB	IVB	VB	VIB	VII B	VIII	Ib	IIb	IIIA	IVA	VA	VIA	VIIA	VIIIA		
H															He		
Li	Be									B	C	N	O	F	Ne		
Na	Mg									Al	Si	P	S	Cl	Ar		
K	Ca	Sc	Ti	V	Cr	Mn	Fe	Co	Ni	Cu	Zn	Ga	Ge	As	Se	Br	Kr
Rb	Sr	Y	Zr	Nb	Mo	Tc	Ru	Rh	Pd	Ag	Cd	In	Sn	Sb	Te	I	Xe
Cs	Ba	La	Hf	Ta	W	Re	Os	Ir	Pt	Au	Hg	Tl	Pb	Bi	Po	At	Rn
Fr	Ra	Ac															
			Ce	Pr	Nd	Pm	Sm	Eu	Gd	Tb	Dy	Ho	Er	Tm	Yb	Lu	
			Th	Pa	U	Np	Pu	Am	Cm	Bk	Cf	Es	Fm	Md	No	Lr	

# STRAINED SILICON





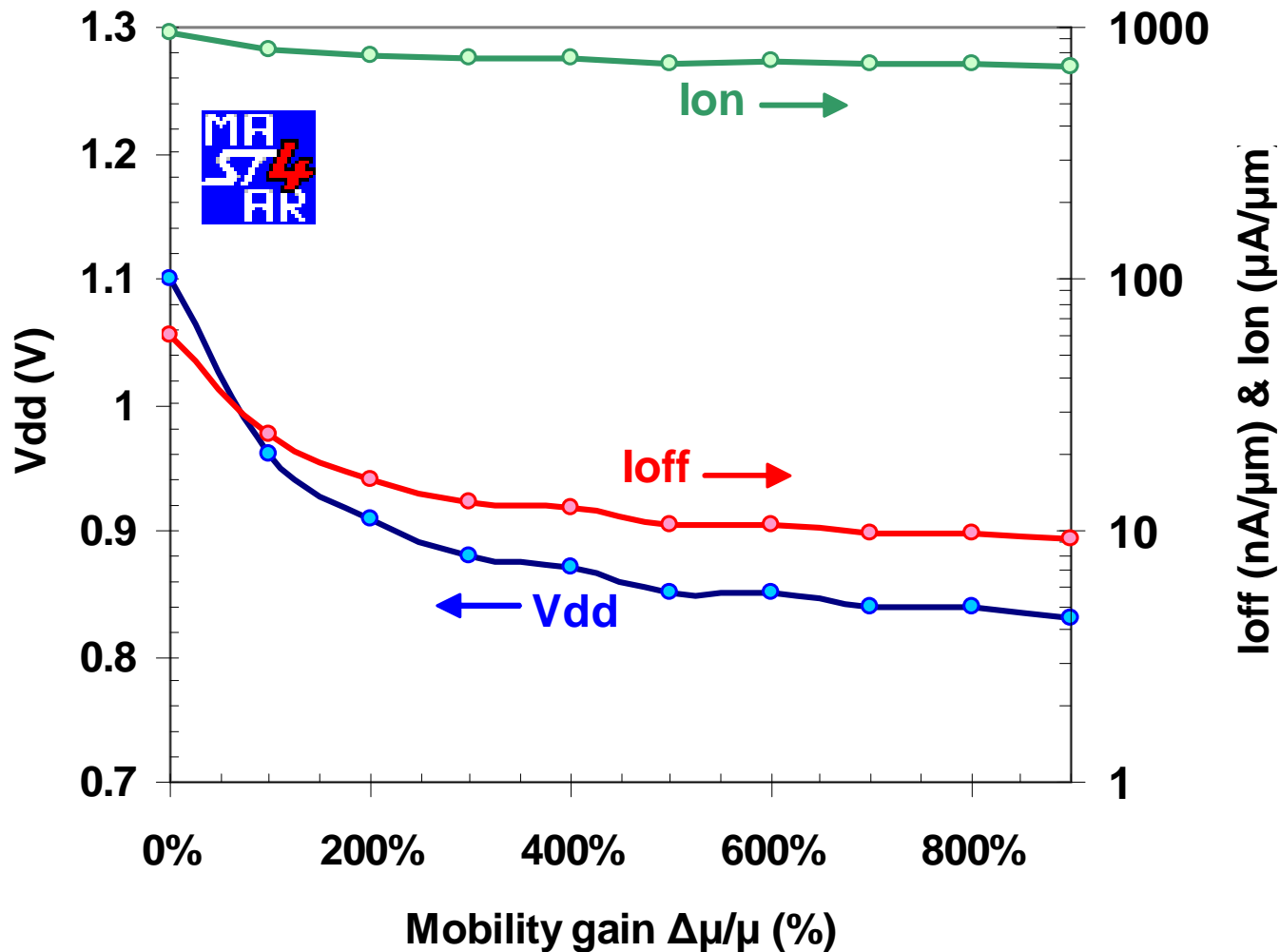
# PRESENT USE OF MOBILITY BOOSTERS



# High $\mu \sim$ Small $E_g \sim$ High $\epsilon$

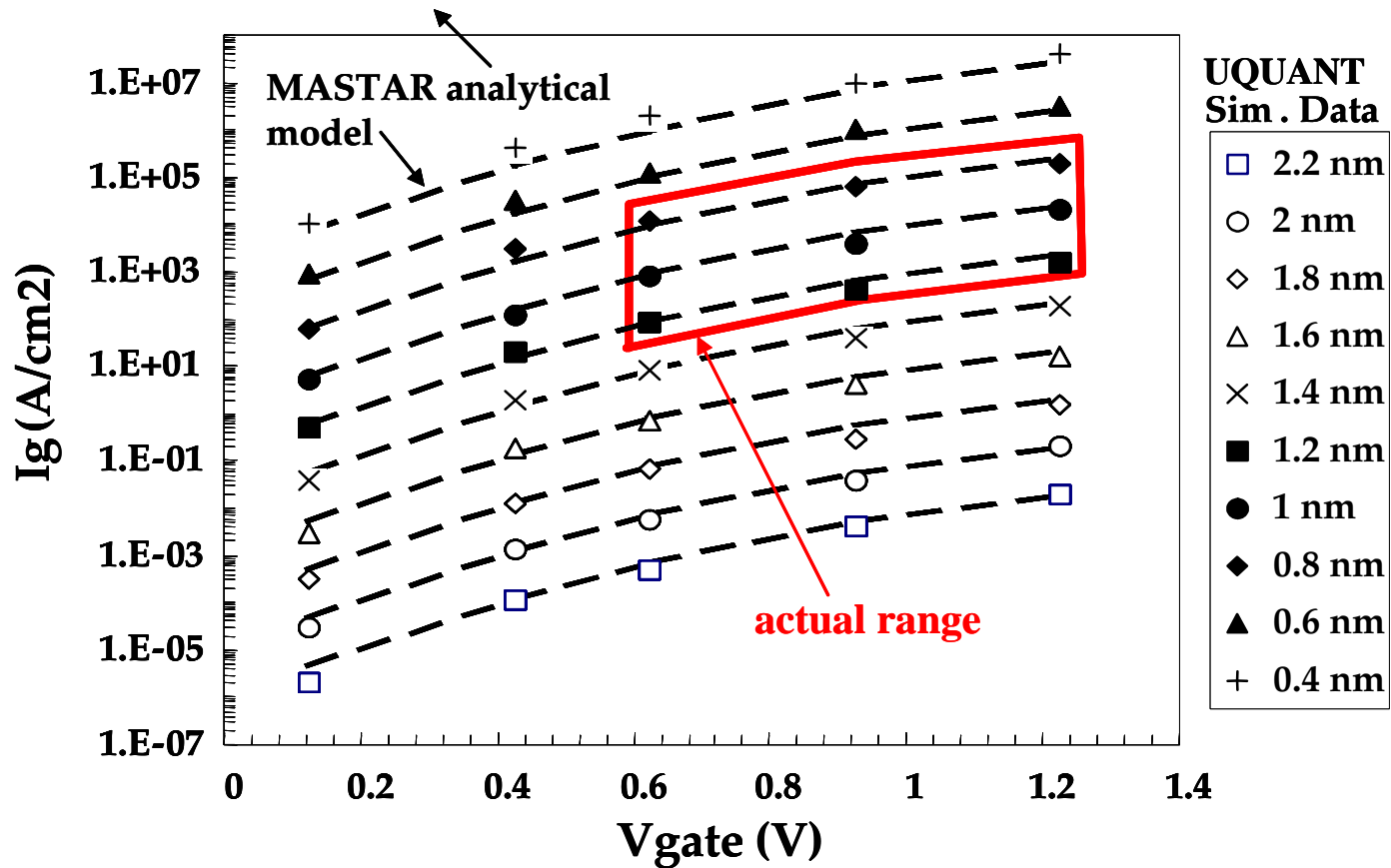
Semiconductor	Type	Bandgap (eV) at 300K	Dielectric constant	Electron bulk mobility (cm <sup>2</sup> /Vs)	Hole bulk mobility (cm <sup>2</sup> /Vs)
InSb	Direct	0.17	15.9	77000	850
InAs	Direct	0.36	12.	30000	450
GaSb	Direct	0.68	14.8	5000	1000
InP	Direct	1.27	12.1	4500	100
GaAs	Direct	1.43	11.5	8000	300
GaN	Direct	3.39	$\epsilon_{\perp} = 9.5$	1000	50
Ge	Indirect	0.66	16	3600	1800
Si	Indirect	1.12	12	1350	480

