IEEE EDS DL Talk

Roadmap for 22nm Logic CMOS and Beyond

January 21, 2009

@IIT-Bombay

Hiroshi Iwai

Tokyo Institute of Technology

2008 IEDM Short Course, Sunday, December 14, 2008
"22 nm CMOS Technology"

Introduction
Kristin De Meyer, IMEC

1. Technology Scaling and Roadmap
Hiroshi Iwai, Tokyo Institute of Technology

2. **22nm Device Architecture and Performance Elements**
Kelin Kuhn, Intel Corporation

3. **Lithography for the 22nm Technology Node**
Kurt Ronse, Geert Vandenberghe, IMEC

4. **BEOL Technology for the 22nm Technology Node**
Jeffrey Gambino, IBM

5. **Device/Circuit Interactions at the 22nm Technology Node**
Kaushik Roy, Purdue University

Outline

1. Scaling

2. ITRS Roadmap

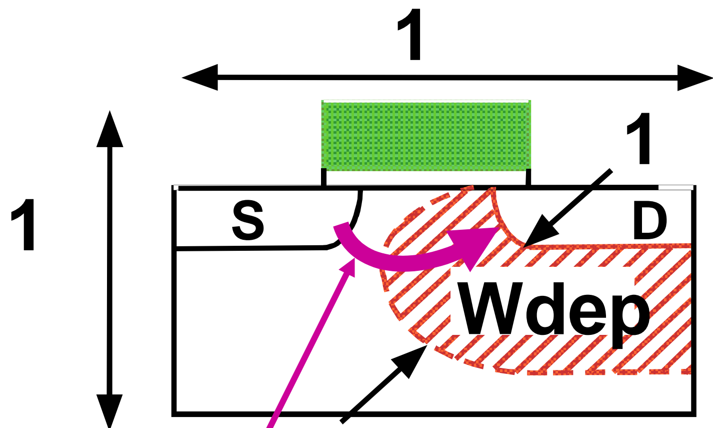
3. Voltage Scaling/ Low Power and Leakage

4. SRAM Cell Scaling

5. Roadmap for further future
   as a personal view

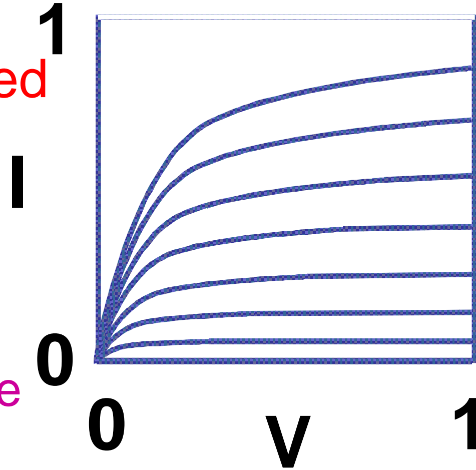
1. Scaling

Scaling Method: by R. Dennard in 1974



Wdep: Space Charge Region (or Depletion Region) Width

Wdep has to be suppressed
Otherwise, large leakage between S and D



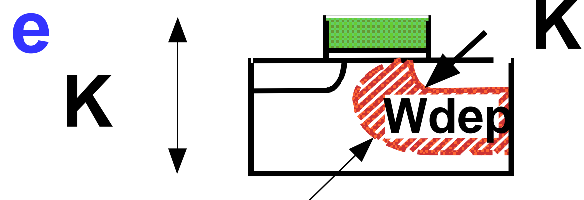
Potential in space charge region is high, and thus, electrons in source are attracted to the space charge region.

K=0.7 for example

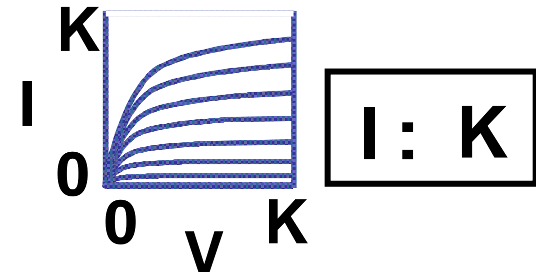
X , Y , Z : K, V : K, Na : 1/K

By the scaling, Wdep is suppressed in proportion, and thus, leakage can be suppressed.

→ Good scaled I-V characteristics



$Wdep \propto \sqrt{V/Na}$
: K



I : K

Downscaling merit: Beautiful!

Geometry & Supply voltage	L_g, W_g T_{ox}, V_{dd}	K	**Scaling K : K=0.7 for example**
Drive current in saturation	I_d	K	$I_d = v_{sat} W_g C_o (V_g - V_{th})$ C_o: gate C per unit area $\rightarrow W_g (t_{ox}^{-1})(V_g - V_{th}) = W_g t_{ox}^{-1}(V_g - V_{th}) = KK^{-1}K = K$
I_d per unit W_g	$I_d/\mu m$	1	I_d per unit $W_g = I_d / W_g = 1$
Gate capacitance	C_g	K	$C_g = \varepsilon_o \varepsilon_{ox} L_g W_g / t_{ox} \rightarrow KK/K = K$
Switching speed	τ	K	$\tau = C_g V_{dd} / I_d \rightarrow KK/K = K$
Clock frequency	f	1/K	$f = 1/\tau = 1/K$
Chip area	A_{chip}	α	α: Scaling factor $\rightarrow$ **In the past, $\alpha > 1$ for most cases**
Integration (# of Tr)	N	α/K^2	$N \rightarrow \alpha/K^2$ **$= 1/K^2$, when $\alpha = 1$**
Power per chip	P	α	$fNCV^2/2 \rightarrow K^{-1}(\alpha K^{-2})K(K^1)^2 = \alpha$ **$= 1$, when $\alpha = 1$**

$k = 0.7$ and $\alpha = 1$	$k = 0.7^2 = 0.5$ and $\alpha = 1$
Single MOFET	
Vdd → 0.7	Vdd → 0.5
Lg → 0.7	Lg → 0.5
Id → 0.7	Id → 0.5
Cg → 0.7	Cg → 0.5
P (Power)/Clock → $0.7^3 = 0.34$	P (Power)/Clock → $0.5^3 = 0.125$
τ (Switching time) → 0.7	τ (Switching time) → 0.5
Chip	
N (# of Tr) → $1/0.7^2 = 2$	N (# of Tr) → $1/0.5^2 = 4$
f (Clock) → $1/0.7 = 1.4$	f (Clock) → $1/0.5 = 2$
P (Power) → 1	P (Power) → 1

- The concerns for limits of down-scaling have been announced for every generation.

- However, down-scaling of CMOS is still the 'royal road'* for high performance and low power.

- Effort for the down-scaling has to be continued by all means.

---

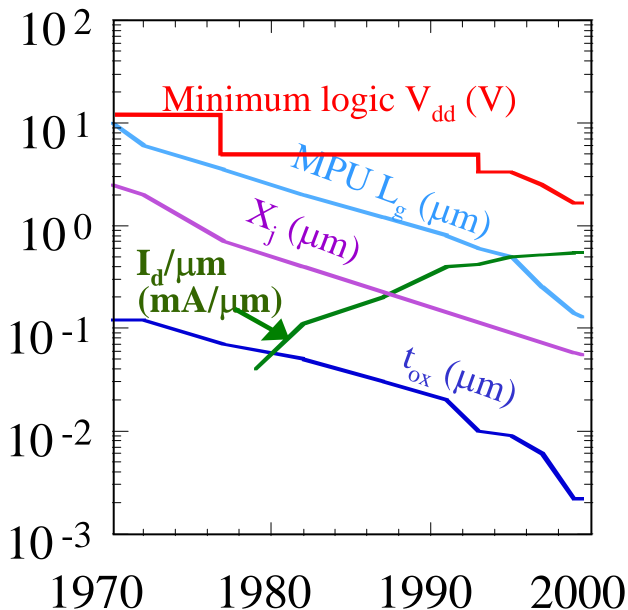
*Euclid of Alexandria (325BC?-265BC?)

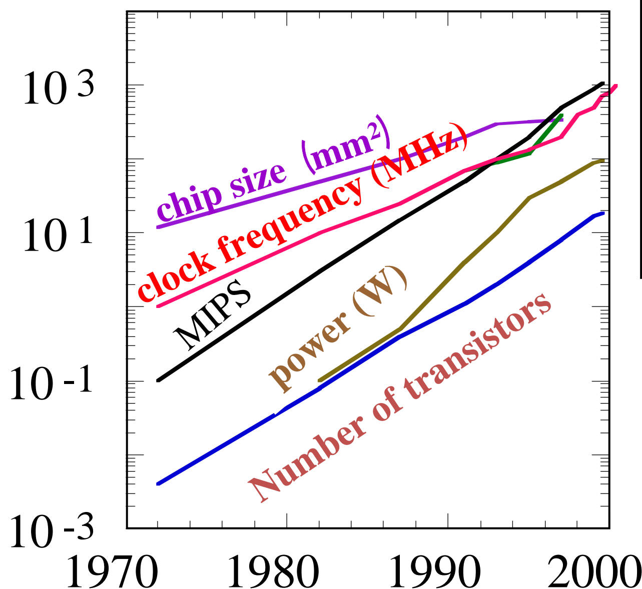
'There is no royal road to Geometry'

Mencius (Meng-zi), China (372BC?-289BC?)

孟子: 王道, 覇道 (Rule of right or virtue vs. Rule of military)

Actual past downscaling trend until year 2000





Past 30 years scaling

Merit: N, f increase

Demerit: P increase

V_{dd} scaling insufficient

Additional significant increase in I_d, f, P

Source. Iwai and S. Ohmi, Microelectronics Reliability 42 (2002), pp.1251-1268

Change in 30 years

	Ideal scaling	Real Change		Ideal scaling	Real Change		Ideal scaling	Real Change
L_g	K	10^{-2}	I_d	K (10^{-2})	10^{-1}	f	1/K(10^2)	10^3
t_{ox}	K(10^{-2})	10^{-2}	I_d/μm	1	10^1	P = fαNCV²	α(10^1)	10^5
V_{dd}	K(10^{-2})	10^{-1}	N	α/K²(10^5)	10^4			
A_{chip}	α	10^1						

Vd scaling insufficient, α increased → N, Id, f, P increased significantly

- Now, power and/or heat generation are the limiting factors of the down-scaling

- Supply voltage reduction is becoming difficult, because Vth cannot be decreased any more, as described later.

- Growth rate in clock frequency and chip area becomes smaller.

2. ITRS Roadmap (for 22 nm CMOS logic)

ITRS Roadmap does change every year!

- 2007 Edition
- 2006 Update
- 2005 Edition
- 2004 Update
- 2003 Edition
- 2002 Update
- 2001 Edition
- 2000 Update

http://www.itrs.net/reports.html

The current latest version: ITRS 2007 Edition

ITRS 2008 Update will be published on the web at the end of Dec 2008 or Jan. 2009

However, discussions for ITRS 2008 Update were open at the '2008 ITRS Summer Public Conference on 16 July 2008' held in SF, US

http://www.itrs.net/Links/2008Summer/Presentations.html

The material in this SC is based on 2008 ITRS Summer Public Conference and ITRS 2007 Edition

2008 ITRS Winter Public Conference, was held on Dec.9. in Seoul. http://www.itrs.net/

The Winter Public Conference data could not be feed-back to this SC material because of print schedule for SC.

What does '22 nm' mean in 22 nm CMOS Logic?

'XX nm CMOS Technology
Commercial Logic CMOS products

Technology name	Starting Year
45 nm	2007
32 nm	2009?
22 nm	2011?~ 2012?
16 nm	2013?~ 2014?

ITRS (Likely in 2008 Update)
for High Performance Logic

Year	Half Pitch (1st Metal)	Physical Gate Length
2007	68 nm	32 nm
2008	59 nm	29 nm
2009	52 nm	27 nm
2010	45 nm	24 nm
2011	40 nm	22 nm
2012	36 nm	20 nm
2013	32 nm	18 nm
2014	29 nm	16 nm

Source: 2008 ITRS Summer Public Conf.

'XX nm' CMOS Logic Technology:
- In general, there is no common corresponding parameter with 'XX nm' in ITRS table, which stands for 'XX nm' CMOS.

What does '22 nm' mean in 22 nm CMOS Logic?

-' XX nm' does not correspond to the 'Half Pitch' nor 'Physical Gate Length' of ITRS.

-'XX nm' is now just a commercial name for CMOS Logic generation of size and its technology.

- Actual parameter values and starting years for commercial products are somewhat different from the above ITRS table, depending on semiconductor companies.

- In 22 and 16 nm technologies, physical gate lengths of high-performance logic device may be close to XX nm.

What does '22 nm' mean in 22 nm CMOS Logic?

8μm → 6μm → 4μm → 3μm → 2μm → 1.2μm → 0.8μm → 0.5μm

- Originally, 'XX' means lithography resolution.
- Thus, 'XX' was the gate length, and half pitch of lines
- 'XX' had shrunk 0.7 in 3 years in average (0.5 in 6 years) those days.
- 'XX' value deviated among companies: example:**1.5μm, 1.2μm, 1μm**

→ 350nm → 250nm → 180nm → 130nm → 90nm → 65nm → 45nm

- 'XX' values were established by NTRS* and ITRS with the term of 'Technology Node**' and 'Cycle***' using typical 'half pitch value'.

  *NTRS: National Tech. Roadmap, **Term 'Technology Node' is not used now.

  ***Cycle: Period or year for which the half pitch becomes X0.71.

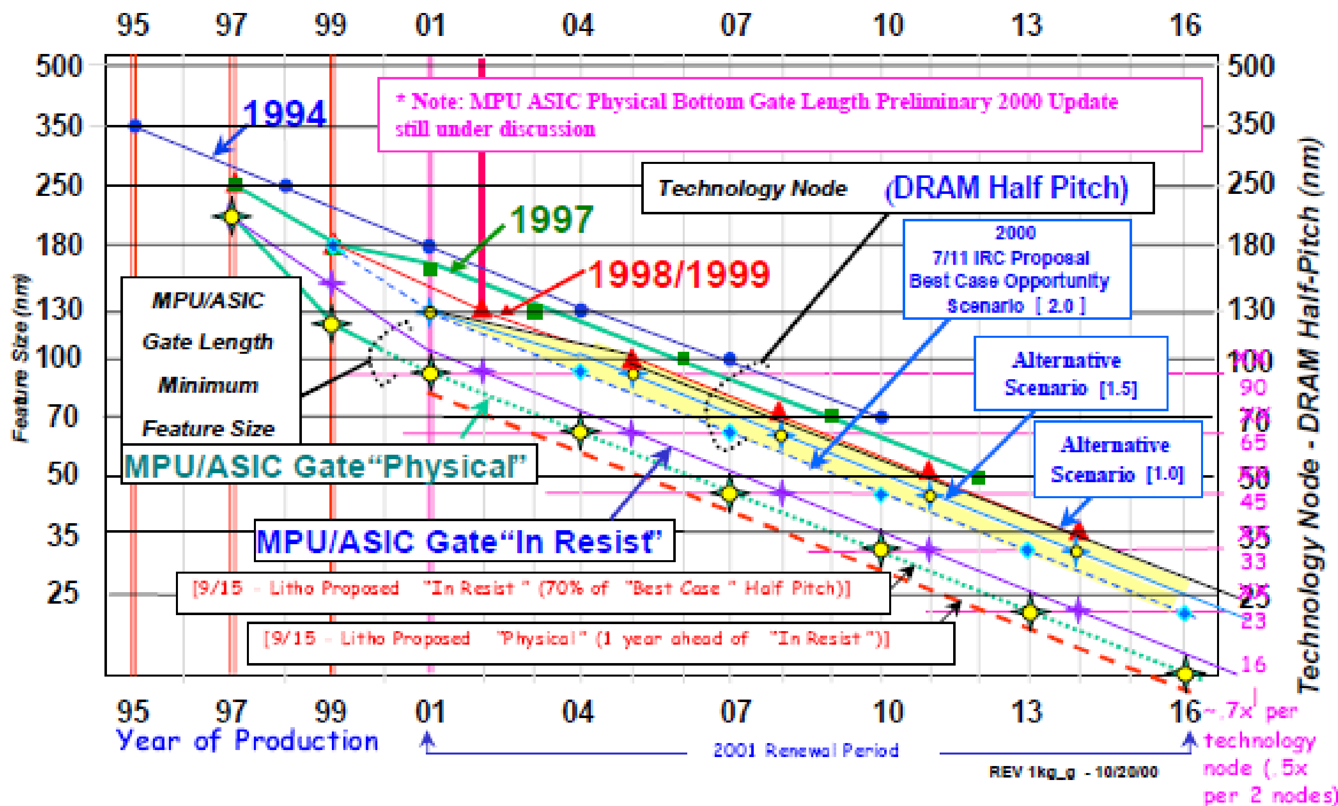
- The gate length of logic CMOS became smaller with one or two generations from the half pitch, and 'XX' names ahead of generations have been used for logic CMOS.

- Memory still keeps the half pitch as the value of 'XX'

→ 32nm → 22nm → 16nm → 11nm → 8nm?? → 5.5nm ??

What does '22 nm' mean in 22 nm CMOS Logic?