# FUTURE ? - MOBILITY BOOST TRADED AGAINST VDD AT CONSTANT CV/I (if fluctuations suppressed)

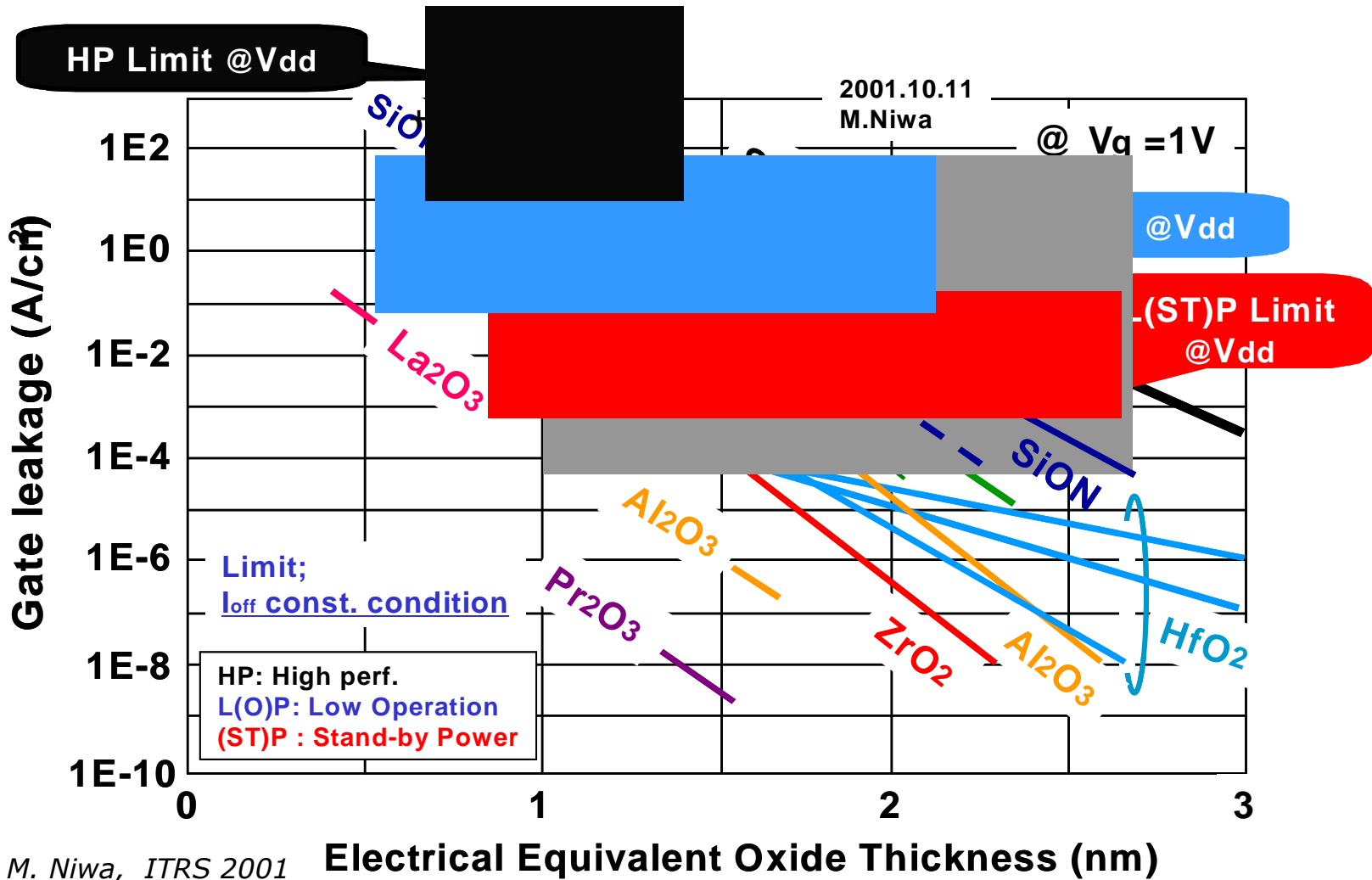


# SiO2 : TOO LEAKY GATE

$$I_g[\text{A/cm}^2] = 1.44e5 * (\text{Exp}(-4.02 * U_g[\text{V}]^2 + 13.05 * U_g[\text{V}]) * \text{Exp}(-1.17 * T_{ox}[\text{\AA}]))$$



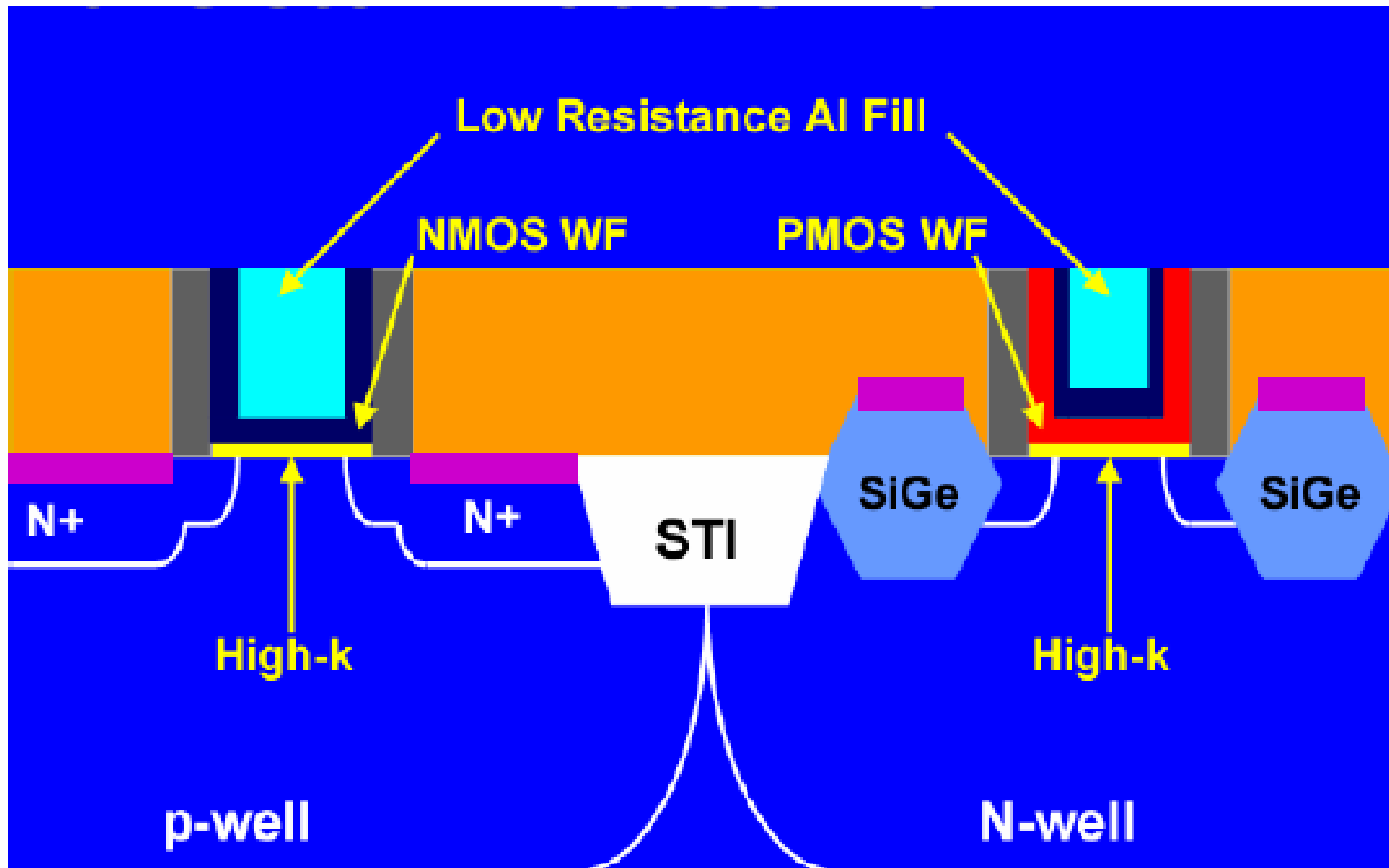
# TODAY Hf BASED OXIDES GO TO PRODUCTION