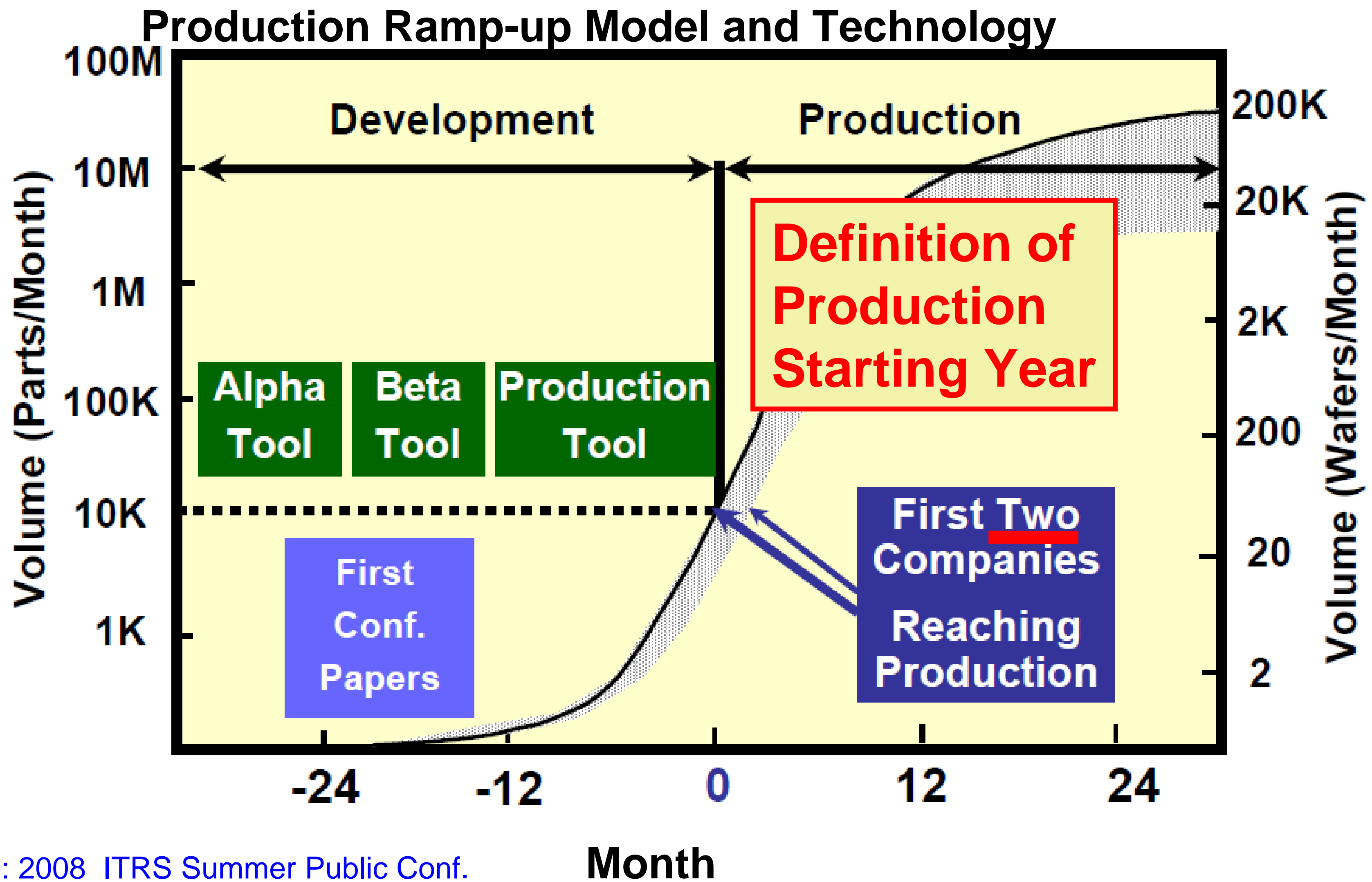
Gate length of Logic CMOS became significantly smaller than lithography resolution or half-pitch using special technique such as resist aching (or trimming) method since 350 nm CMOS.



Source: ITRS 2001 Update

Resist Ashing

Resist

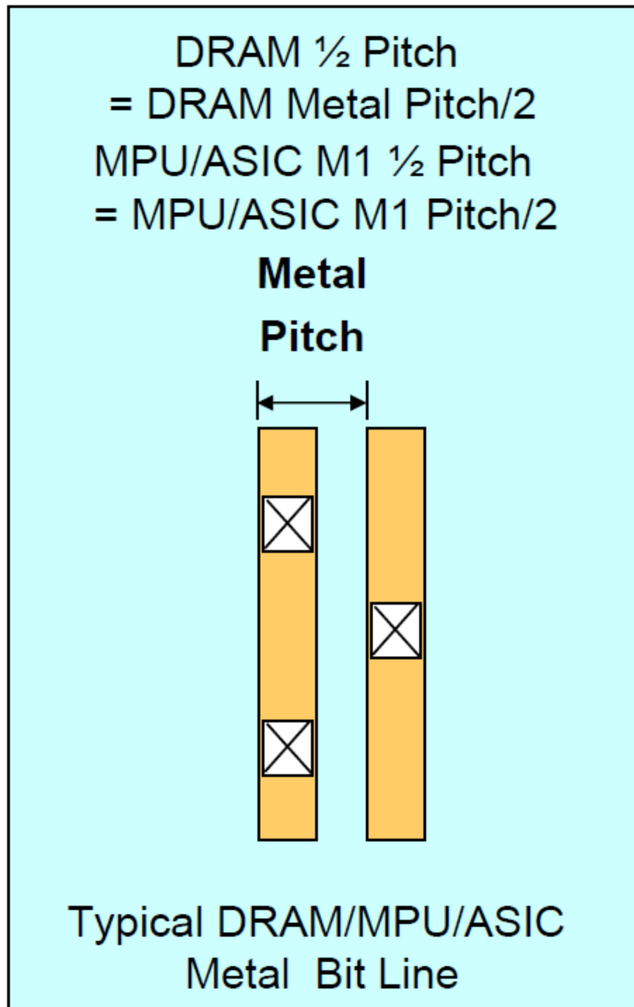


Source: 2008 ITRS Summer Public Conf.

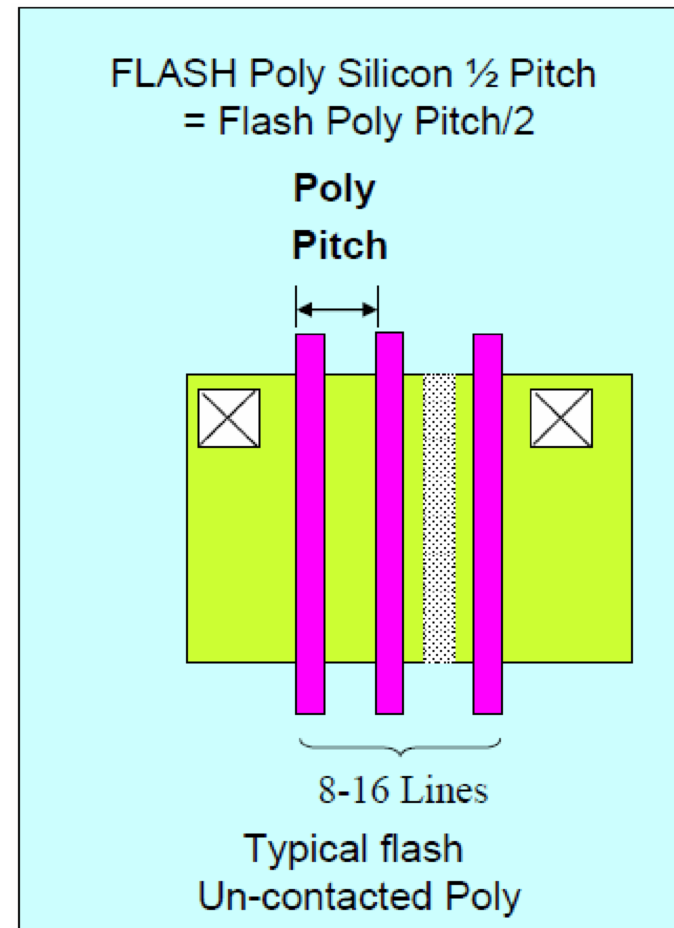
Some Problem: Number of most advanced logic CMOS companies is decreasing in generations.

Definition of the Half Pitch

Logic 1st Metal Half Pitch

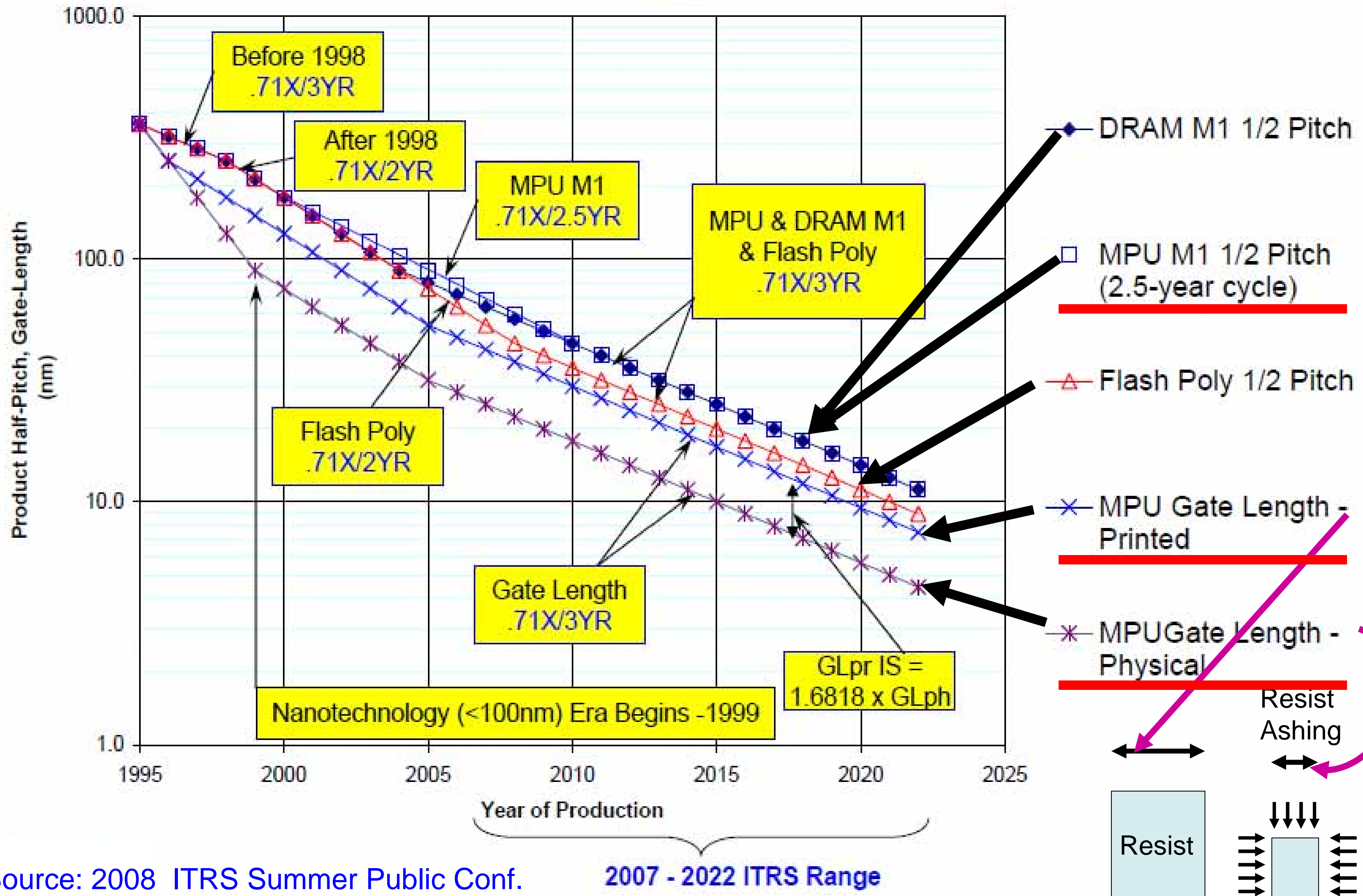


Flash Poly Gate Half Pitch



Source: 2008 ITRS Summer Public Conf.

For example, Typical Half Pitches at ITRS 2007

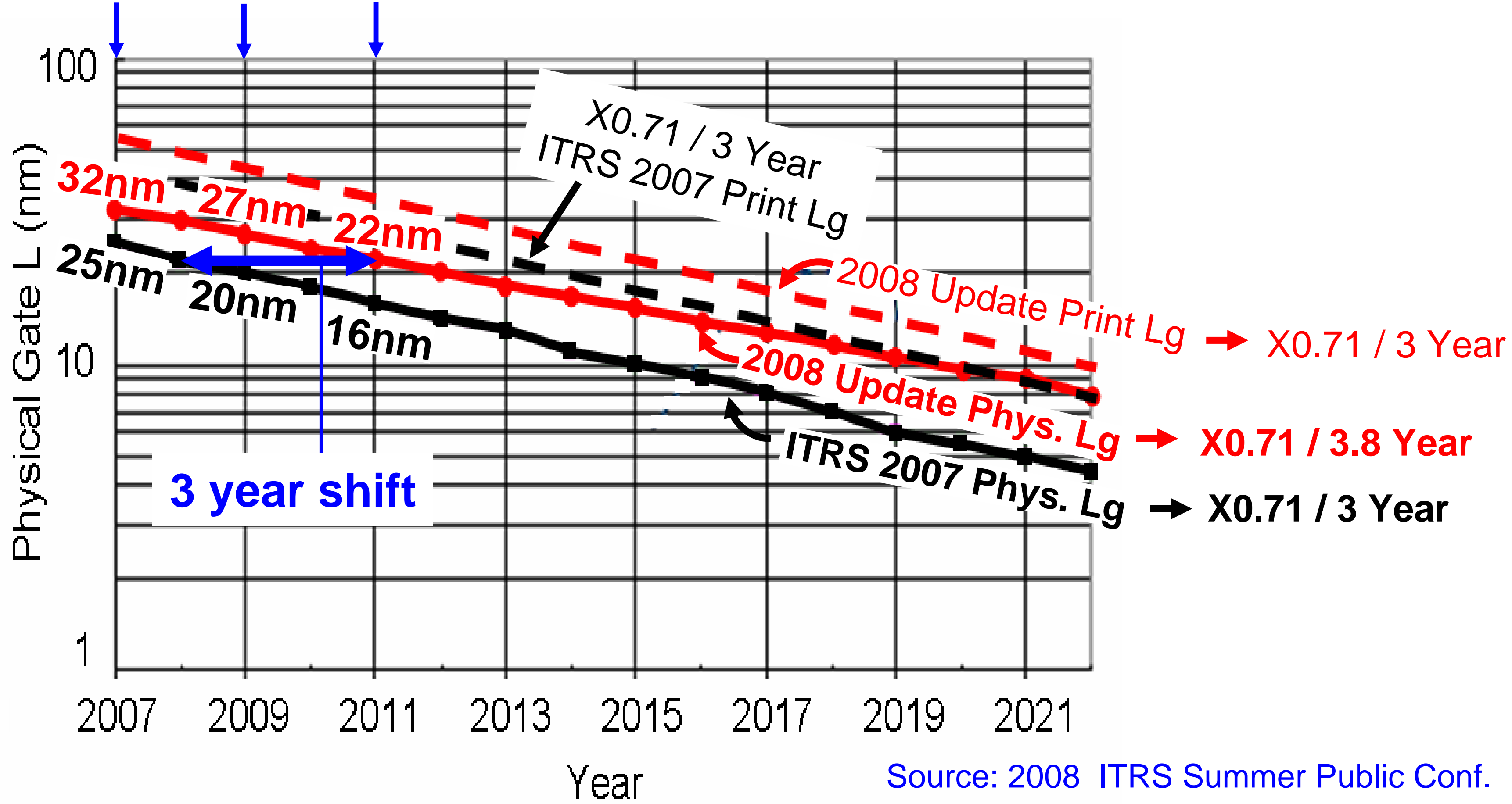


Source: 2008 ITRS Summer Public Conf.

Physical gate length in past ITRS was too aggressive.

The dissociation from commercial product prediction will be adjusted.

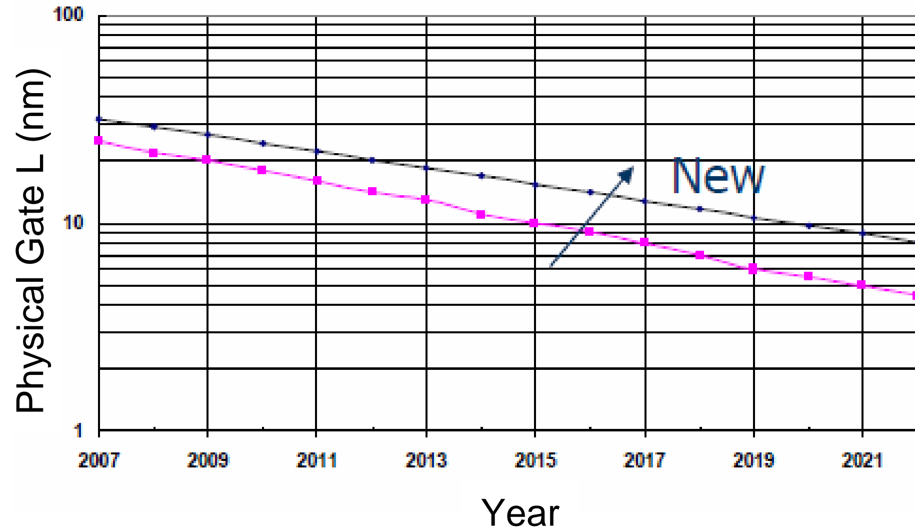
Physical gate length of High-Performance logic will shift by 3-5 yrs.



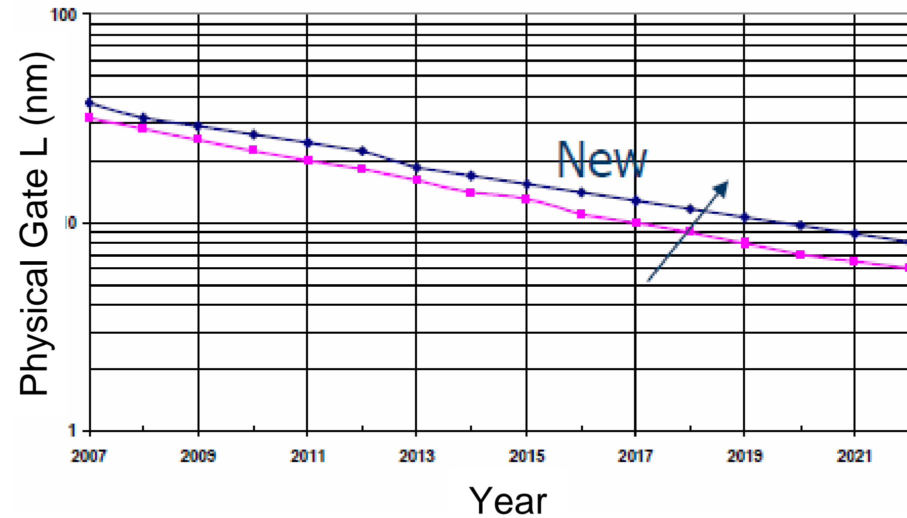
Source: 2008 ITRS Summer Public Conf.

L of HP and LOP will shift, but LSTP will not.