Source : M. Niwa, ITRS 2001

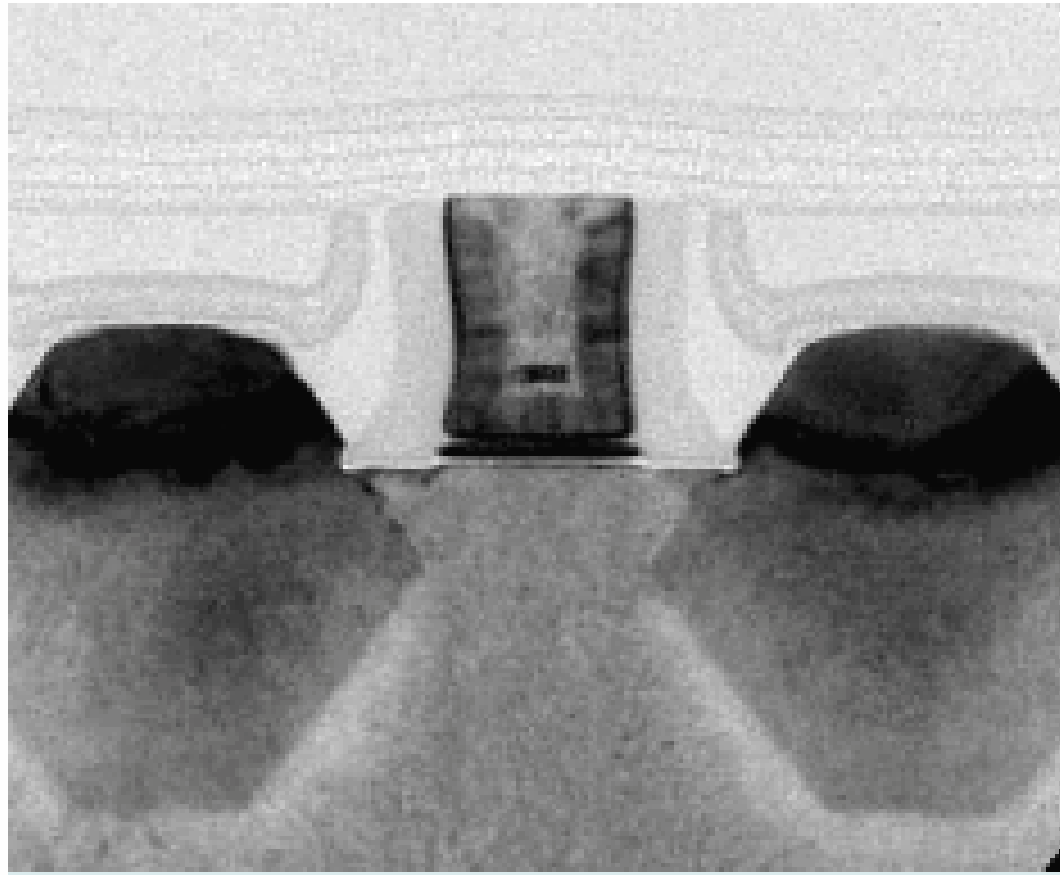
Electrical Equivalent Oxide Thickness (nm)

# INTEL HK-MG : the first IN-PRODUCTION solution



From Hermes Microvision

# INTEL's PMOSFET with HK/MG and SiGe S/D



*From INTEL's IEDM 2007 presentation 10.2*

# OUTLINE

 **INTRODUCTION**

 **MANUFACTURABILITY PROBLEMS**

 **POWER / SPEED PROBLEMS**

 **LOOKING FOR SOLUTIONS – MATERIALS**

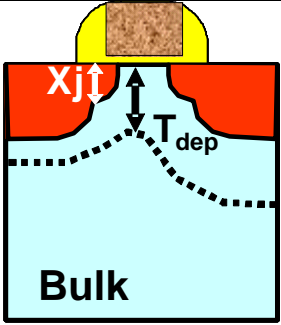
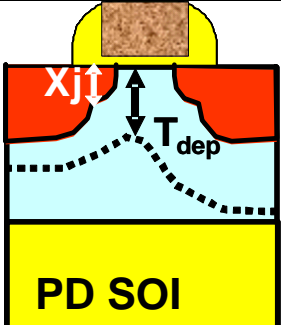
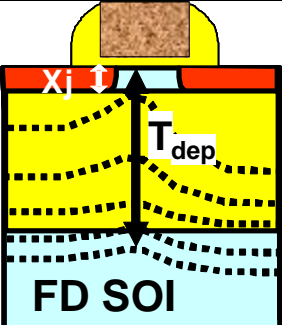
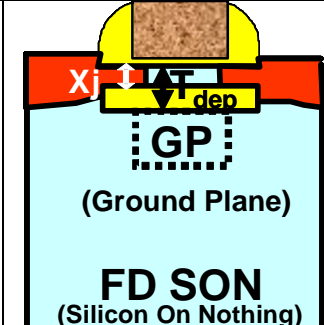
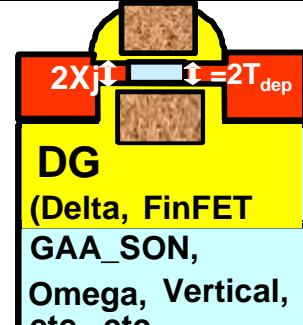
 **LOOKING FOR SOLUTIONS – DEVICES**

 **LOOKING FOR SOLUTIONS – EMERGING**

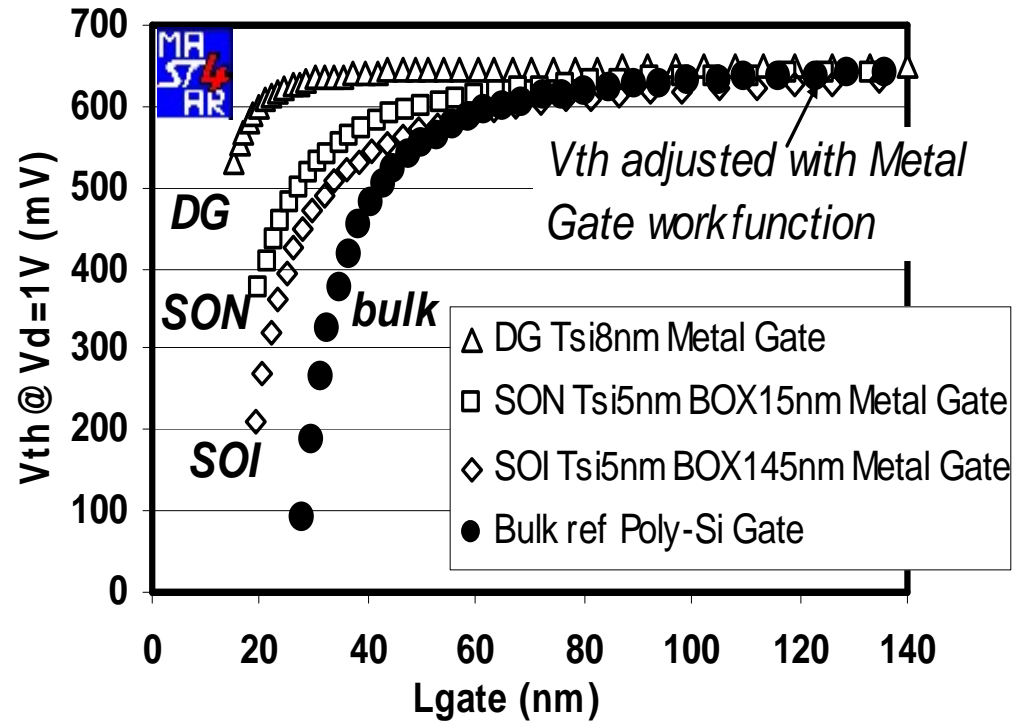
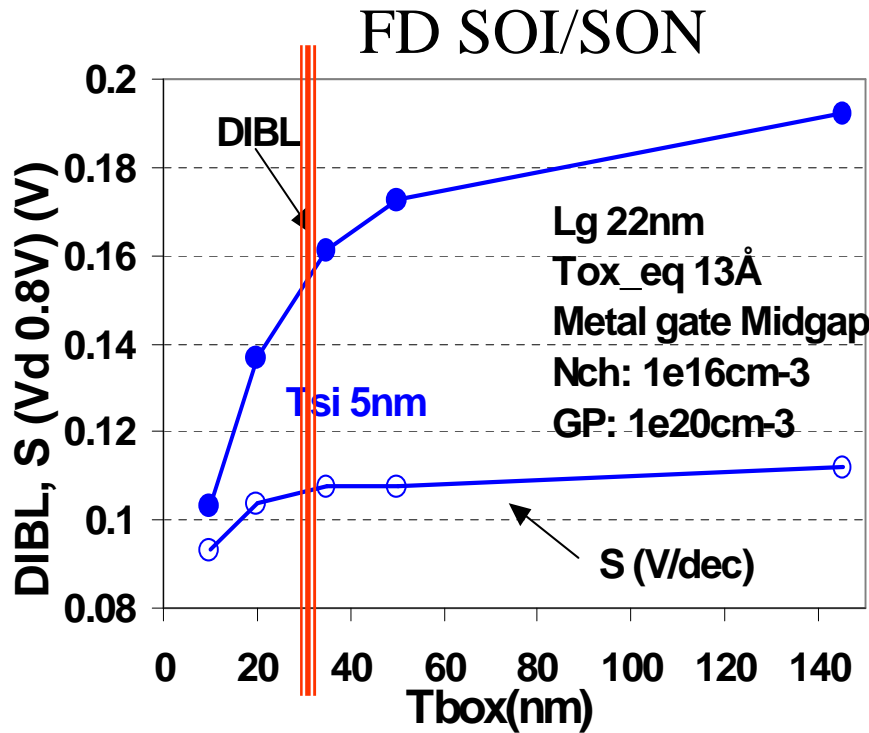
 **NEW PARADIGM FOR 32nm and BEYOND**

 **LONGER TERM**

# DEVICE STRUCTURES

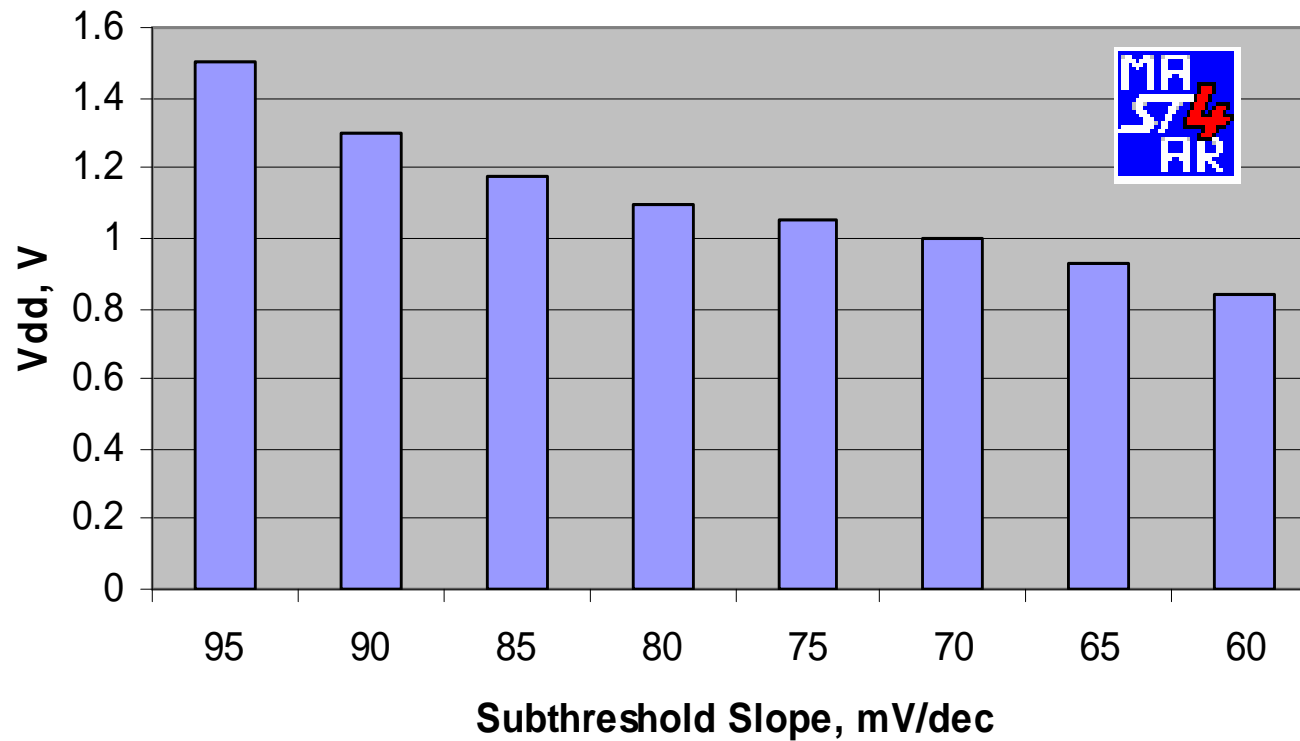
 <p><b>Bulk</b></p>	 <p><b>PD SOI</b></p>	 <p><b>FD SOI</b></p>	 <p><b>FD SON</b> (Silicon On Nothing)</p>	 <p><b>DG</b> (Delta, FinFET, GAA_SON, Omega, Vertical, etc., etc.)</p>
$SCE = 0.64 \frac{\epsilon_{si}}{\epsilon_{ox}} \times EI \times \Phi_d$		$DIBL = 0.8 \frac{\epsilon_{si}}{\epsilon_{ox}} \times EI \times V_{ds}$		
$EI = 1 \times \left( 1 + \frac{X_j^2}{L_{el}^2} \right) \times \frac{T_{ox}}{L_{el}} \frac{T_{dep}}{L_{el}}$	$EI = 1 \times \left( 1 + \frac{X_j^2}{L_{el}^2} \right) \times \frac{T_{ox}}{L_{el}} \frac{T_{dep}}{L_{el}}$	$EI = 1 \times \left( 1 + \frac{T_{si}^2}{L_{el}^2} \right) \times \frac{T_{ox}}{L_{el}} \frac{T_{si} + \lambda T_{box}}{L_{el}}$	$EI = 1 \times \left( 1 + \frac{T_{si}^2}{L_{el}^2} \right) \times \frac{T_{ox}}{L_{el}} \frac{T_{si}}{L_{el}}$	$EI = 1 \times \left( 1 + \frac{T_{si}^2/4}{L_{el}^2} \right) \times \frac{T_{ox}}{L_{el}} \frac{T_{si}/2}{L_{el}}$
$L_{el, \min}(DIBL=100mV/V) = 2.82T_{dep} = 34nm$ <p>if max ch. doping is <math>10^{19}cm^{-3} \Rightarrow T_{ep} \sim 12nm</math></p>	$L_{el, \min}(DIBL=100mV/V) = 2.82T_{dep} = 34nm$ <p>if max ch. doping is <math>10^{19}cm^{-3} \Rightarrow T_{ep} \sim 12nm</math></p>	$L_{el, \min}(DIBL=100mV/V) \gg 1.82T_{si} \gg 5.5nm$ <p>if minimum feasible <math>T_{si}</math> is supposed 3nm (exact numbers depend on <math>\lambda T_{box}</math>)</p>	$L_{el, \min}(DIBL=100mV/V) = 1.82T_{si} \approx 5.5nm$ <p>if minimum feasible <math>T_{si}</math> is supposed 3nm</p>	$L_{el, \min}(DIBL=100mV/V) = 0.91T_{si} \approx 2.7nm$ <p>if minimum feasible <math>T_{si}</math> is supposed 3nm</p>