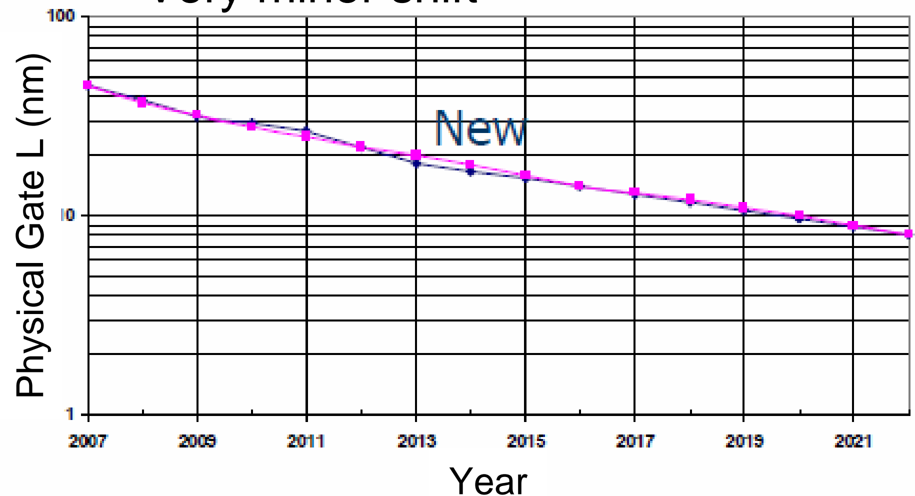
HP (High Performance) Logic
shifts by 3 – 5 years



LOP (Low Operation Power) Logic
shifts by 1 – 3 years



LSTP (Low Standby Power) Logic
Very minor shift



Source: 2008 ITRS Summer Public Conf.

EOT and Xj shift backward, corresponding to Lg shift

EOT: 0.55 nm → 0.88 nm, Xj: 8 nm → 11 nm @ 22nm CMOS

Likely in 2008 Update

Correspond to 22nm

Source: 2008/ ITRS Summer Public Conf.

Year of Production	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022
2007 MPU/ASIC Lg (nm)	25	23	20	18	16	14	13	11	10	9	8	7	6.3	5.6	5	4.5
2008 MPU/ASIC Lg (nm)	32	29	27	24	22	20	18	17	15	14.0	12.8	11.7	10.7	9.7	8.9	8.1
Shift/Interpolate Formua	2005	intrp	intrp	intrp	intrp	2009	2010	intrp	intrp	2012	intrp	intrp	intrp	intrp	intrp	intrp
EOT w/3E20 poly, bulk MPU (nm)	1.2	0.71	0.54	0.41												
EOT w/3E20 poly, bulk MPU (nm)	1.3	1.2	1.2	1	0.68	0.54	0.41									
EOT w/metal gate, bulk MPU (nm)		0.9	0.75	0.65	0.55	0.50										
EOT w/metal gate, bulk MPU (nm)			1.0	0.95	0.88	0.75	0.65	0.60	0.53	0.5						
Drain Ext. X_j bulk MPU (nm)	12.5	11	10	9	8	7										
Drain Ext. X_j bulk MPU (nm)	11	11	11	11	11		9	8.5	7.7	7						

Likely in 2008 Update

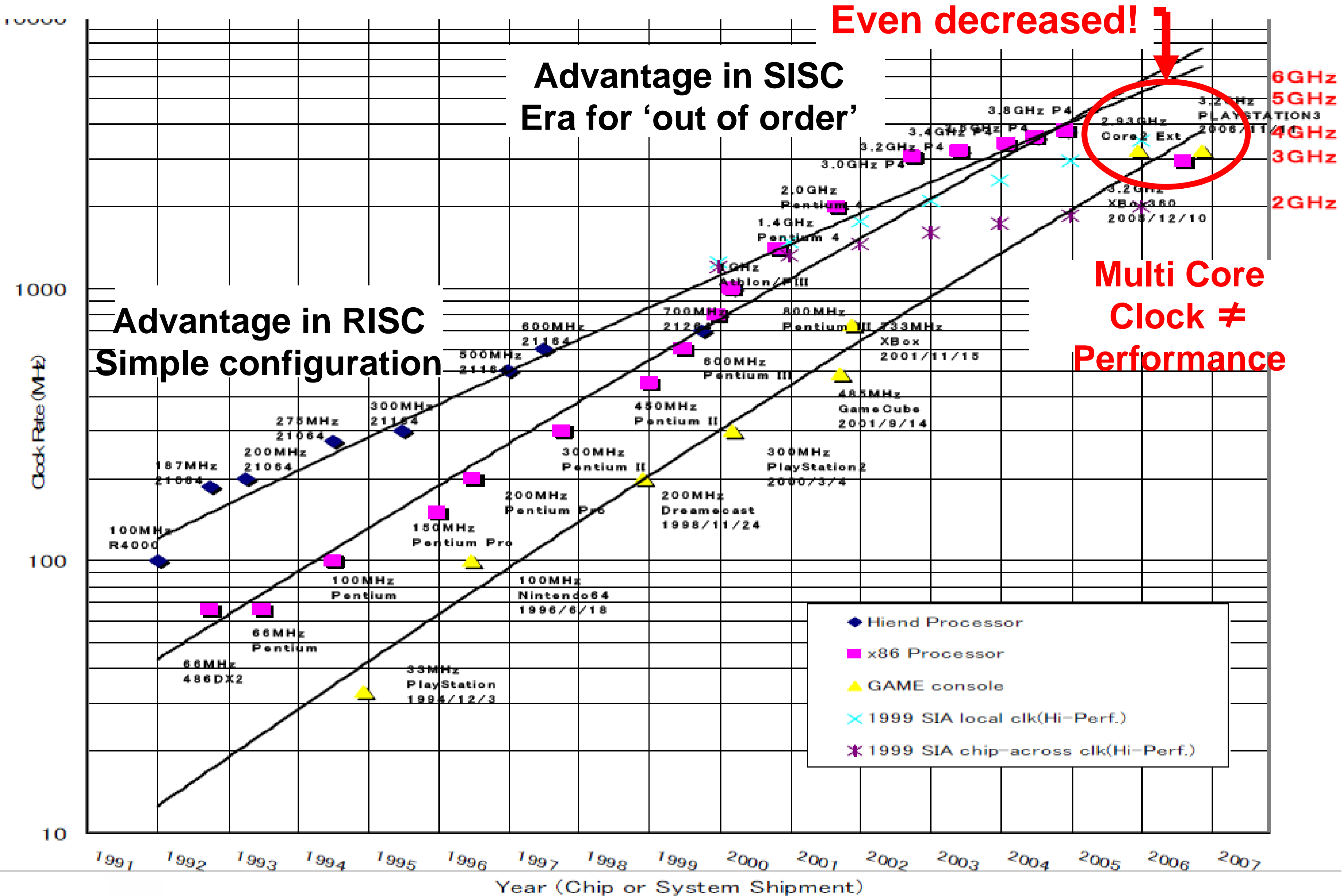
Likely in 2008 Update

Likely in 2008 Update

non-steady trend corrected

filled in for metal gate EOT for 2009/10 based on latest conference presentations

Clock frequency does not increase aggressively anymore.



Source: Mitsuo Saito, Toshiba

MPU "GHz" by "Cores" ITRS2007

Continued?

Core Clock Frequency

Actual

Forecast

CLK (GHz)
Clk Roadmap
Cores * Clock

Chip Frequency

CPU "GHz"

1000
100
10
1
0.1
0.01
0.001
0.0001

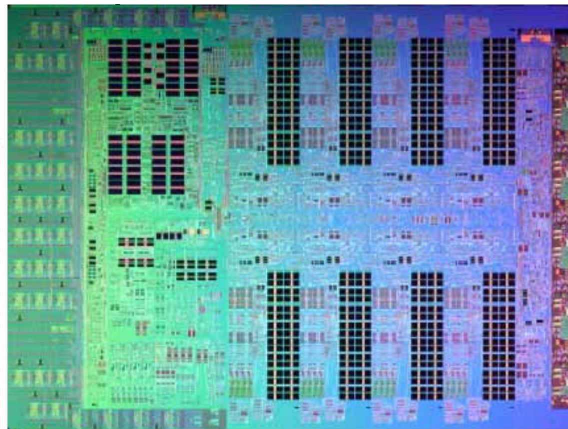
1970 1980 1990 2000 2010 2020 2030

8080

4004

Pentium

Cell Broadband Engine



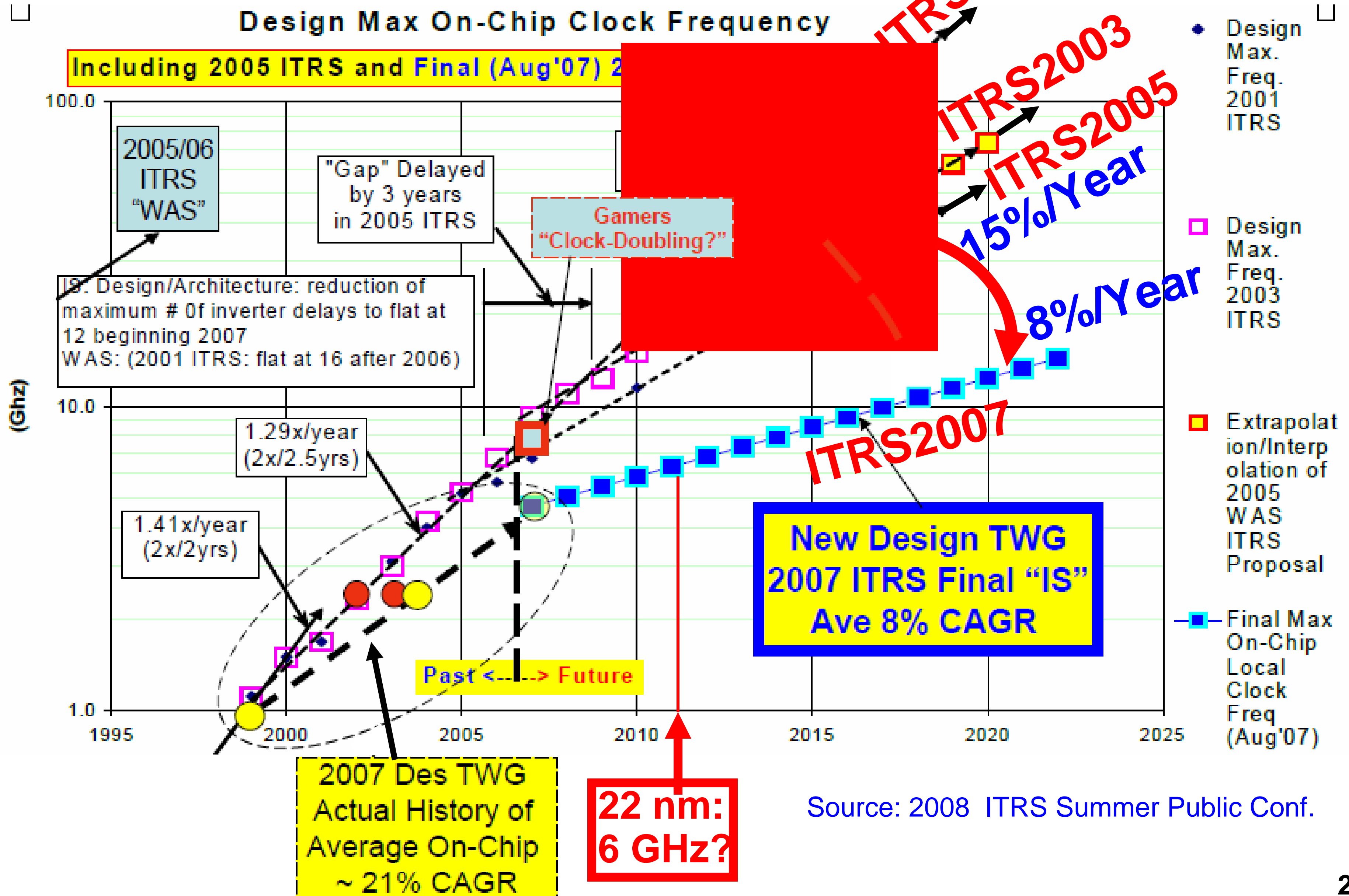
6GHz capability for SRAM

Source:
IBM, Toshiba, Sony
ISSCC2007 and 08

Source: 2007 ITRS Winter Public Conf.

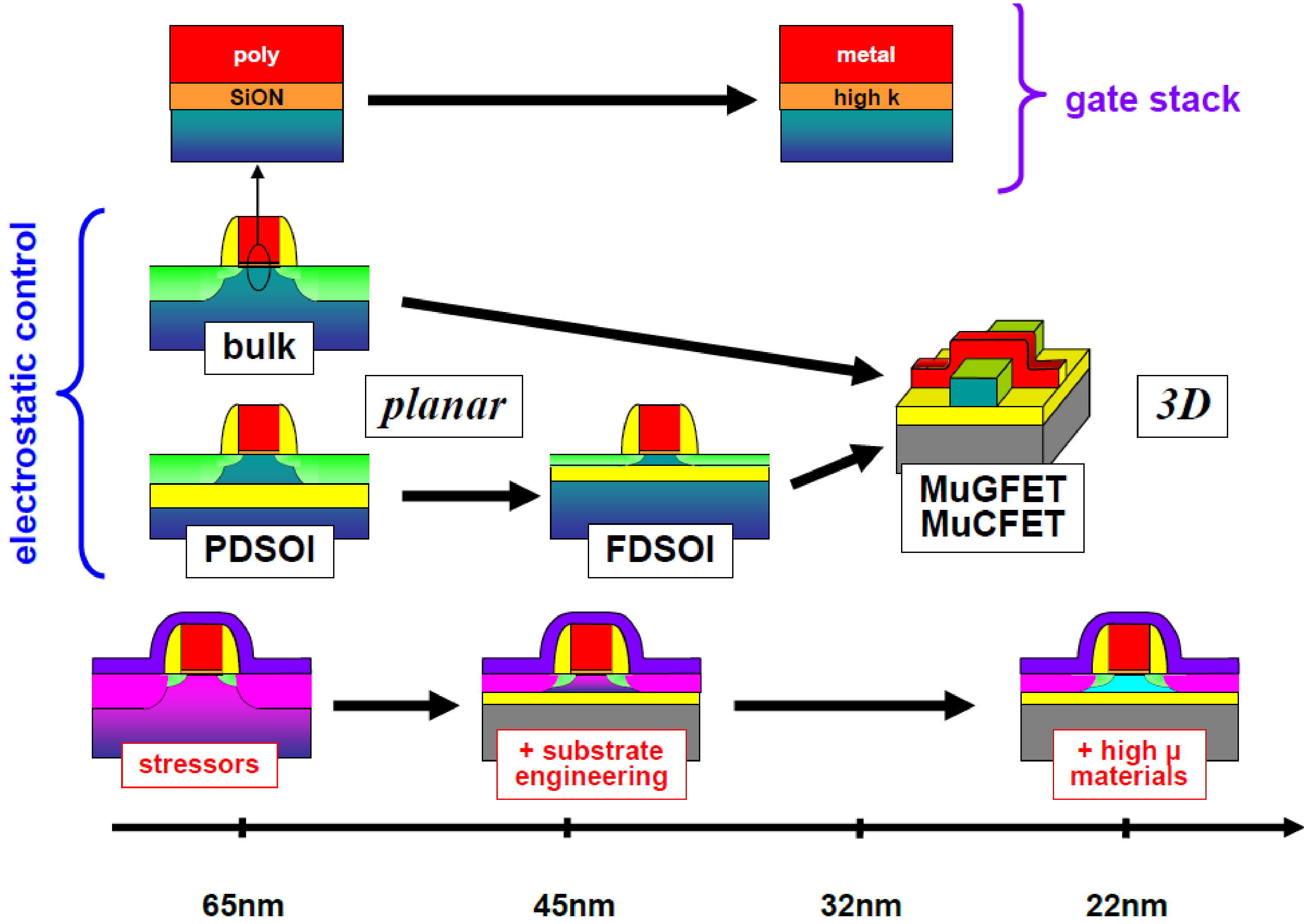
Clock frequency Change in the past ITRS

(Max on chip frequency or 'Core clock')



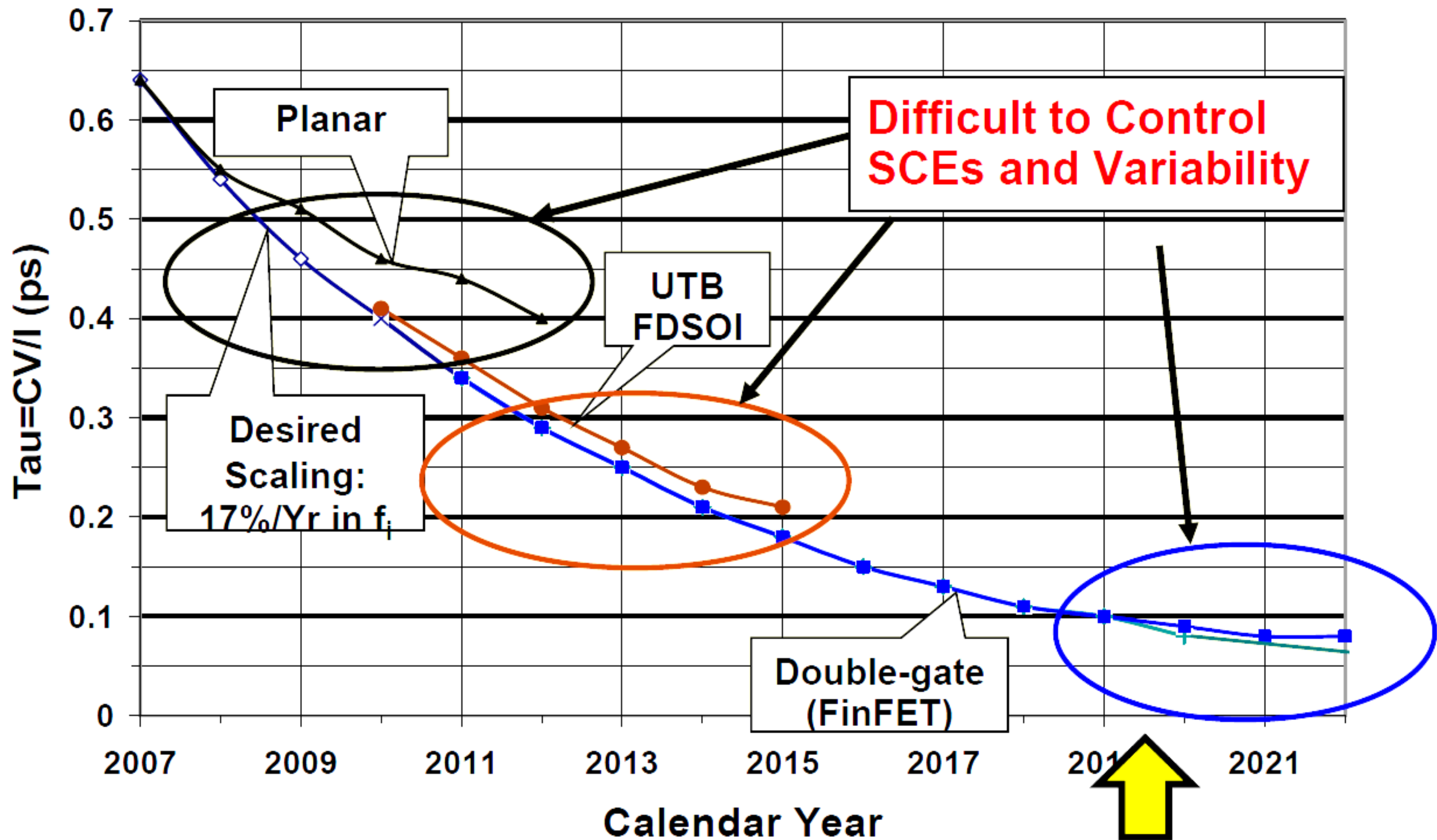
Source: 2008 ITRS Summer Public Conf.