# THIN BOX – How thin for SCALABILITY

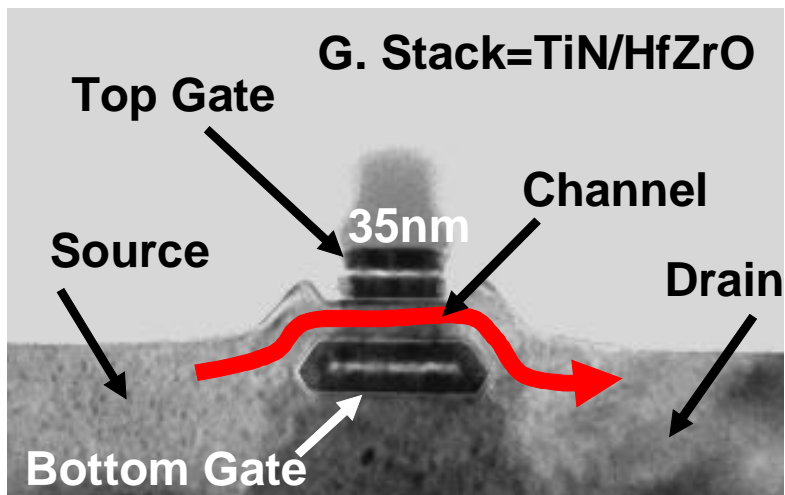
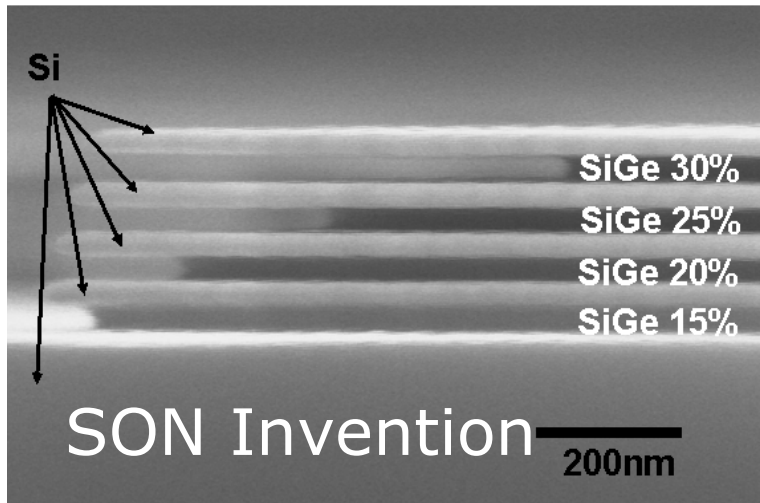


# SLOPE is key to decrease Vdd, AND THUS POWER

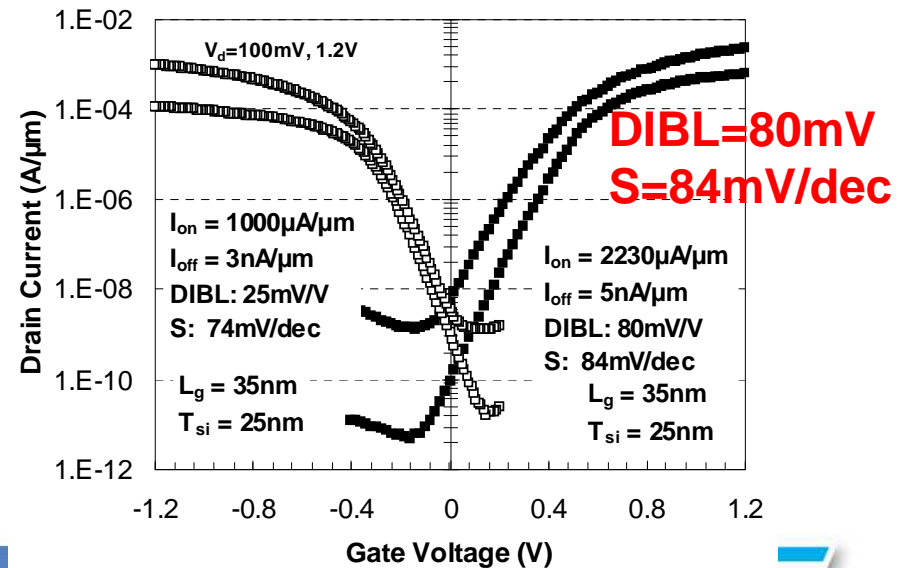
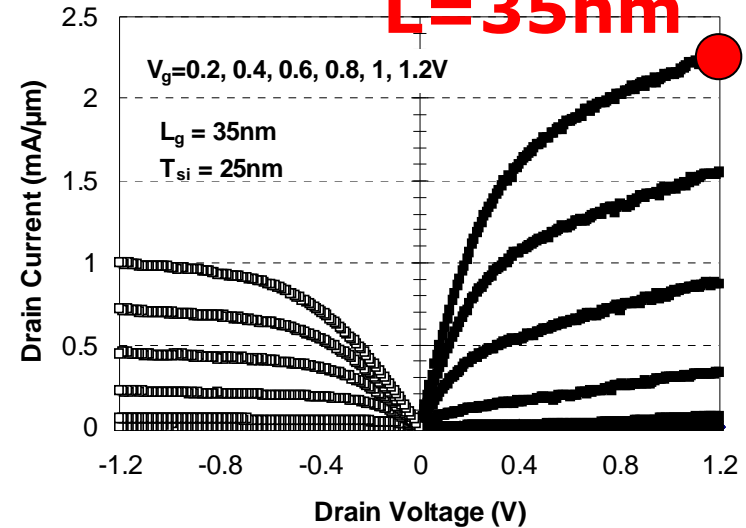
Vdd-vs-SS @ CV/I=const (HP22nm)



# BEYOND 32nm



**2300  $\mu\text{A}/\mu\text{m}$**   
**L=35nm**



# OUTLINE

 **INTRODUCTION**

 **MANUFACTURABILITY PROBLEMS**

 **POWER / SPEED PROBLEMS**

 **LOOKING FOR SOLUTIONS – MATERIALS**

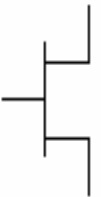
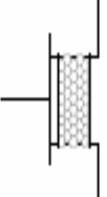
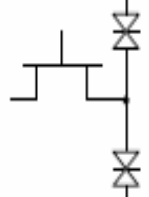
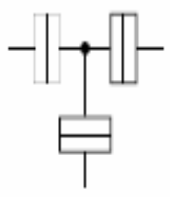



 **LOOKING FOR SOLUTIONS – DEVICES**

 **LOOKING FOR SOLUTIONS – EMERGING**

 **NEW PARADIGM FOR 32nm and BEYOND**

 **LONGER TERM**

## SIMILAR TO CMOS LIMITATIONS (ITRS 2005):

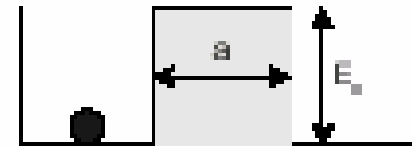
Device								
		FET [B]	1D structures	Resonant Tunneling Devices	SET	Molecular	Ferromagnetic logic	Spin transistor
Cell Size (spatial pitch)	Projected	100 nm	100 nm [C]	100 nm [C]	40 nm [L]	10 nm [Q]	140 nm [U]	100 nm [C]
	Demonstrated	590 nm	~1.5 μm [D]	3 μm [H]	~700 nm [M]	~2 μm [R]	250 nm [V, W]	100 μm [X]
Density (device/cm <sup>2</sup> )	Projected	1E10	4.5E9	4.5E9	6E10	1E12	5E9	4.5E9
	Demonstrated	2.8E8	4E7	1E7	2E8	2E7	1.6E9	1E4
Switch Speed	Projected	12 THz	6.3 THz [E]	16 THz [I]	10 THz [M]	1 THz [S]	1 GHz [U]	40 GHz [Y]
	Demonstrated	1 THz	200 MHz [F]	700 GHz [J]	2 THz [N]	100 Hz [R]	30 Hz [V, W]	Not known
Circuit Speed	Projected	61 GHz	61 GHz [C]	61 GHz [C]	1 GHz [L]	1 GHz [Q]	10 MHz [U]	Not known
	Demonstrated	5.6 GHz	220 Hz [G]	10 GHz [Z]	1 MHz [F]	100 Hz [R]	30 Hz [V]	Not known
Switching Energy, J	Projected	3E-18	3E-18	>3E-18	1×10 <sup>-18</sup> [L] [>1.5×10 <sup>-17</sup> ] [O]	5E-17 [T]	~1E-17 [V]	3E-18
	Demonstrated	1E-16	1E-11 [G]	1E-13 [K]	8×10 <sup>-17</sup> [P] [>1.3×10 <sup>-14</sup> ] [O]	3E-7 [R]	6E-18 [W]	Not known

From ITRS 2005

# Lowest Barrier: Distinguishability Barrier

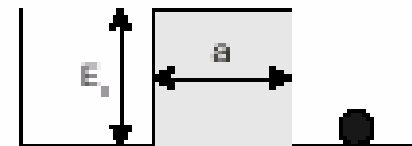


*Distinguishability D* implies low probability  $\Pi$  of spontaneous transitions between two wells (error probability)



$D = \max, \Pi = 0$

$D = 0, \Pi = 0.5$  (50%)



*Classic distinguishability:*

$$\Pi_{classic} = \exp\left(-\frac{E_b}{k_B T}\right)$$

*Minimum distinguishable barrier:*

$$\frac{1}{2} = \exp\left(-\frac{E_b}{k_B T}\right) \longrightarrow E_b = k_B T \ln 2$$

*REF.: R. Cavin, V. Zhirnov, J. Hutchby & G. Burianoff, SRC*

## REF.: R. Cavin & V. Zhirnov, SRC, Essderc 2005 invited talk

From these equations, the **minimum size**  $x_{\min}$ , of a switch is :

$$x_{\min} = \frac{\hbar}{\Delta p} = \frac{\hbar}{\sqrt{2m_e E_{bit}}} = \frac{\hbar}{\sqrt{2m_e k_B T \ln 2}} = \mathbf{1.5 \text{ nm}} \quad (T=300 \text{ K})$$

This minimum size corresponds to a **maximum integration density** of switches:

$$n_{\max} = \frac{1}{x_{\min}^2} = \mathbf{4.7 \times 10^{13} \text{ devices/cm}^2} \quad (\text{ITRS 22nm node } 2.2 \times 10^9 / \text{cm}^2)$$

The **minimum switching time** is estimated as:

$$t_{\min} = \frac{\hbar}{\Delta E} = \frac{\hbar}{k_B T \ln 2} = \mathbf{0.04 \text{ ps}} \quad (\text{ITRS 22nm node } 0.15 \text{ ps})$$

The **power dissipation** per unit area of this limit technology is given by:

$$P = \frac{n_{\max} E_{bit}}{t_{\min}} = \mathbf{3.7 \times 10^6 \text{ W/cm}^2 \text{ ULTIMATE}} \quad (\text{ITRS 22nm node } 100 \text{ W/cm}^2 \text{ Pratical})$$