Structure and technology innovation (ITRS 2007)



Source: 2008 ITRS Summer Public Conf.

Technology innovation described in ITRS 2007

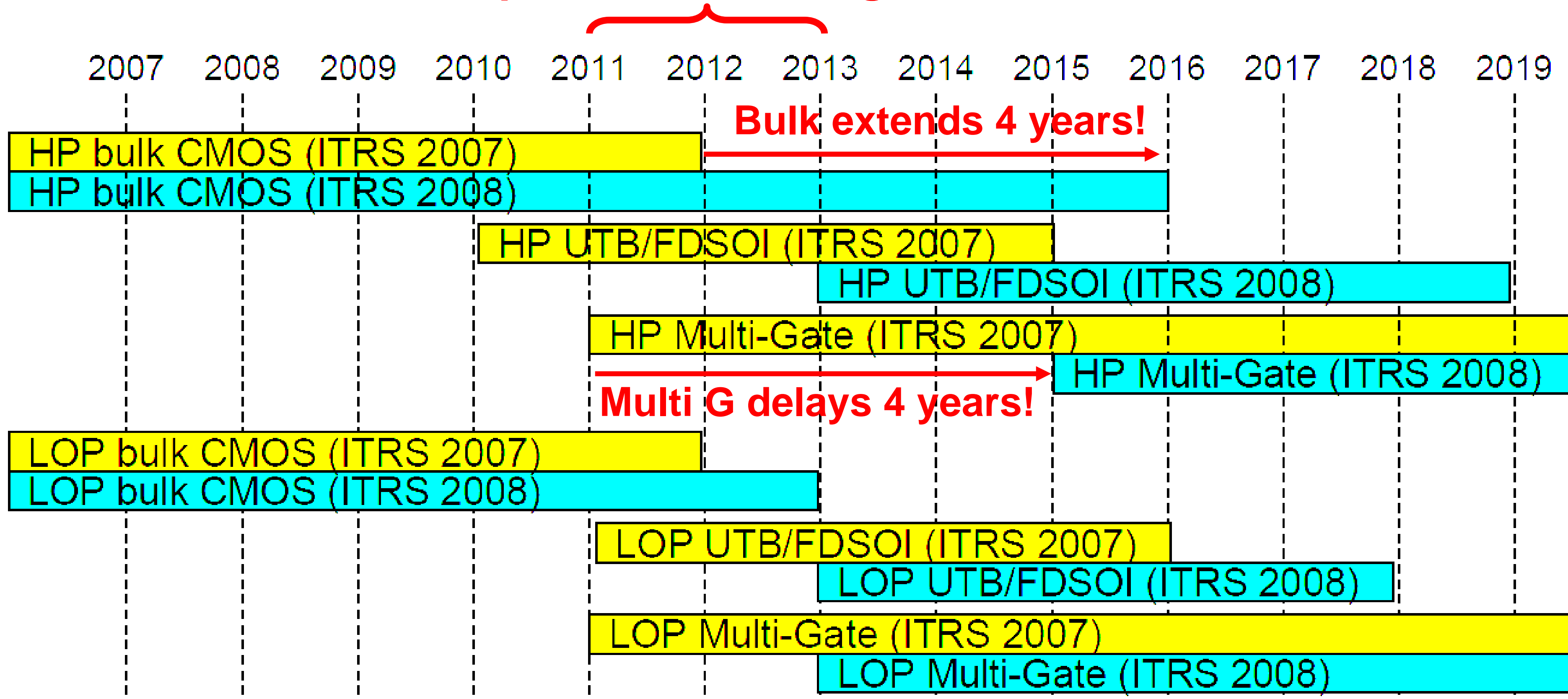


Alternative material (Ge, III-V) and structure (Nanowire) in channel region.

Source: 2007 ITRS Winter Public Conf.

Timing of CMOS innovations shifts backward.

Bulk CMOS has longer life now!



Source: 2008 ITRS Summer Public Conf.

Wafer size (ITRS 2007)

Correspond to 22nm

Year of Production	*2007*	*2008*	*2009*	*2010*	*2011*	*2012*	*2013*	*2014*	*2015*
MPU/ASIC Metal 1 (M1) ½ Pitch (nm)(contacted)	*68*	*59*	*52*	*45*	*40*	*36*	*32*	*28*	*25*
MPU Physical Gate Length (nm)	*25*	*23*	*20*	*18*	*16*	*14*	*13*	*11*	*10*
MPU High-Performance Total Chip Area(mm^2)	*310*	*246*	*195*	*310*	*246*	*195*	*310*	*246*	*195*
MPU High-Performance Active Transistor Area(mm^2)	*31.7*	*25.1*	*20.0*	*31.7*	*25.1*	*20.0*	*31.7*	*25.1*	*20.0*
*General Characteristics * (99% Chip Yield)*									
Maximum Substrate Diameter (mm)—High-volume Production (>20K wafer starts per month)**	300	300	300	300	300	450	450	450	450

Source: ITRS 2007

??

Maybe delay??

Gate CD (Critical Dimension) Control

ITRS 2007

Correspond to 22nm Logic

Year of Production	*2007*	*2008*	*2009*	*2010*	*2011*	*2012*	*2013*	*2014*	*2015*
MPU/ASIC Metal 1 (M1) ½ Pitch (nm)(contacted)	*68*	*59*	*52*	*45*	*40*	*36*	*32*	*28*	*25*
MPU Physical Gate Length (nm)	*25*	*23*	*20*	*18*	*16*	*14*	*13*	*11*	*10*
Total maximum allowable etch 3σ (nm), including photoresist trim and gate etch [AA]	**1.5**	**1.38**	**1.2**	**1.08**	**0.96**	**0.84**	**0.78**	**0.66**	**0.6**

Source: ITRS 2007

2008 Update

Correspond to 22nm Logic

Year of Production	*2007*	*2008*	*2009*	*2010*	*2011*	*2012*	*2013*	*2014*	*2015*
MPU Physical Gate Length (nm)	32	29	27	24	22	20	18	17	15
L_{gate} *3 σ variation (nm) [Z]*	3.82	3.49	3.18	2.9	2.65	2.42	2.21	2.02	1.84

Source: 2008 ITRS Summer Public Conf.

Gate CD control color changed to 'white' through 2011 and to 'yellow' for 2012 reflecting the new Lg scaling

ITRS2008 Low-k Roadmap Update

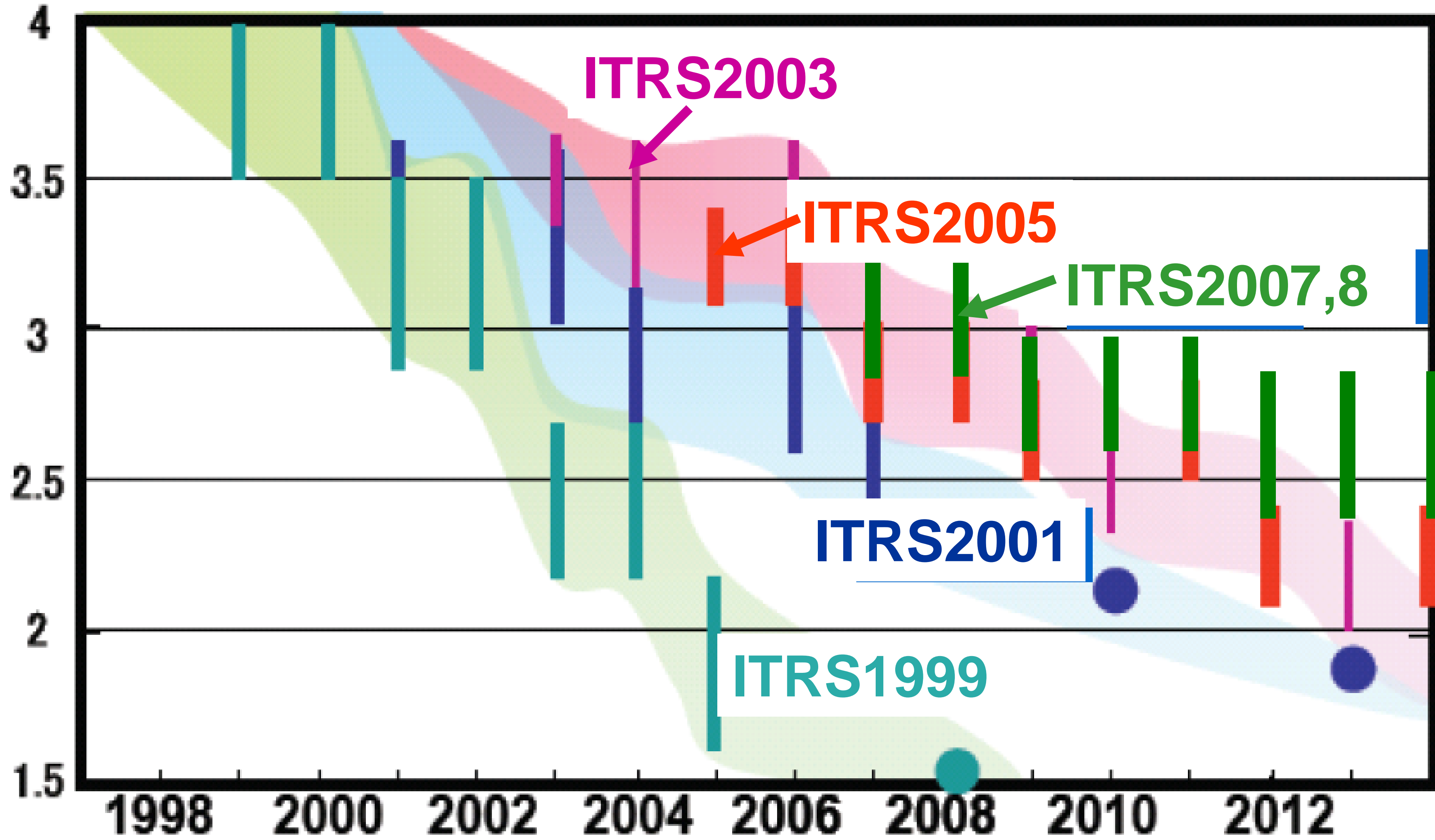
Correspond to 22nm Logic

		Near-term					
ITRS 2007	*Year of Production*	*2008*	*2009*	*2010*	*2011*	*2012*	*2013*
	Interlevel metal insulator – effective dielectric constant (κ)	2.7-3.0	2.5-2.8	2.5-2.8	2.5-2.8	2.1-2.4	2.1-2.4
Update 2008	Interlevel metal insulator – effective dielectric constant (κ)	2.9-3.3	2.6-2.9	2.6-2.9	2.6-2.9	2.4-2.8	2.4-2.8
ITRS 2007	Interlevel metal insulator – bulk dielectric constant (κ)	2.3-2.7	2.1-2.4	2.1-2.4	2.1-2.4	1.8-2.1	1.8-2.1
Update 2007	Interlevel metal insulator – bulk dielectric constant (κ)	2.5-2.8	2.3-2.6	2.3-2.6	2.3-2.6	2.1-2.4	2.1-2.4

Source: 2008 ITRS Summer Public Conf.

k value increases by 0.1 ~ 0.3

Historical Transition of ITRS Low-k Roadmap



Source: 2008 ITRS Summer Public Conf.

Roadmap towards 22nm technology and beyond

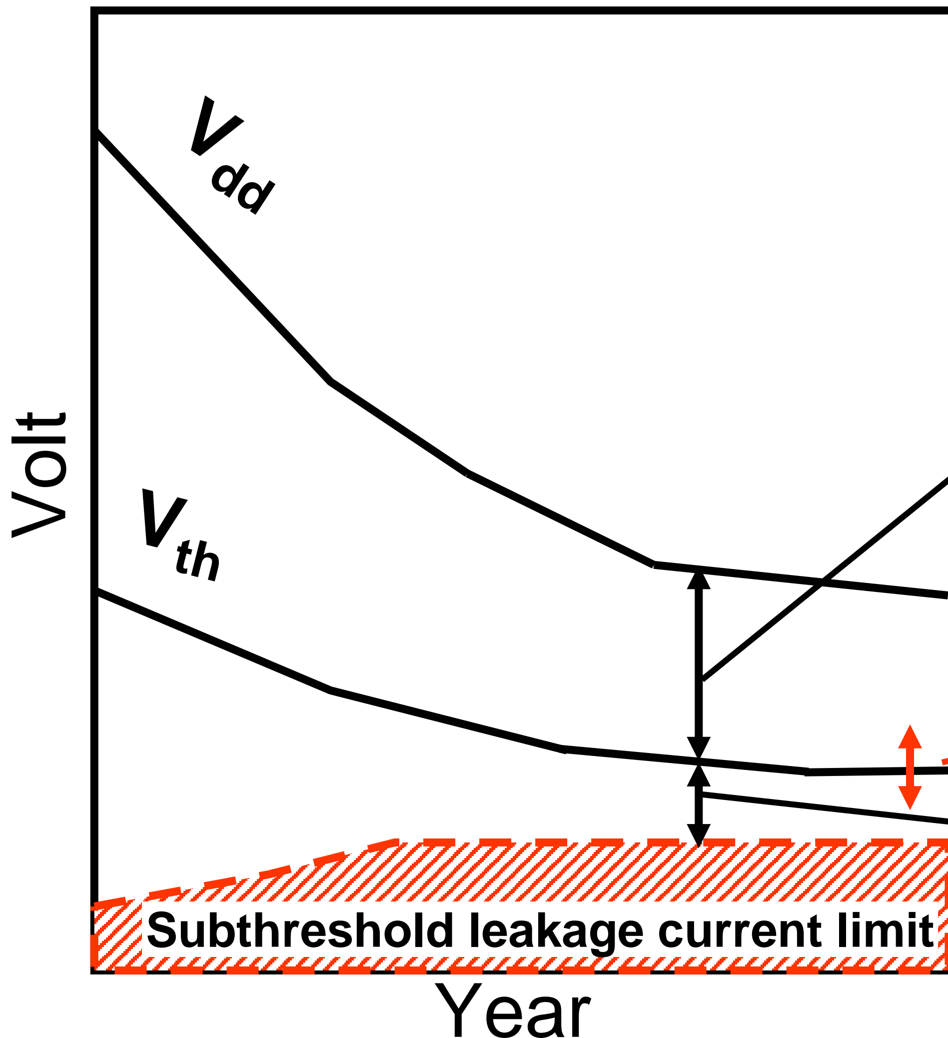
- Physical gate length downsizing rate will be less aggressive.

- Corresponding to the above, performance increase would slow down – Clock frequency, etc.

- Introduction of innovative structures – UTB SOI and DG delayed, and bulk CMOS has longer life than predicted by previous ITRS roadmaps.

3. Voltage Scaling / Low Power and Leakage

Difficulty in Down-scaling of Supply Voltage: Vdd



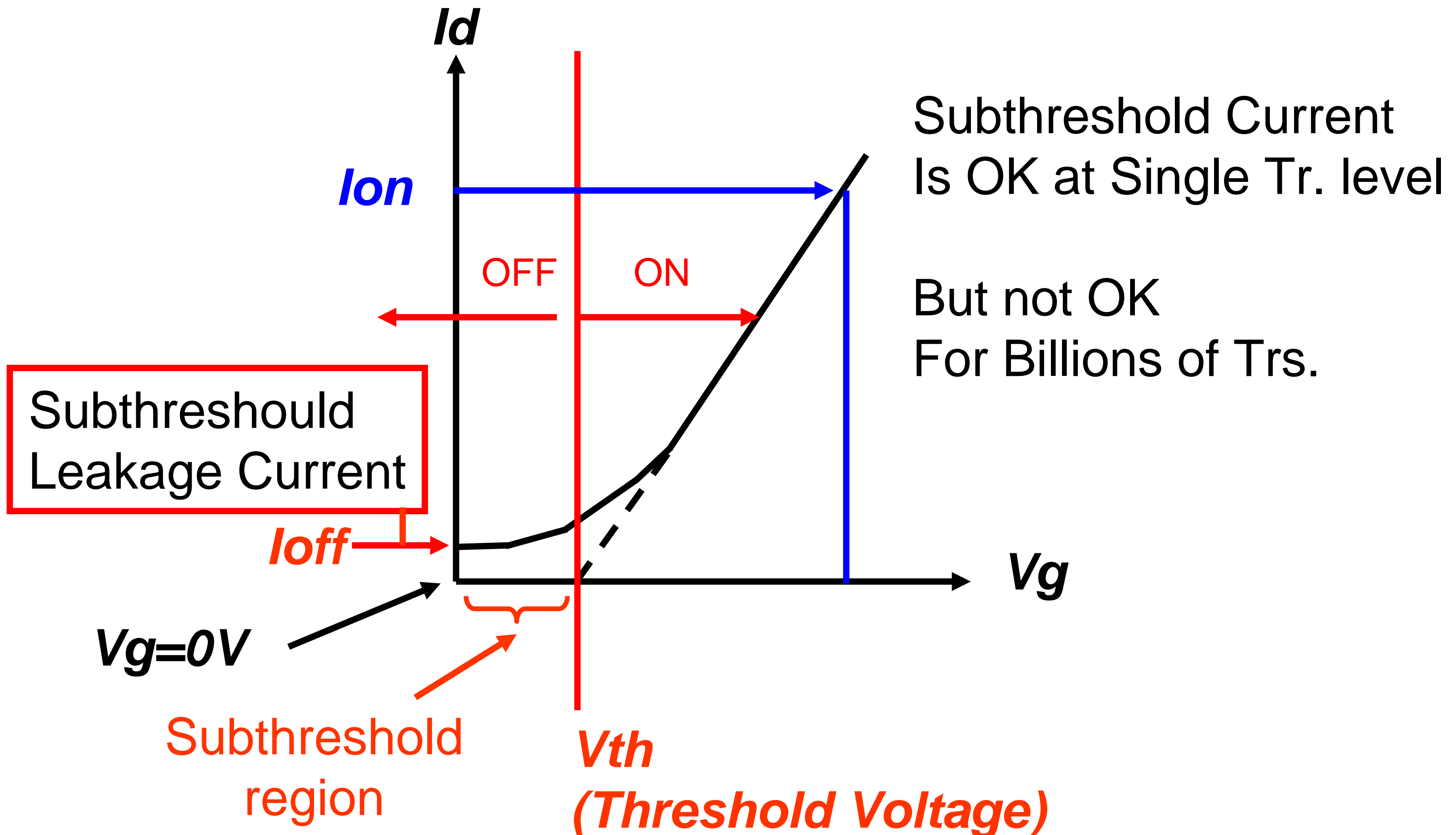
Because, V_{th} cannot be down-scaled anymore, V_{dd} down-scaling is difficult.