# OUTLINE

 **INTRODUCTION**

 **MANUFACTURABILITY PROBLEMS**

 **POWER / SPEED PROBLEMS**

 **LOOKING FOR SOLUTIONS – MATERIALS**

 **LOOKING FOR SOLUTIONS – DEVICES**

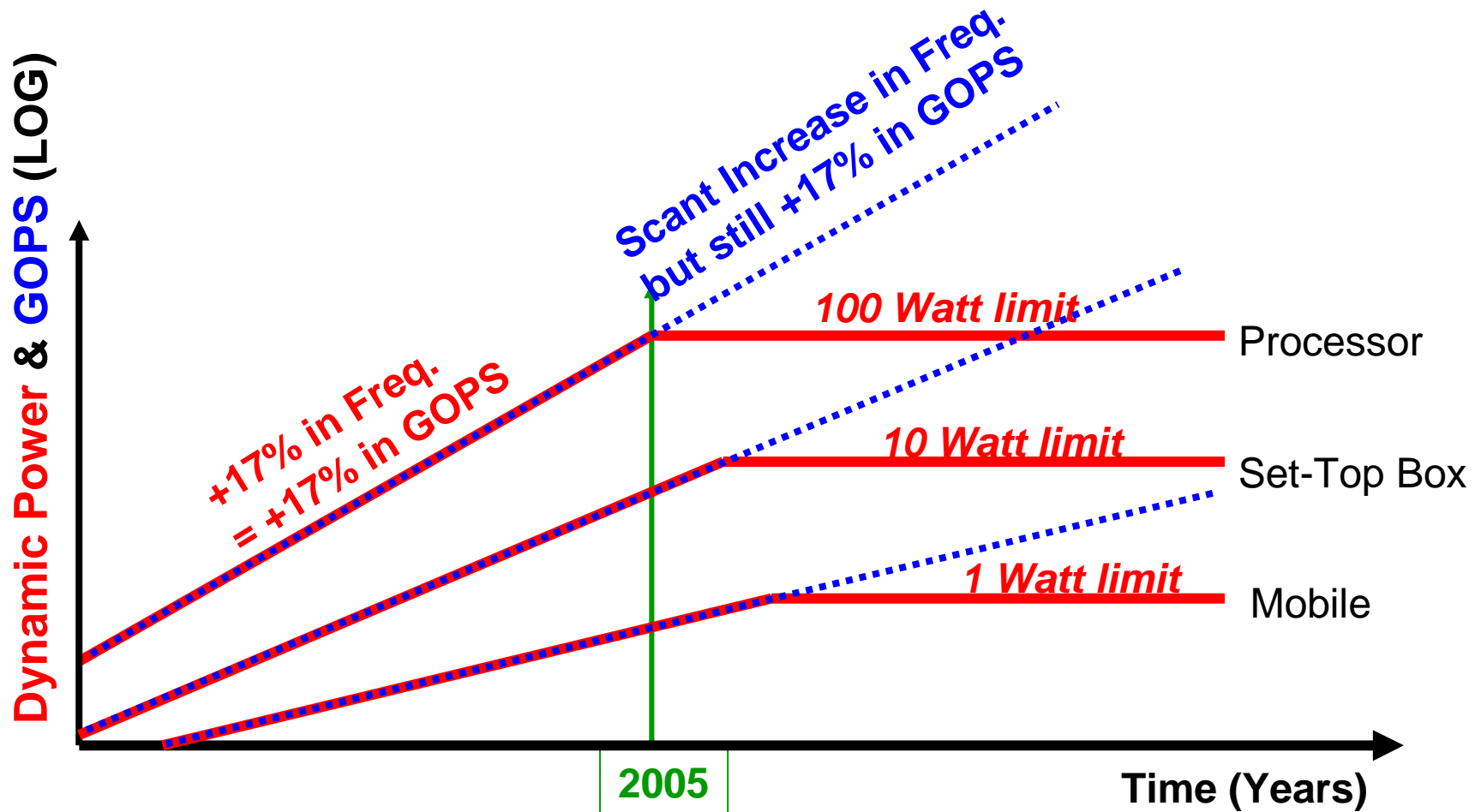
 **LOOKING FOR SOLUTIONS – EMERGING**

 **NEW PARADIGM FOR 22nm and BEYOND**

 **LONGER TERM**

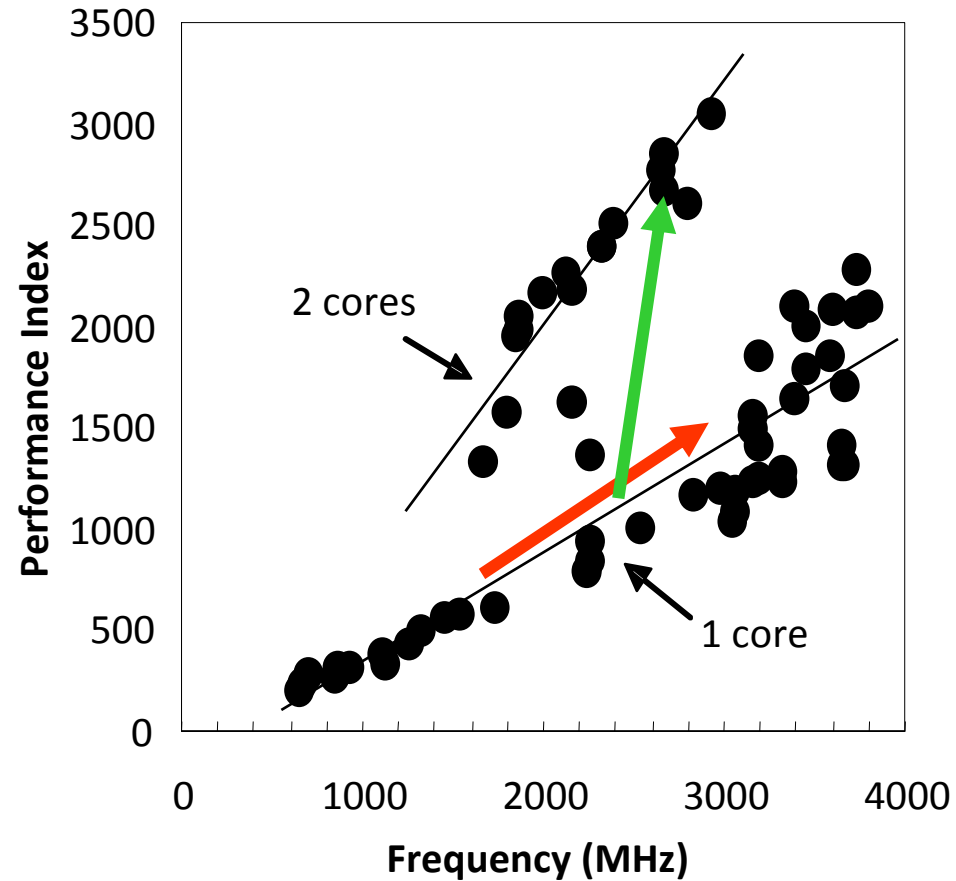
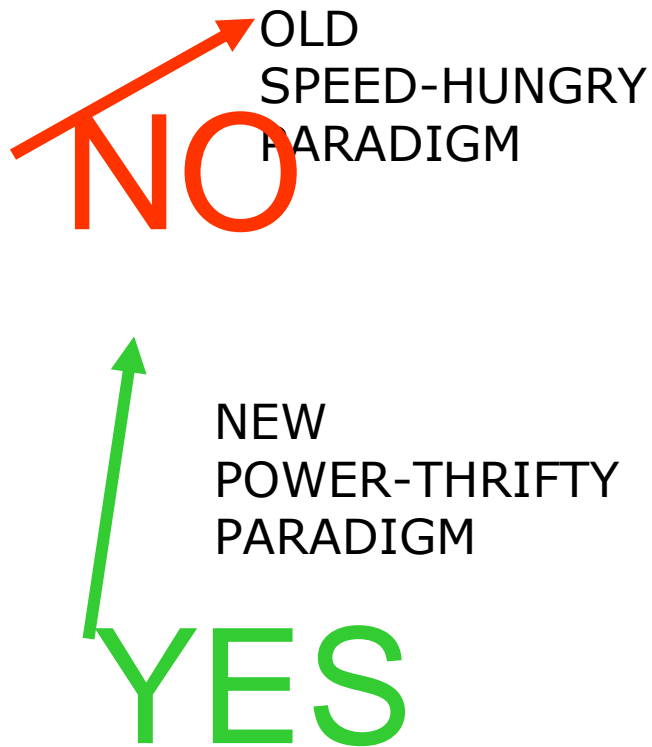
**IN the PAST - Moore's Law powered by power consumption**

**In the FUTURE - will be powered by other tricks**



# DO THE MOORE'S LAWS CONTINUE ?

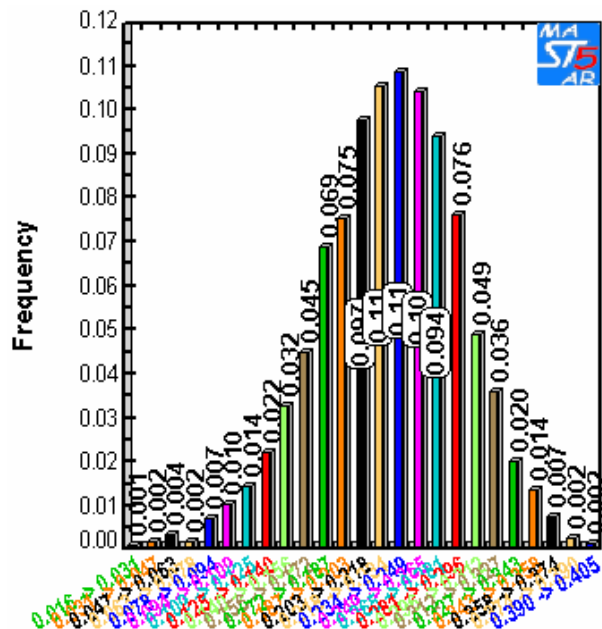
▣ We sell to customers system performance rather than frequency



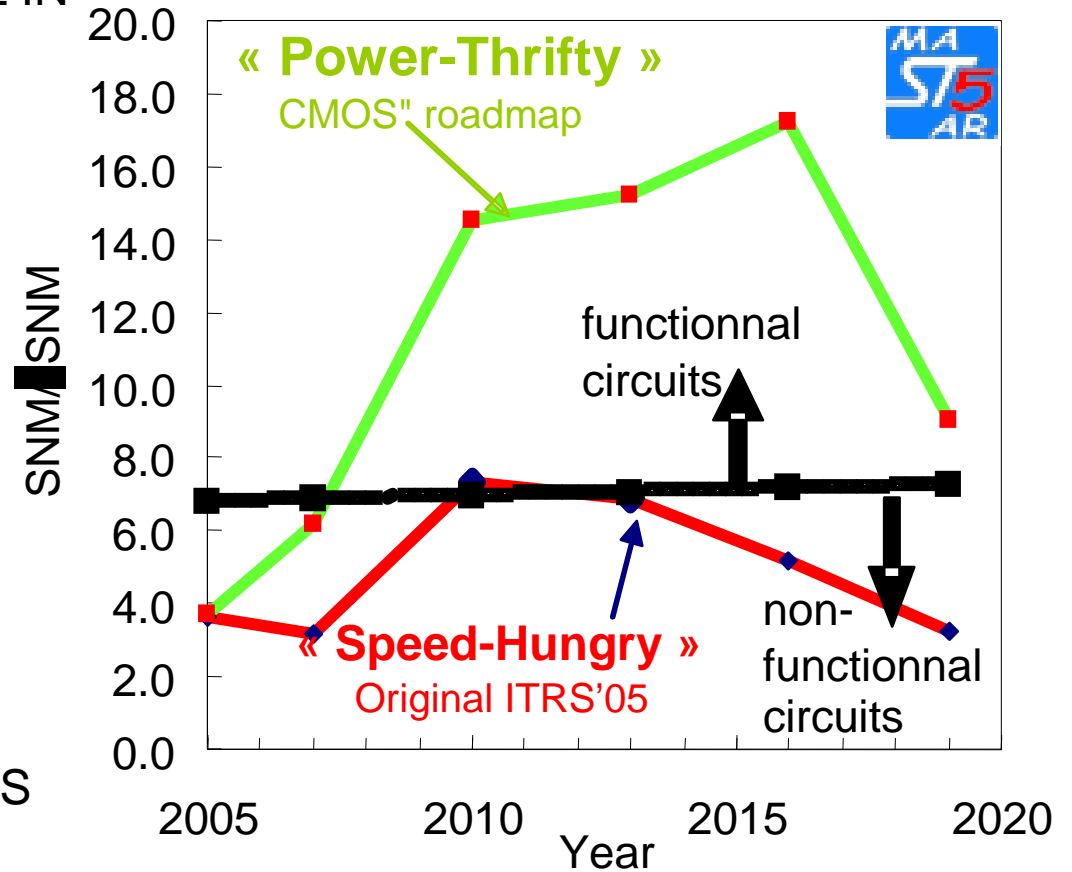
*Datas from Internet benchmarkings on commercial  $\mu$ processors*

# IMPROVED MANUFACTURABILITY :

DEVICES FOR VLSI NEED TO BE REPRODUCIBLE, ALL IDENTICAL IN BILLION PARTS !