$V_{dd} - V_{th}$ determines the performance (High Id) and cannot be too small.

ΔV_{th}: V_{th} variation

$> \Delta V_{th}$
Margin for V_{th} variation is necessary

Subtheshold leakage current of MOSFET



Vth cannot be decreased anymore

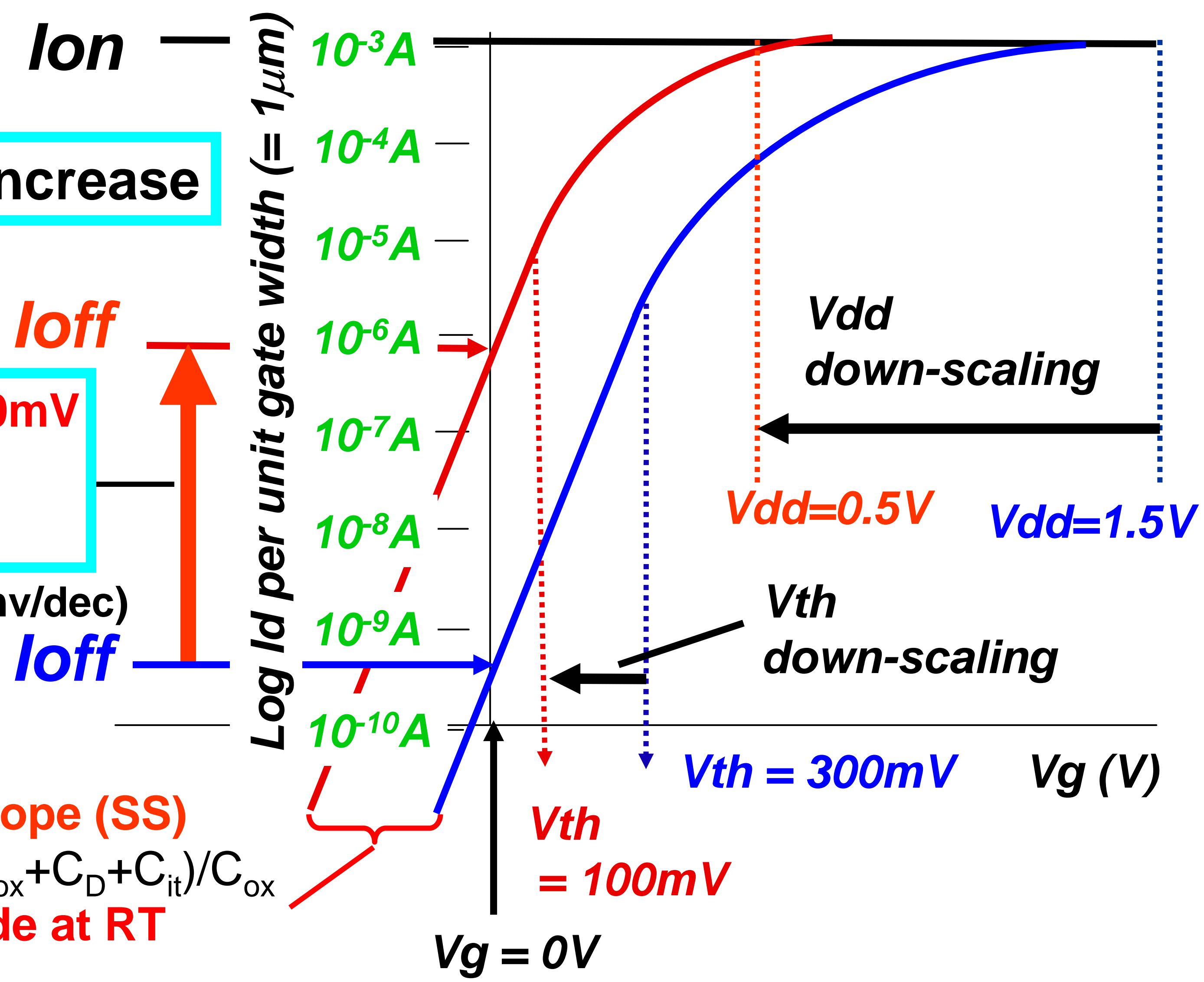
Log scale Id plot

significant Ioff increase

Vth: 300mV → 100mV

Ioff increases with 3.3 decades

(300 – 100)mV/(60mv/dec) = 3.3 dec



Subthreshold slope (SS)

$= (\mathrm{Ln}10)(kT/q)(C_{ox}+C_D+C_{it})/C_{ox}$

> ~ 60 mV/decade at RT

SS value:

Constant and does not become small with down-scaling

ITRS for HP logic

S-D leakage current

Isd-leak has to be stay less than 1μA/μm

Id-sat growth will be modest in 2008 update

Saturated Drain current

Leakage current (μA/μm)

1.0E+1
1.0E+0
1.0E-1
1.0E-2
1.0E-3

2001
2005, 2007, 2008
2005 DG
1999

2004 2007 2010 2013 2016 2019 2022

Year

- 2008up (bulk)
- 2008up (UTB)
- 2008up (DG)
- 2007 (bulk)
- 2007 (UTB)
- 2007 (DG)
- 2005 (bulk)
- 2008 (UTB)
- 2005 (DG)
- 2003
- 2001
- 1999

Saturation current (μA/μm)

3000
2500
2000
1500
1000
500

2005 DG
2007 DG
2005 Blk
2008 DG
2008 UTB
2008 Blk
2001
1999

2004 2007 2010 2013 2016 2019 2022

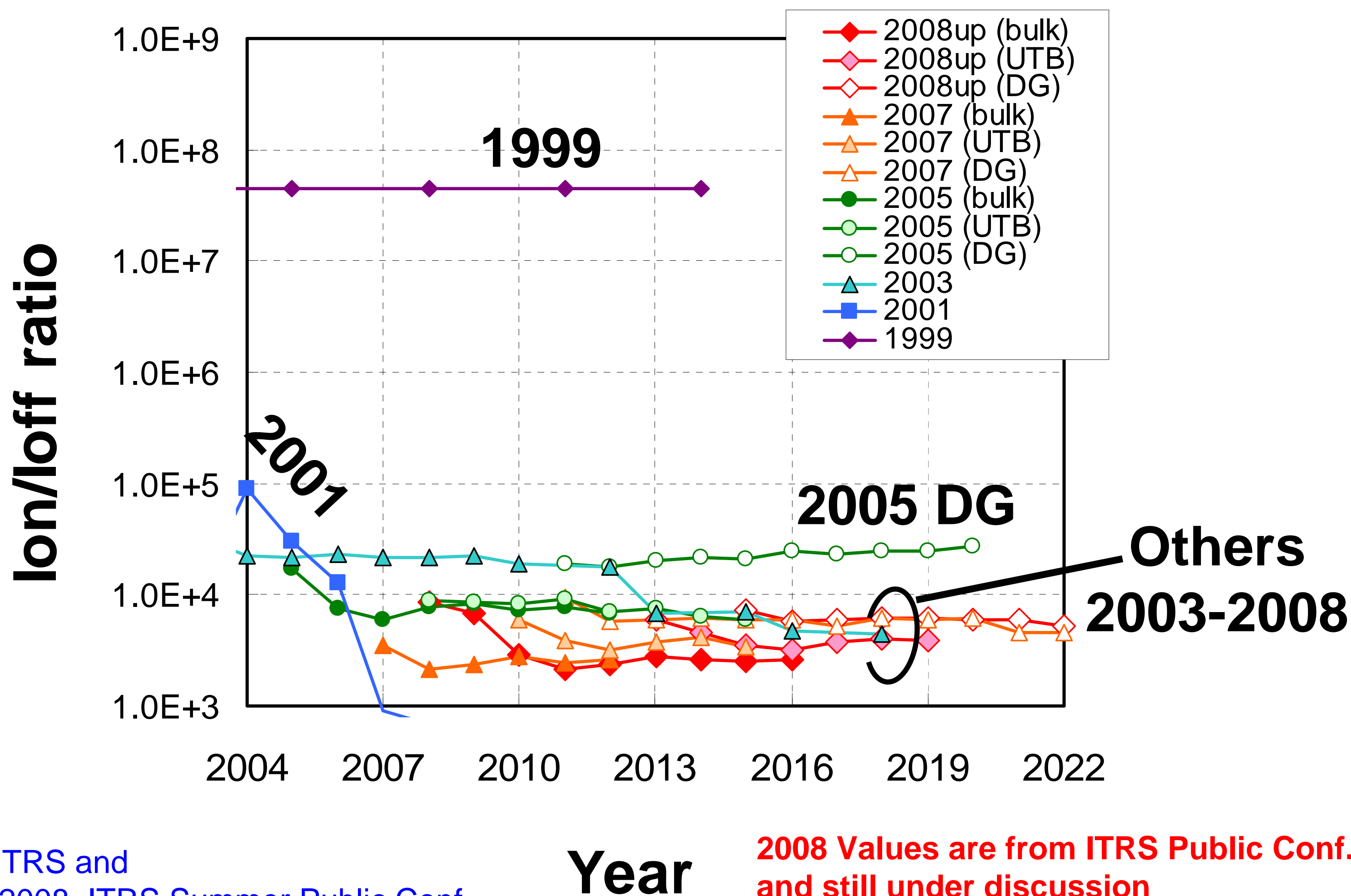
Year

- 2008up (bulk)
- 2008up (UTB)
- 2008up (DG)
- 2007 (bulk)
- 2007 (UTB)
- 2007 (DG)
- 2005 (bulk)
- 2005 (UTB)
- 2005 (DG)
- 2003
- 2001
- 1999

Source: ITRS and 2008 ITRS Summer Public Conf.

2008 Values are from ITRS Public Conf. and still under discussion

ITRS for HP logic



Source: ITRS and 2008 ITRS Summer Public Conf.

2008 Values are from ITRS Public Conf. and still under discussion

ITRS for HP logic

Vdd

Vdd will stay higher in 2008 update

2008

2003, 2005, 2007

1999

2001

Vdd (V)

1.2
1
0.8
0.6
0.4
0.2
0

2004 2007 2010 2013 2016 2019 2022

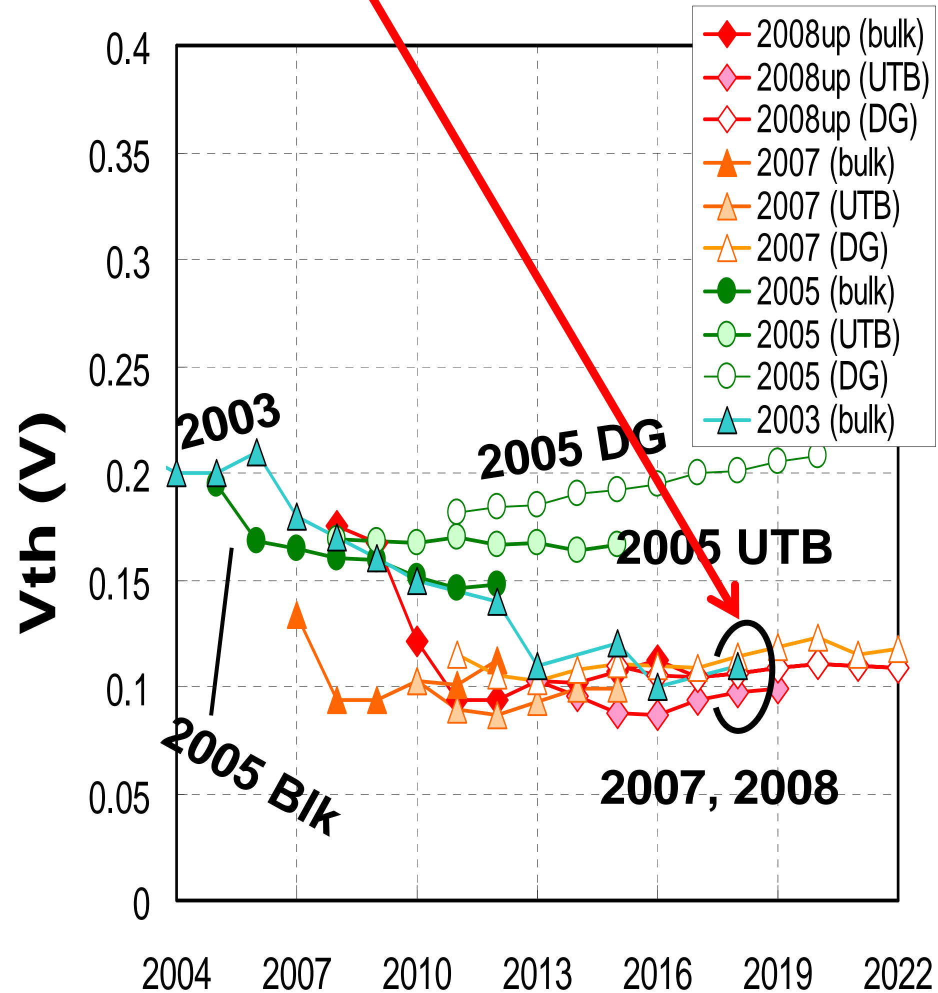
2008up (bulk)
2008up (UTB)
2008up (DG)
2007 (bulk)
2007 (UTB)
2007 (DG)
2005 (bulk)
2005 (UTB)
2005 (DG)
2003
2001
1999

Year

2008 Values are from ITRS Public Conf. and still under discussion

Vth-sat will be around 0.1V

Saturated Vth

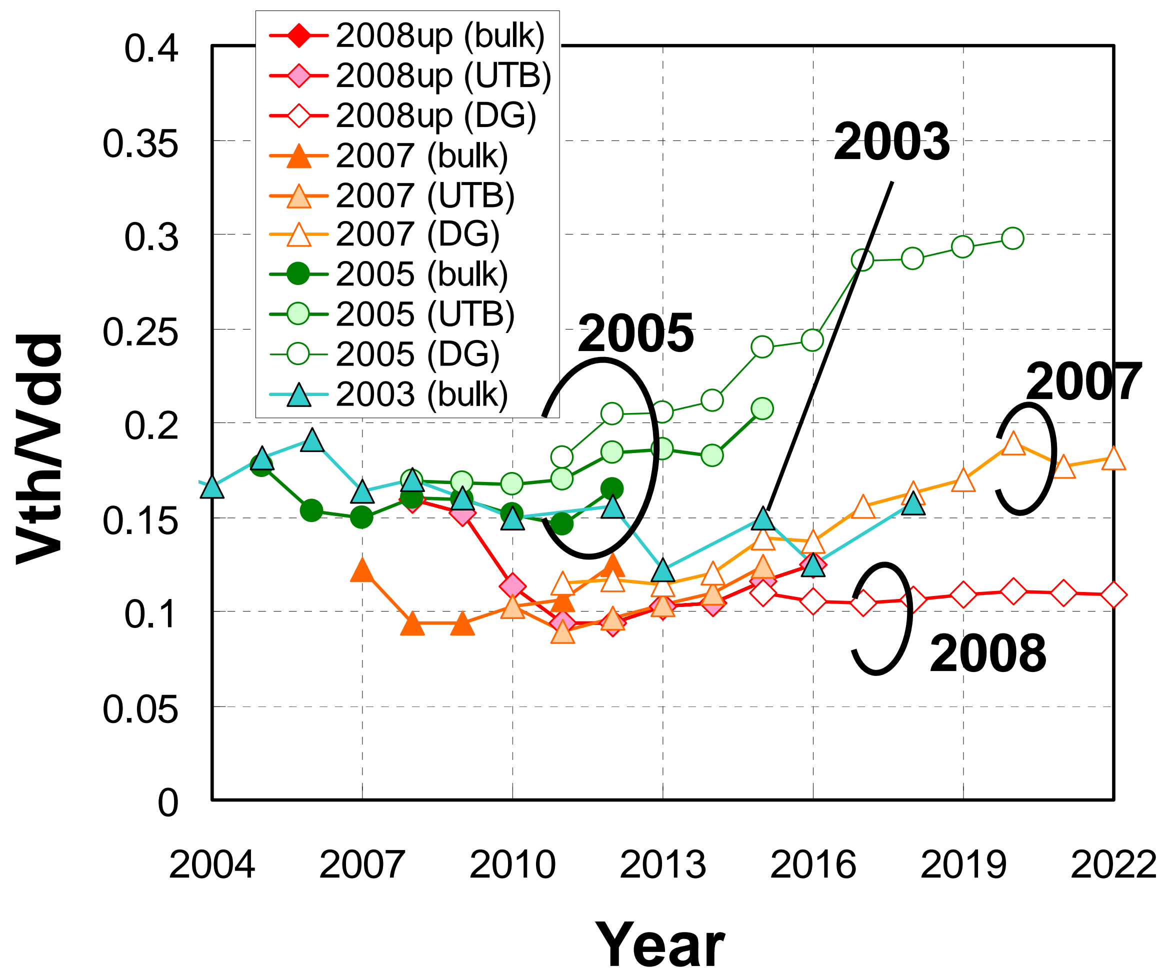


Year

Source: ITRS and 2008 ITRS Summer Public Conf.