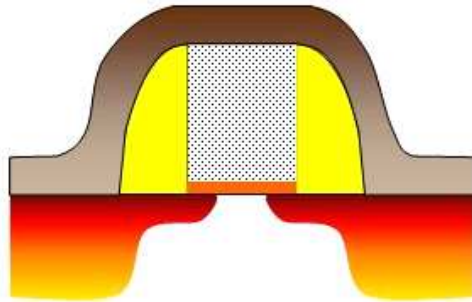


OTHERWISE THEY ARE USELESS EVEN IF SUPER PERFORMANT

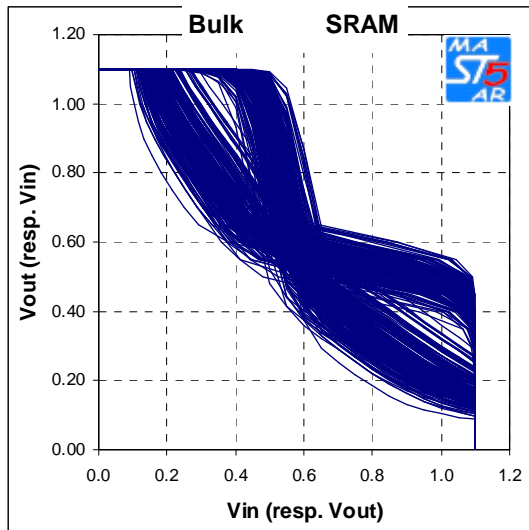


# VARIABILITY -vs- DEVICE STRUCTURE

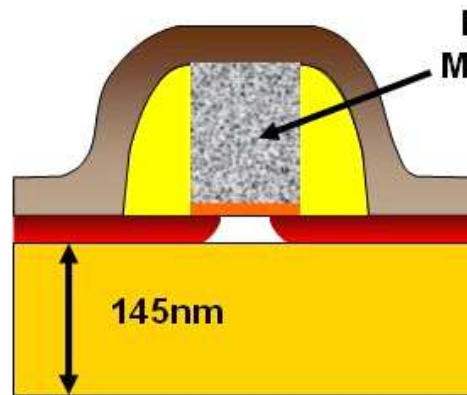
Bulk



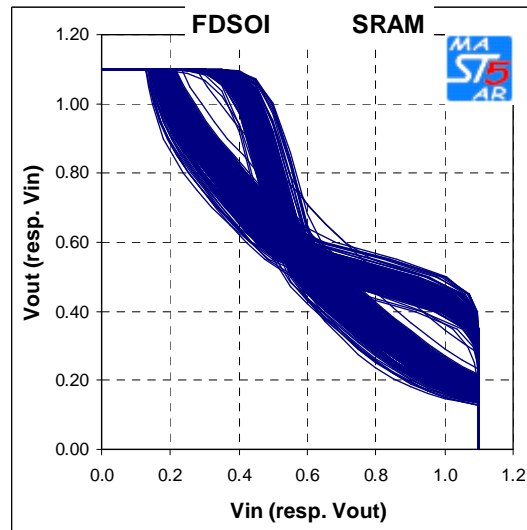
$N_{ch} \sim 3e18 / cm^2$



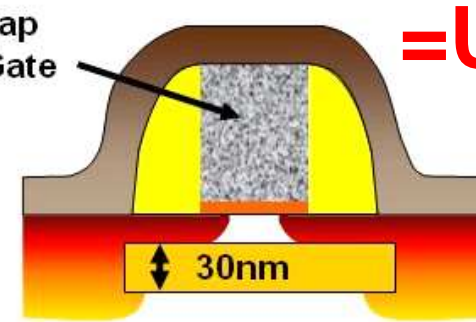
FDSOI Thick BOX



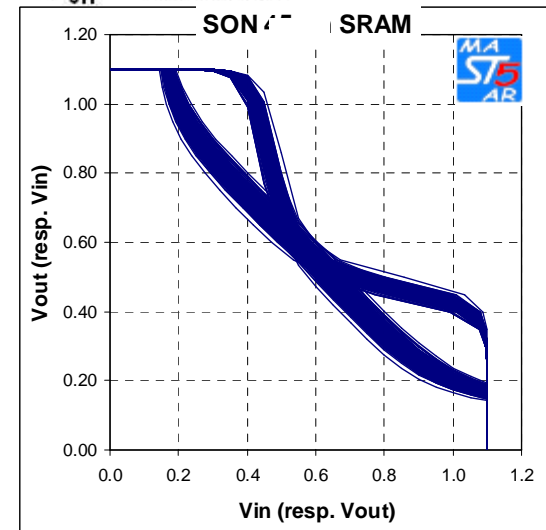
$N_{ch} \sim 1e18 / cm^2$



FDSON/SOI Thin BOX



$N_{ch} \sim 1e16 / cm^2$



$=UTB^2$

# OUTLINE

 **INTRODUCTION**

 **MANUFACTURABILITY PROBLEMS**

 **POWER / SPEED PROBLEMS**

 **LOOKING FOR SOLUTIONS – MATERIALS**

 **LOOKING FOR SOLUTIONS – DEVICES**

 **LOOKING FOR SOLUTIONS – EMERGING**

 **NEW PARADIGM FOR 32nm and BEYOND**

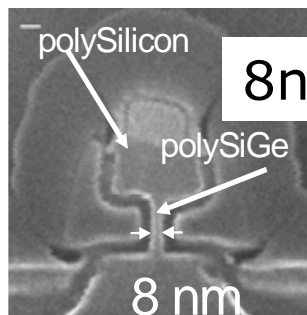
 **LONGER TERM**

# BEYOND 32nm – HOW FAR CAN WE GO ?

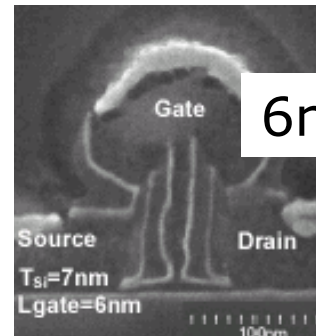
▣ Fundamental limits  $\Delta x \Delta p \geq h/2\pi$  (R. Cavin/V. Zhirnov) -> **1.5nm**

▣ Lateral S-D tunneling (H. Kawaura et al., VLSI'00) -> **2-3nm**

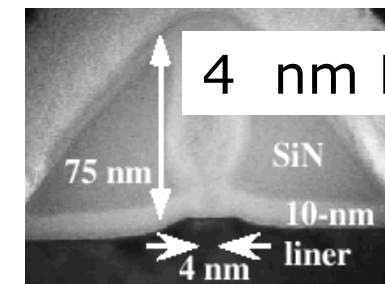
▣ MOSFET  $EI = \left(1 + \frac{X_j^2}{L_{el}^2}\right) \frac{T_{ox\_el}}{L_{el}} \frac{T_{dep}}{L_{el}}$  (Skotnicki) -> **2.7nm**



8nm STM



6nm IBM



4 nm NEC

**DOWN TO 2-4nm - NO NEED FOR ANYTHING ELSE BUT MOSFET,  
BEYOND ? – « DEMOCRACY » - EVERYONE IS EQUAL & NOBODY  
CAN GO BEYOND**

# Conclusions

- ❑ **Circuit and system rather than devices present issues for CMOS scaling**
- ❑ **Power is the key issue for CMOS scaling**
- ❑ **We need NEW CONCEPTS and DEVICES to enable V<sub>dd</sub> lowering**
- ❑ **Devices DON'T NEED so much to be FASTER but rather less vulnerable to fluctuations & less leaky !**
- ❑ **Not enough research is dedicated to intrinsically LP devices, abrupt SS, low DIBL, low leakage, etc.**
- ❑ **SPEED HUNGRY paradigm have to be replaced by POWER THRIFTY (using // $\mu$ P, etc.)**

## ACKNOWLEDGEMENTS

The authors acknowledge kind help and contributions from :

- Abhijid Pethe, Tejas Krishnamohan, Philip H-S. Wong and Krishna Saraswat (Stanford University),
- Viktor Zhirnov, SRC
- Asen Asenov (Glasgow University),
- Konstantin Likharev (Stony Brook University),
- David Esseni (University of Udine),
- Raphaël Clerc and Marlene Ferrier (IMEP),
- Frederic Boeuf, Fabrice Payet, Jean-Pierre Schoellkopf, Manuel Sellier, Estelle Batail, Richard Ferrent, Enrico Perea and Alexandre Villaret, Stephane Monfray, Michel Haond (ST)
- Arnaud Pouydebasque, Peter Stolk (NXP).