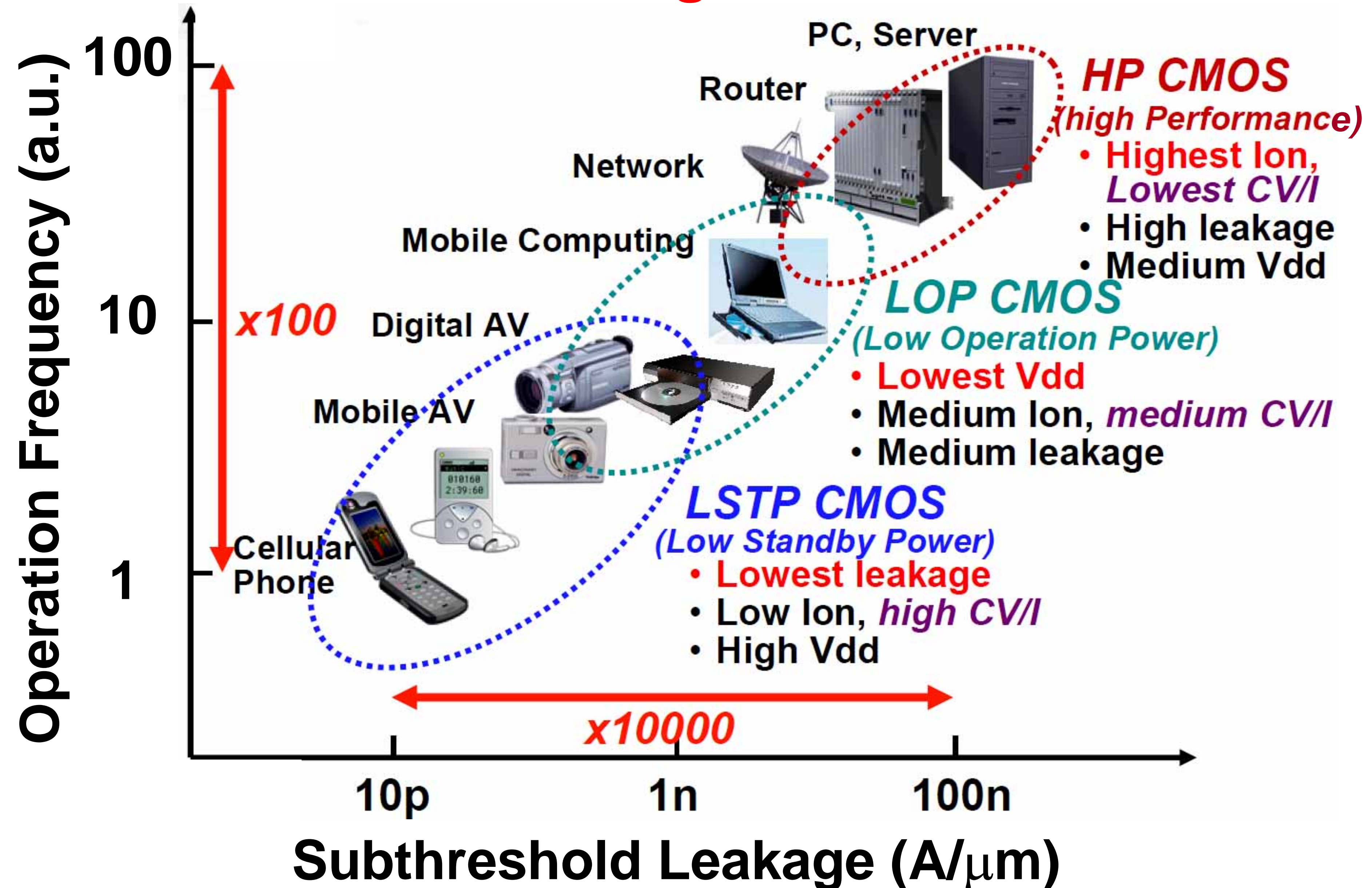
ITRS for HP logic

2008 Values are from ITRS Public Conf. and still under discussion



Source: ITRS and
2008 ITRS Summer Public Conf.

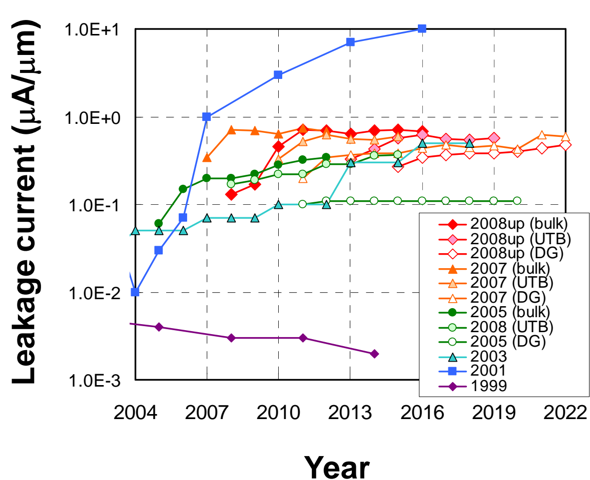
HP, LOP, LSTP for Logic CMOS



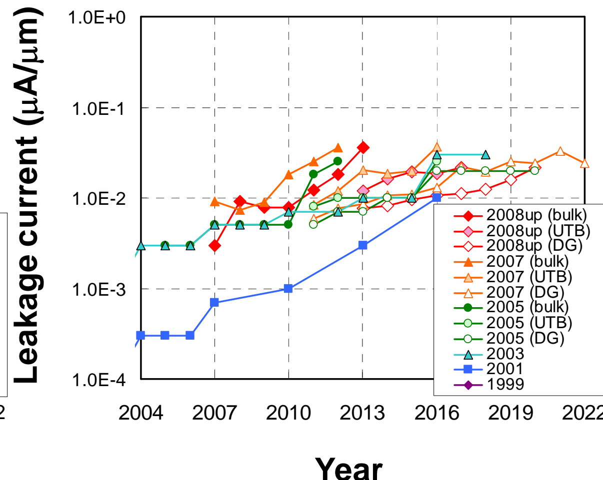
Source: 2007 ITRS Winter Public Conf.

Comparison of Isd-leak and Id-sat for HP, LOP, LSTP

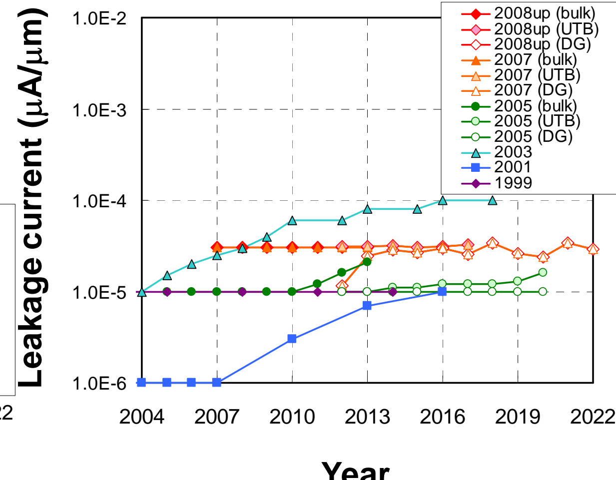
Isd_leak for HP



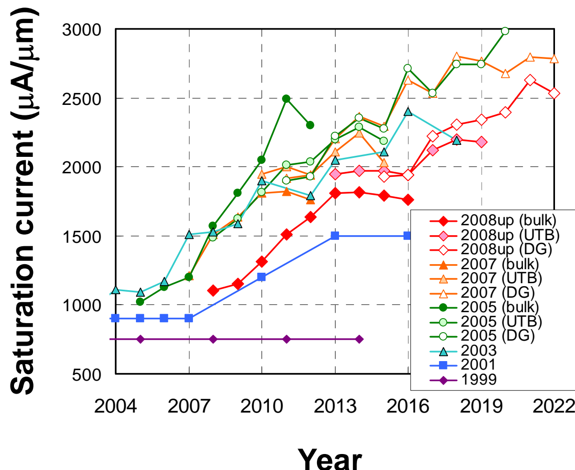
Isd_leak for LOP



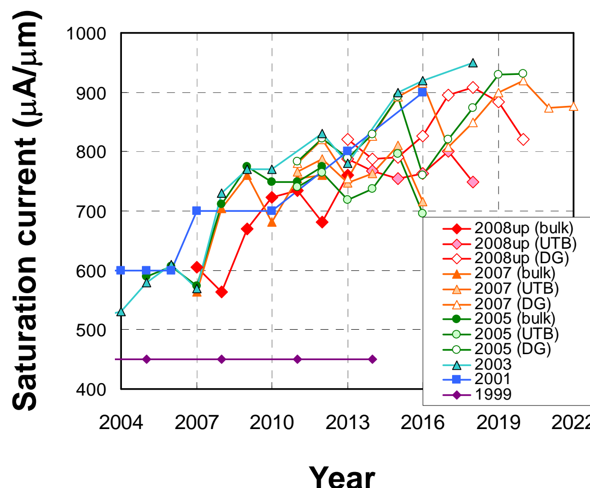
Isd_leak for LSTP



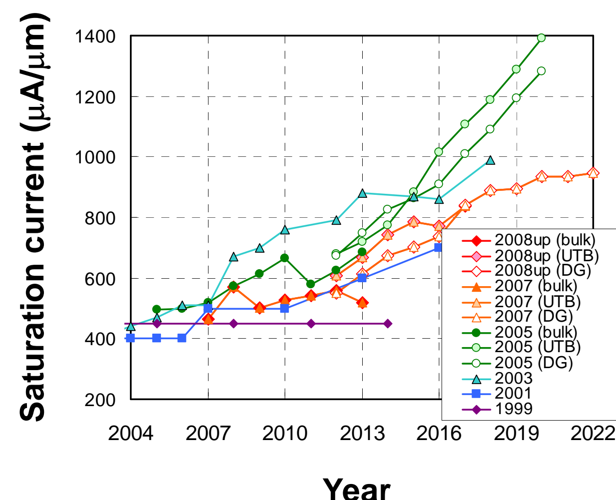
Id_sat for HP



Id_sat for LOP



Id_sat for LSTP

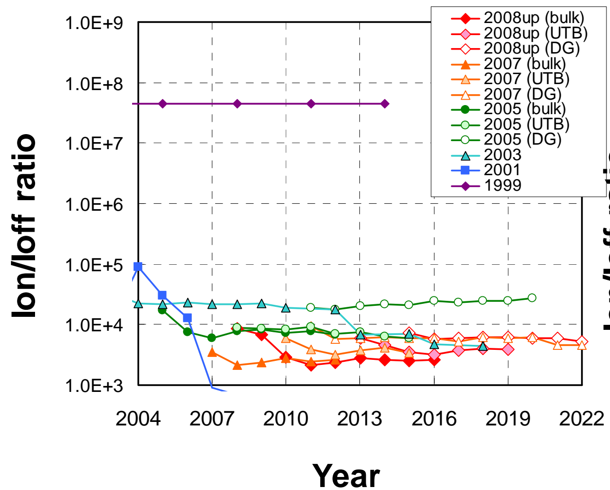


2008 Values are from ITRS Public Conf. and still under discussion

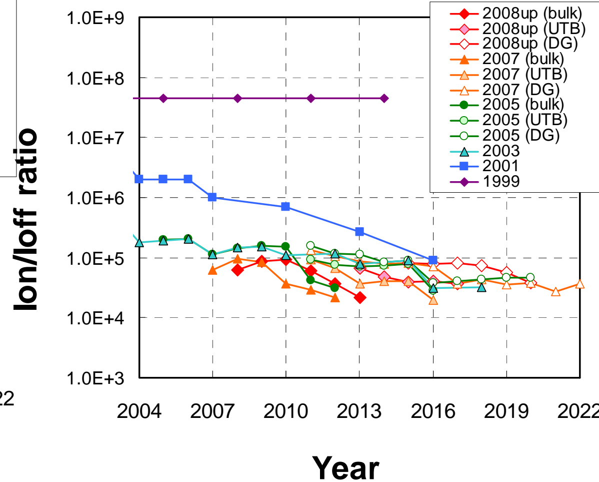
Source: ITRS and 2008 ITRS Summer Public Conf.

Comparison of Ion/Ioff for HP, LOP, LSTP

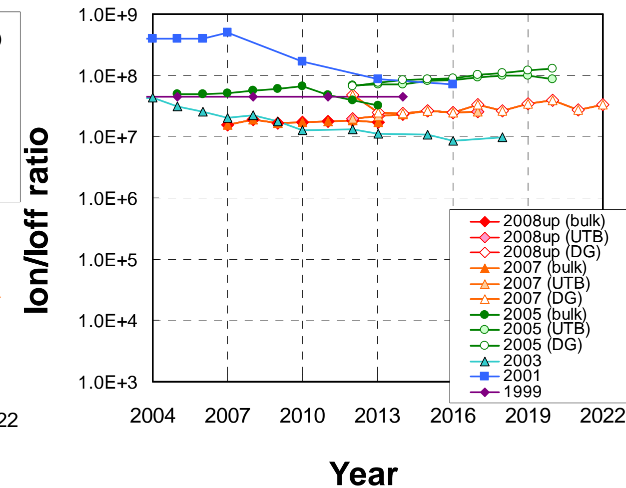
Ion/Ioff for HP



Ion/Ioff for LOP



Ion/Ioff for LSTP

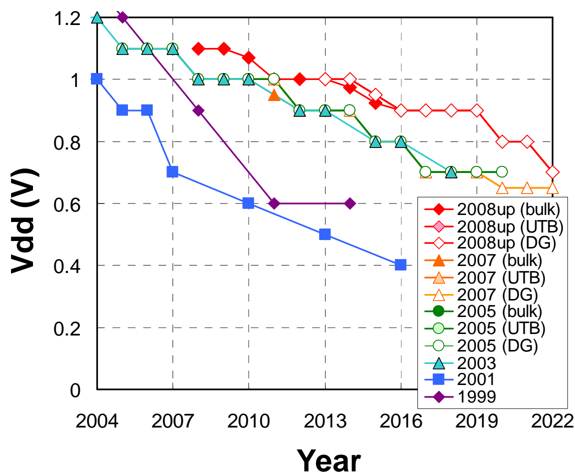


Source: ITRS and
2008 ITRS Summer Public Conf.

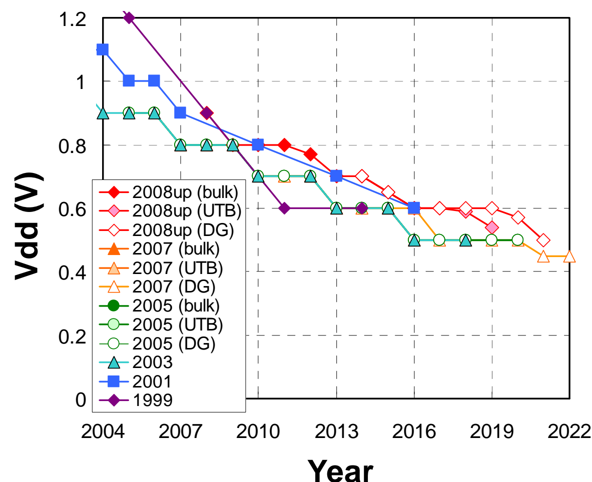
**2008 Values are from ITRS Public Conf.
and still under discussion**

Comparison of Vdd and Vth-sat for HP, LOP, LSTP

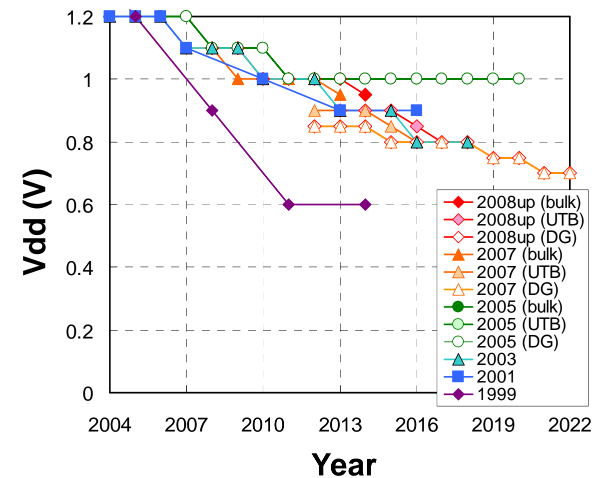
Vdd for HP



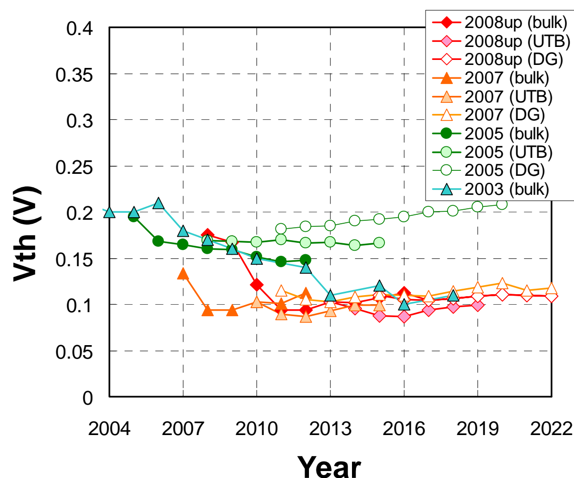
Vdd for LOP



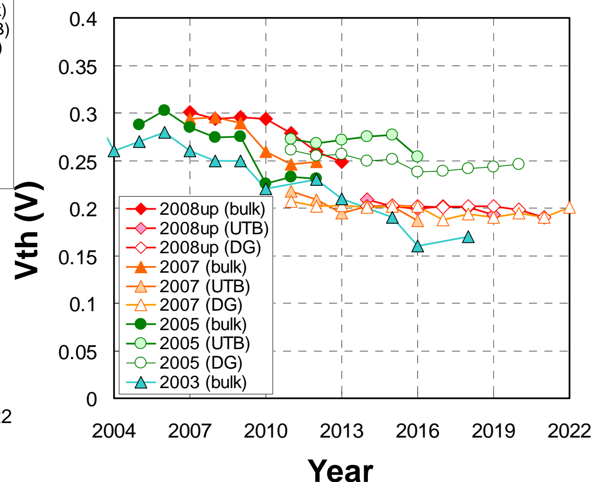
Vdd for LSTP



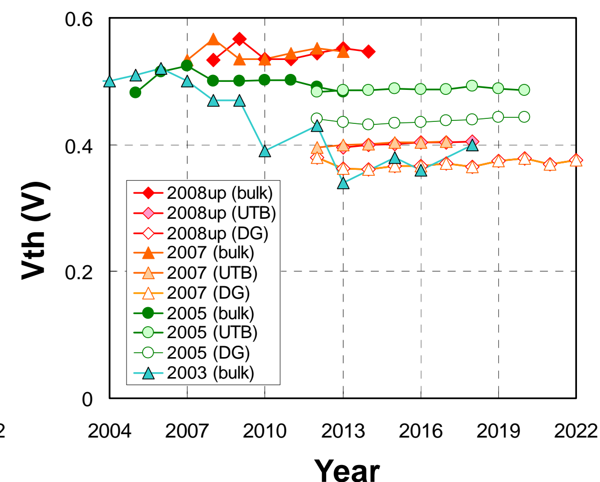
Vth for HP



Vth for LOP



Vth for LSTP

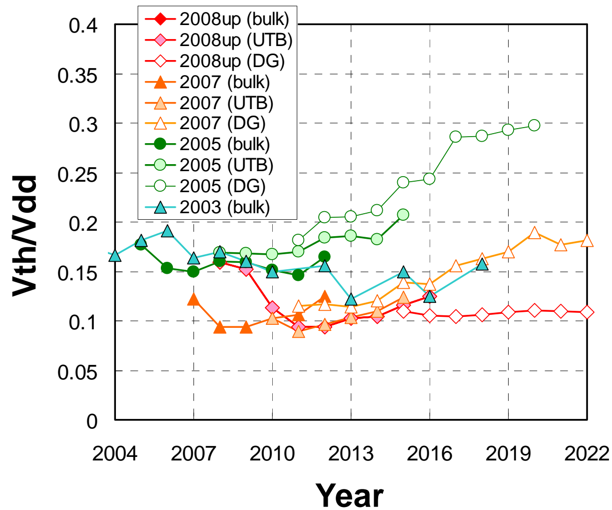


2008 Values are from ITRS Public Conf. and still under discussion

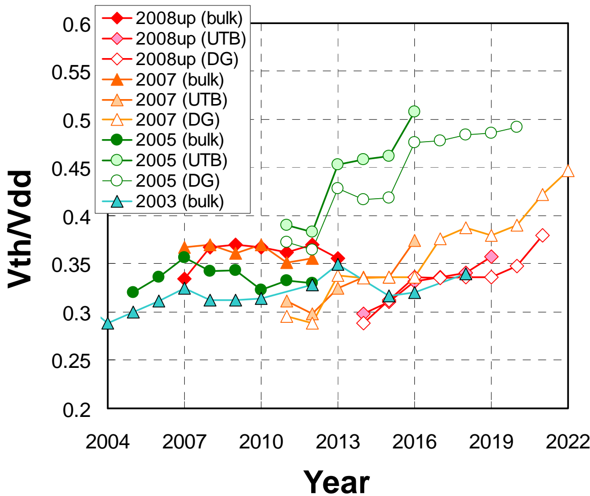
Source: ITRS and
2008 ITRS Summer Public Conf.

Comparison of Vth-sat/Vdd for HP, LOP, LSTP

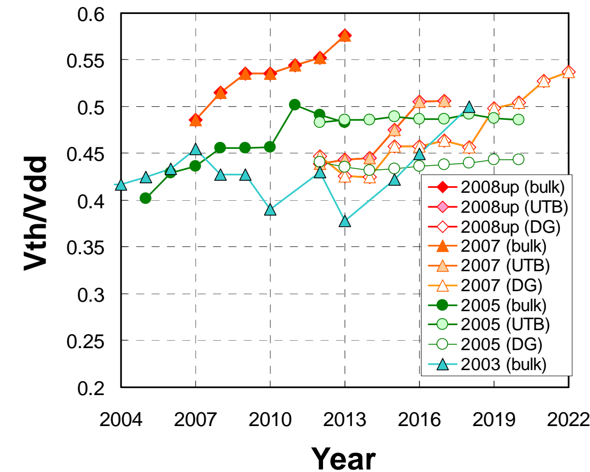
Vth/Vdd for HP



Vth/Vdd for LOP



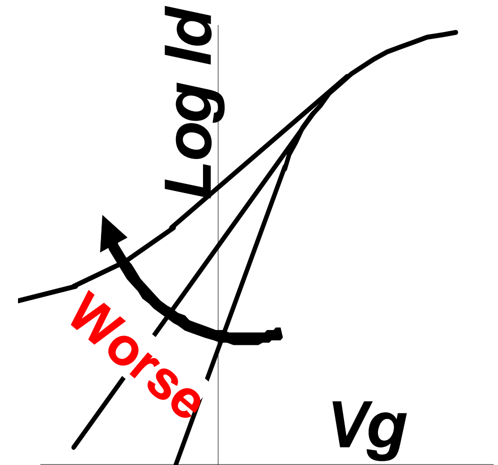
Vth/Vdd for LSTP



Source: ITRS and
2008 ITRS Summer Public Conf.

2008 Values are from ITRS Public Conf.
and still under discussion

SS (Subtheshold Slope) becomes worse in the following cases



1. Improper down-scaling

Ex. When T_{ox}, W_{dep}, or V_{dd} is not scaled

2. High impurity doping in channel or substrate

High impurity Conc.

→ C_D increase

→ SS increase

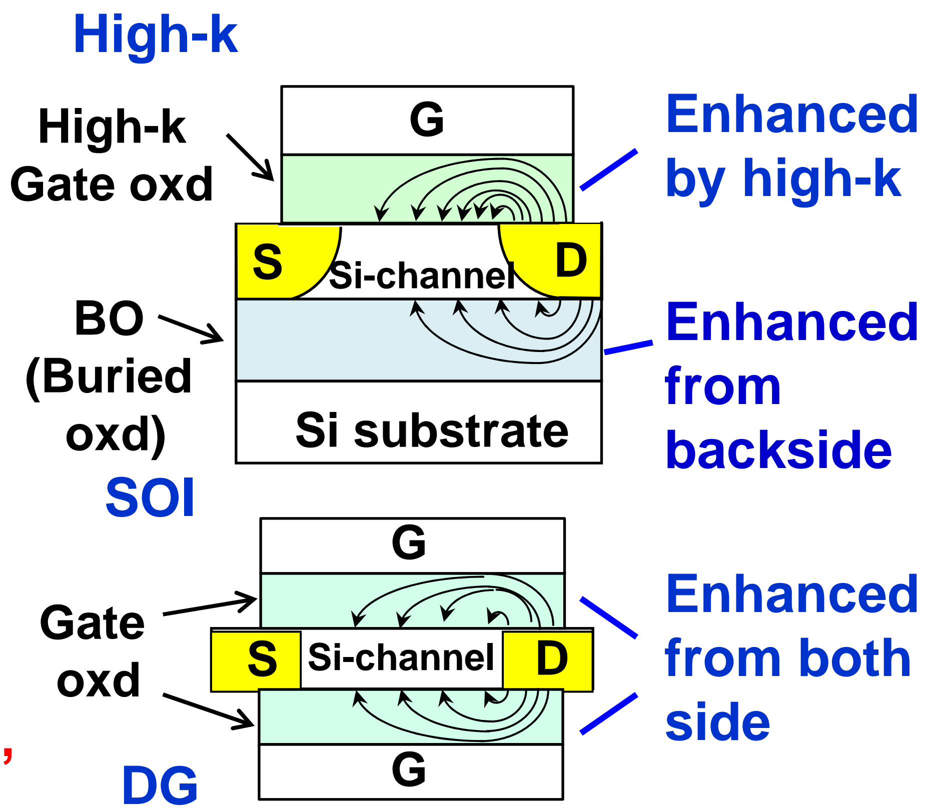
$SS = (Ln10)(kT/q)(C_{ox}+C_D+C_{it})/C_{ox}$

3. Enhanced Drain-Electric-field penetration through oxide

Ex. High-k, SOI,

Multi-gate (Double gate: DG)

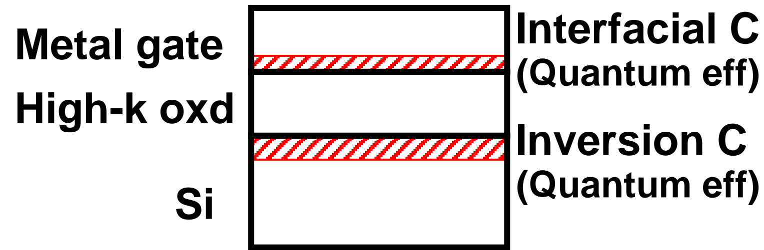
DG and SOI often show better SS, but be careful!

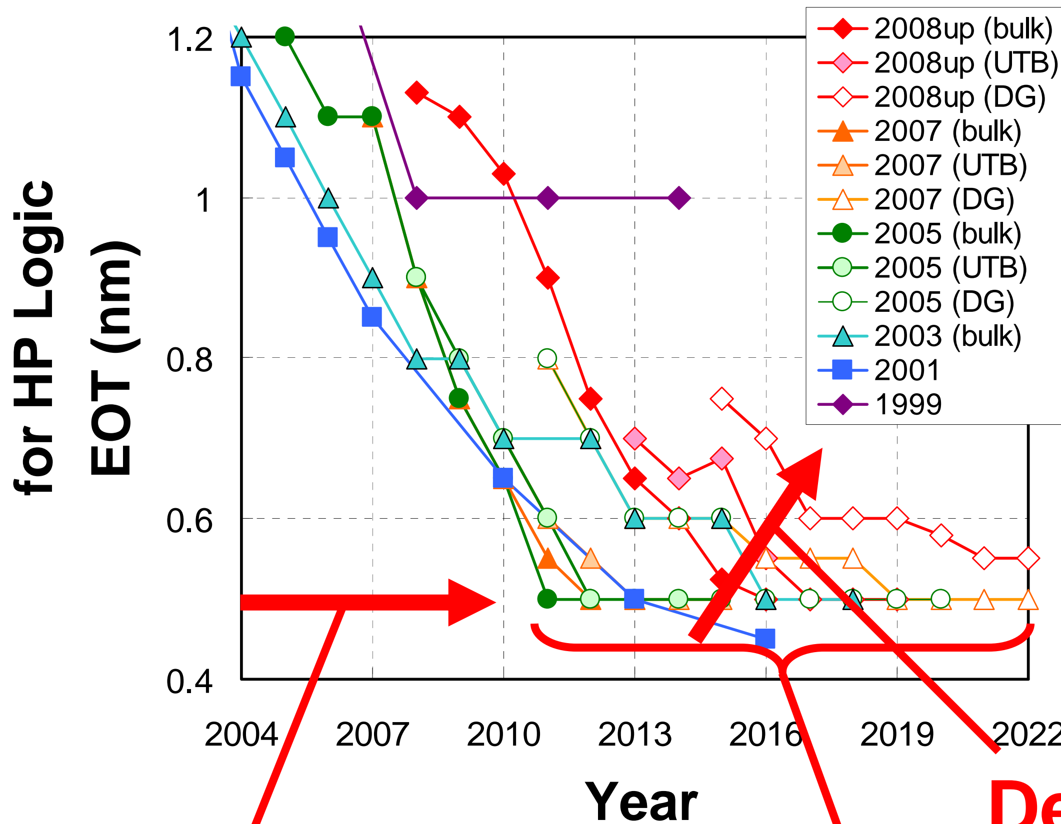


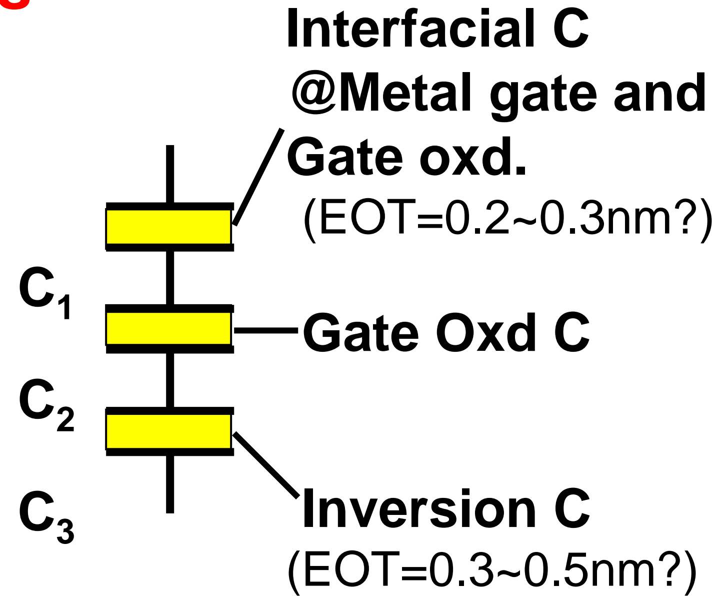
Improper down-scaling

Could we squeeze technologies for ultimate CMOS scaling?

Saturation of EOT thinning is a serious roadblock to proper down-scaling.







$EOT(C_1) + EOT(C_3) > 0.5nm$

Small effect to decrease $EOT(C_2)$ beyond 0.5nm?

Is 0.5nm real limit?

EOT<0.5nm with Gain in Drive Current is Possible

La_2O_3 gate insulator

(a) EOT=0.37nm** V_{th}=-0.04V

W/L=2.5/50μm
PMA 300ºC (30min)

insufficient* compensation region

Drain current (mA)

Drain voltage (V)

(b) EOT=0.43nm** V_{th}=-0.03V

4%up

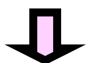
Drain voltage (V)

(c) EOT=0.48nm V_{th}=-0.02V

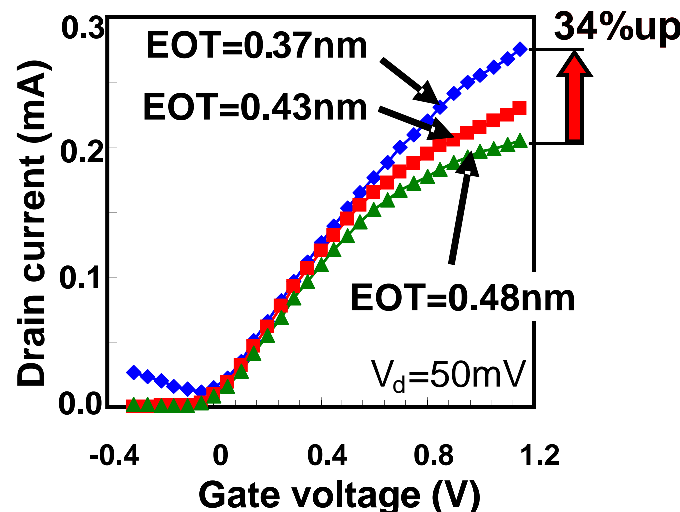
14%up

Drain voltage (V)

EOT scaling below 0.5nm



Still useful for larger drain current



Source: K. Kakushima, K. Okamoto, K. Tachi, P. Ahmet, K. Tsutsui, N.i Sugii, T. Hattori, and H. Iwai, IWDTF 2008, Tokyo, November, 2008

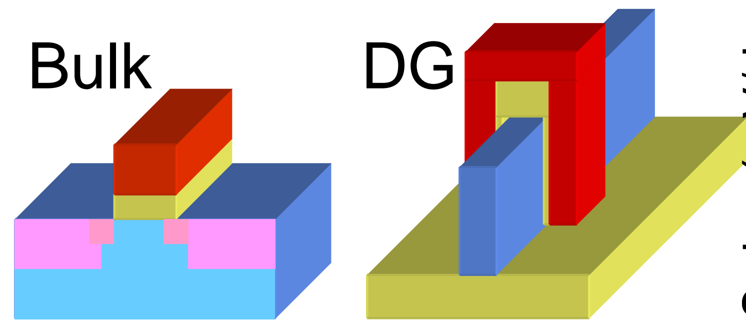
* Because Lg is very large (2.5μm), gate leakage is large in case (a). The gate leakage component was subtracted from measured data for case (a). However, if we make small gate length, the gate leakage current should become sufficiently small to be ignored compared with Id as we verified with SiO_2 gate before (Momose et al.,IEDM 1994). The gate leakage could be suppressed by modifying material and process in future.

** Estimated by Id value

Thus, in future, maybe continuous development of new techniques could make more proper down-scaling possible.

It is difficult to say, but EOT and Vdd may become smaller than expected today.

Enhanced D-Electric-field



Same parameter condition for both

(2006 ITRS Bulk parameters are used for both Bulk and DG)

Lg=16nm, tox(EOT)=0.5nm, Dopant@Channel=8.1X10[18]cm[-2]

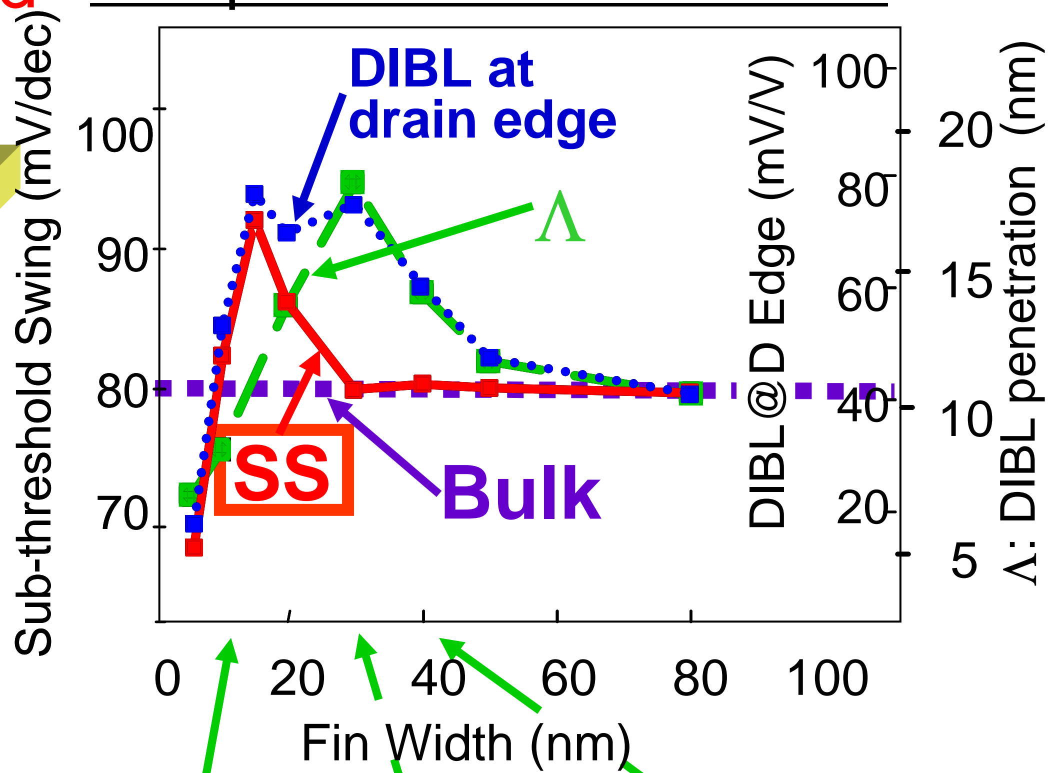
Source: ECS Fall Meeting, Oct 2008, Honolulu, Y. Kobayashi, A. B. Sachid, K. Tsutsui, K. Kakushima, P. Ahmet, V. Ramgopal Rao and H. Iwai.

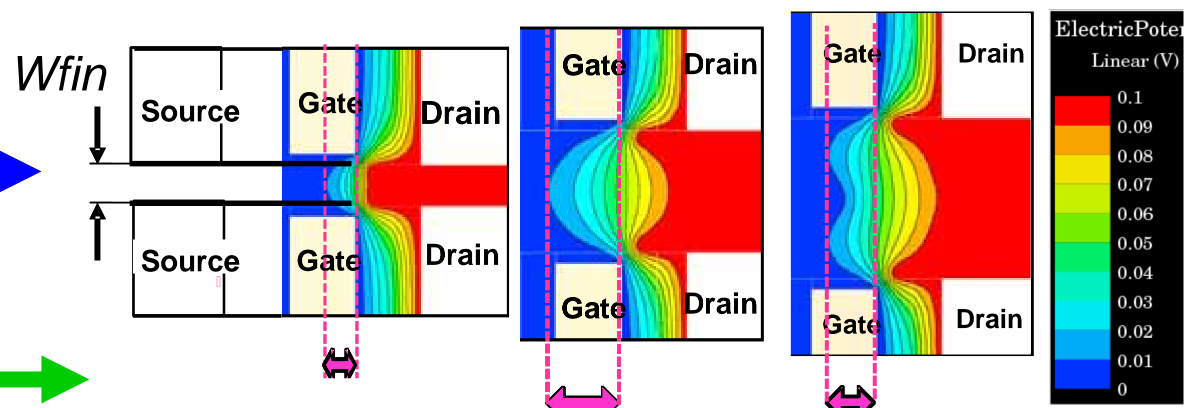
DIBL: Drain Induced Barrier Lowering

$$\left.\frac{\partial V(x,y)}{\partial V_d}\right|_{V_d=1V}$$

Λ: Penetration Depth of DIBL

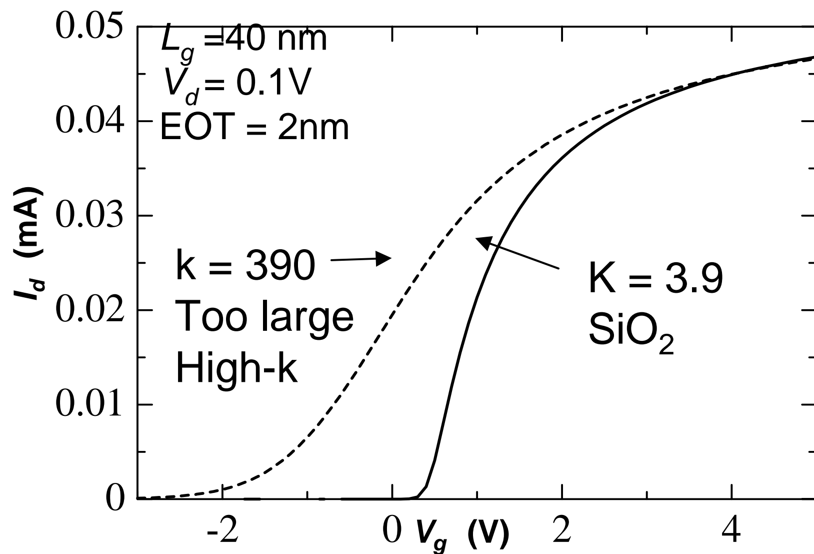
Comparison of Bulk and DG

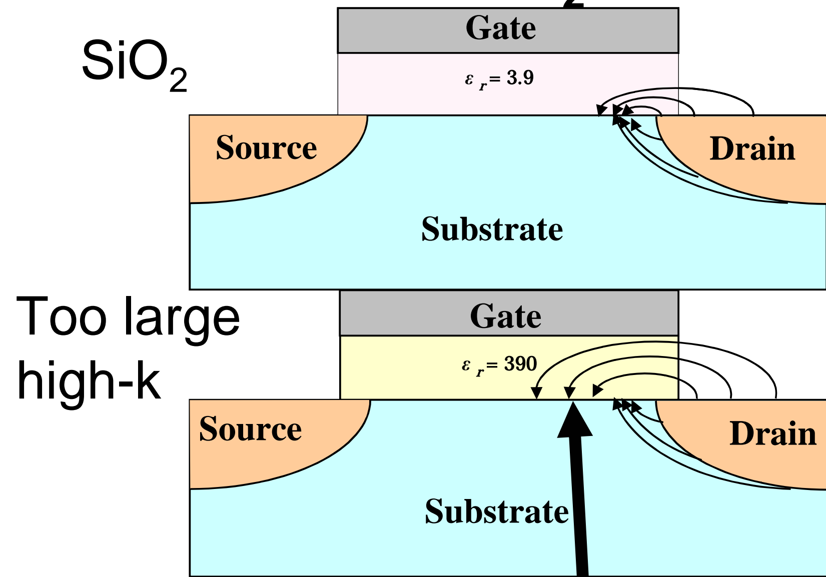


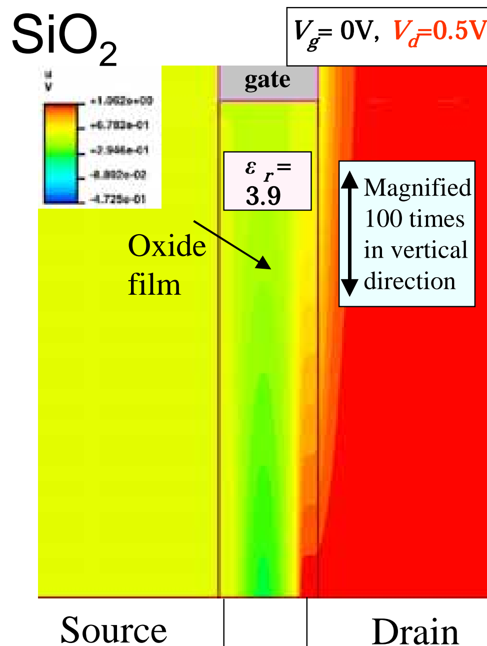


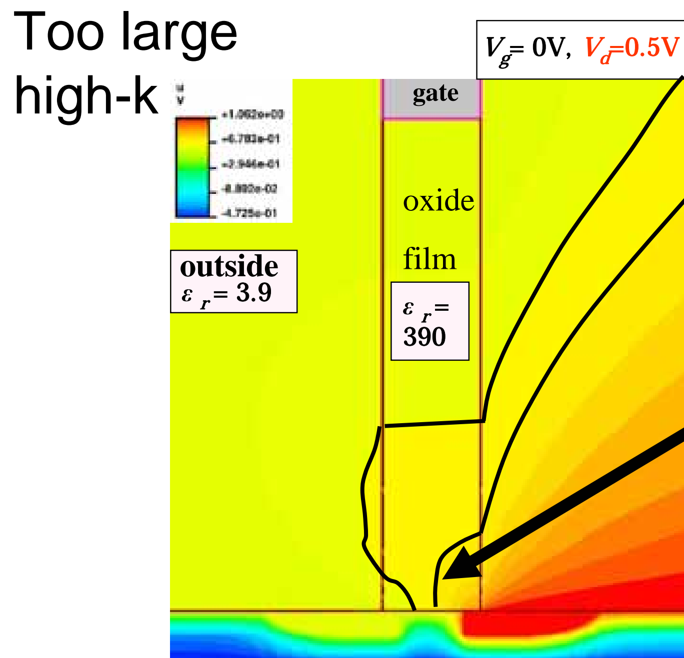
Enhanced D-Electric-field

Comparison of High-k and SiO_2 MOSFETs









Penetration of lateral field from Drain through high-k causes significant short channel effects

R. Fujimura, M. Takeda, K. Sato, S. Ohmi, H. Ishiwara, and H. Iwai, ECS Symp. on ULSI Process Integration II, Volume 2001-2, pp.313-323, 2001,

V_{dd} will stay higher than predicted by previous ITRS roadmaps.

Solution towards Low V_{dd}

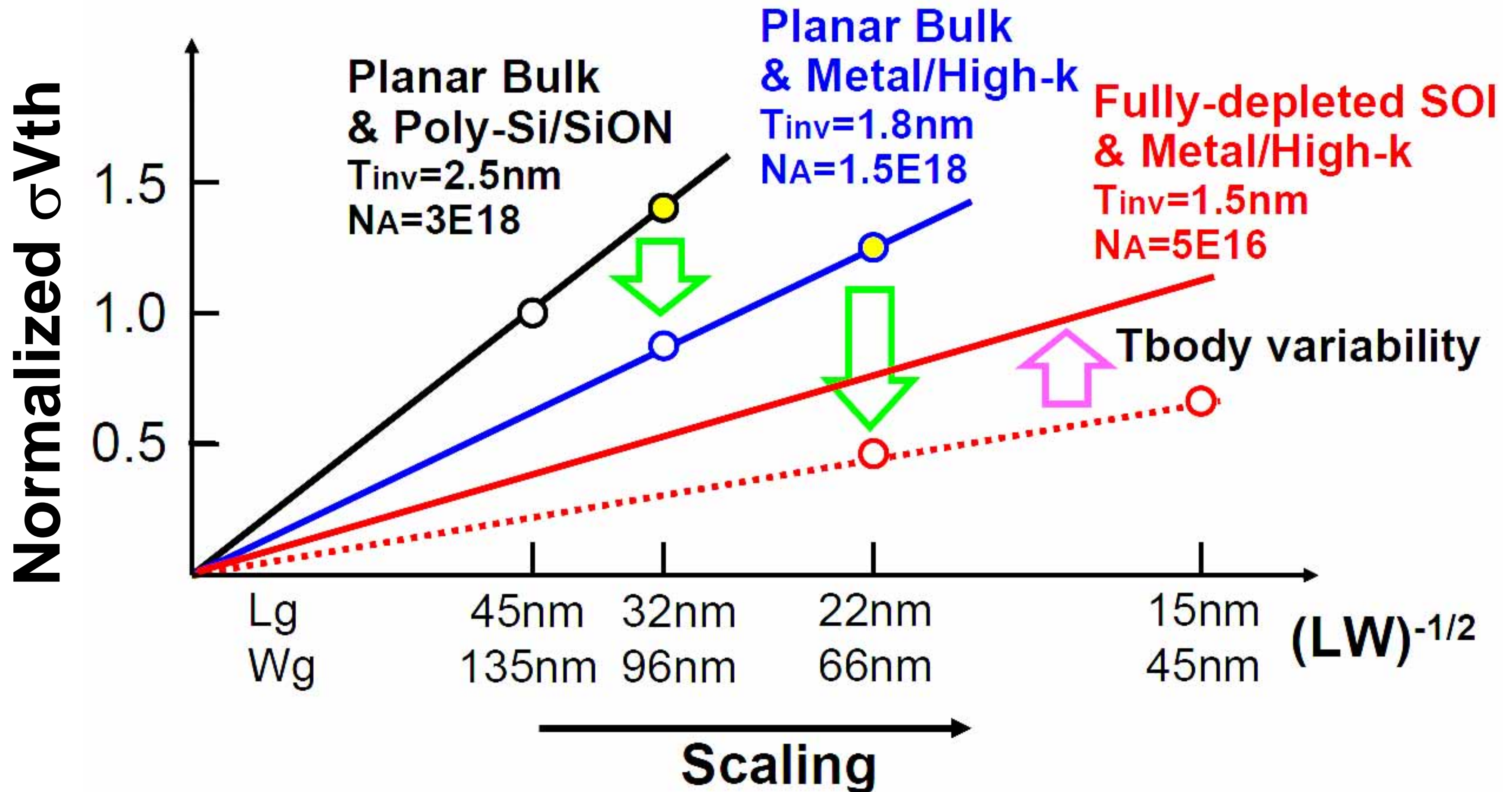
Effort to reduce $I_{sd\text{-}leak}$ and increase $I_{d\text{-}sat}$ is important

- Scaling: Proper down-scaling
  - Introduction of Next generation high-k, S/D etc.
  - CD* variation control by lithography and etching techniques

*** CD: Critical dimension**

- Structure: Bulk → UTB-SOI → DG → Nanowire
- Variation: Proper scaling by new tech. – High-k, litho. Etc.
  V_{th} adjustment by V_{sub} control
- Circuit techniques: Dynamic and local Multi-V_{dd}, etc.

Random Variability Reduction Scenario in ITRS 2007



Source: 2007 ITRS Winter Public Conf.

Another concern for Low V_{dd} besides $I_{sd\text{-}leak}$ increase

→ Huge power loss for voltage conversion to such low V_{dd}

4. SRAM cell scaling

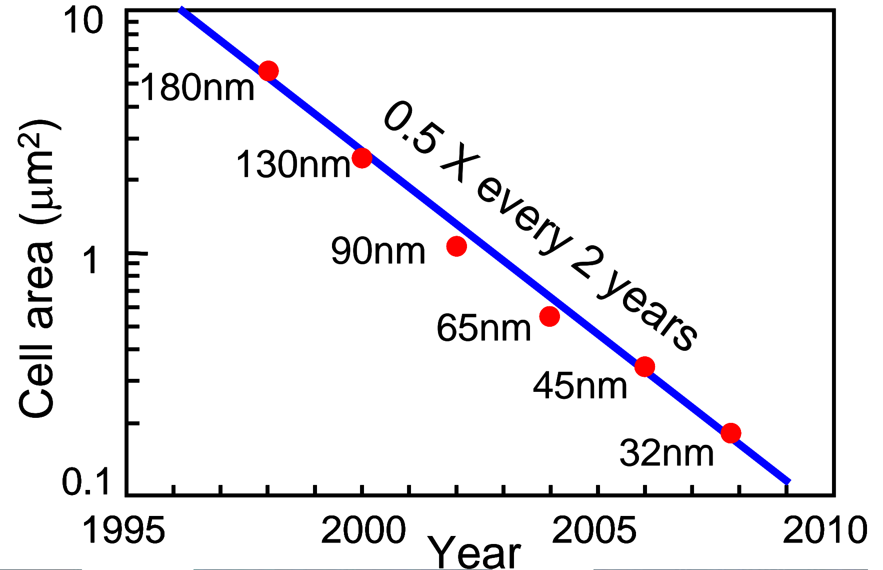
Intel's **SRAM** test chip trend

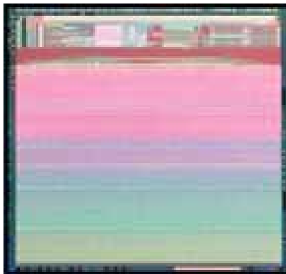
Source: B. Krzanich, S. Natrajan, Intel Developer's Forum 2007
http://download.intel.com/pressroom/kits/events/idffall_2007/Briefing
Silicon&TechManufacturing.pdf

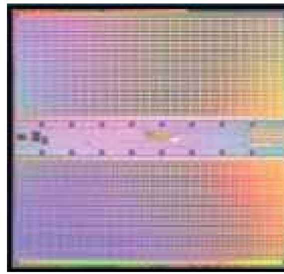
SRAM down-scaling trend has been kept until 32nm and probably so to 22nm

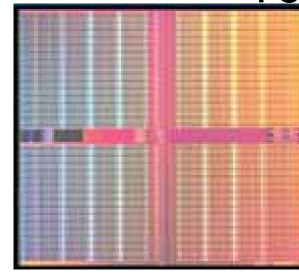
Process name	Lithography	1st production
P1264	65nm	2005
P1266	45nm	2007
P1268	32nm	2009
P1270	22nm	2011

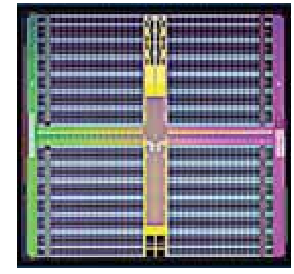
Only schedule has been published











Technology	**90 nm Process**	**65 nm Process**	**45 nm Process**	**32 nm Process**
Cell size	**1.0 μm²cell**	**0.57 μm²cell**	**0.346 μm²cell**	**0.182 μm²cell**
Capacity	50 Mbit	70 Mbit	153 Mbit	291 Mbit
Chip area	109 mm²	110 mm²	119 mm²	118 mm²
Functional Si	February '02	April '04	January '06	September '07

22 nm technology 6T SRAM Cell: Size = 0.1μm

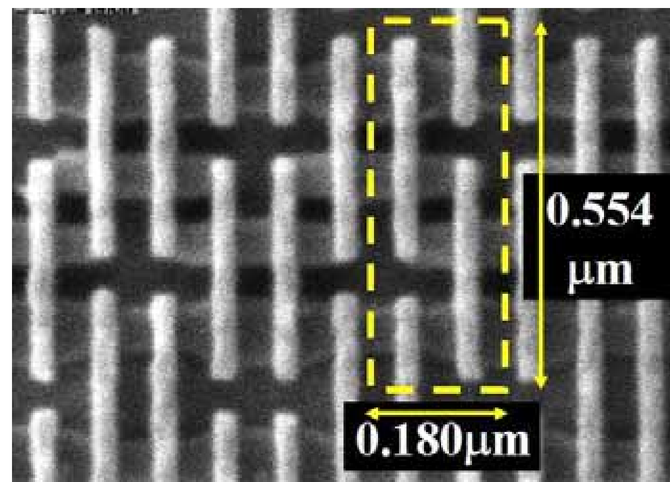
Source: http://www-03.ibm.com/press/us/en/pressrelease/24942.wss

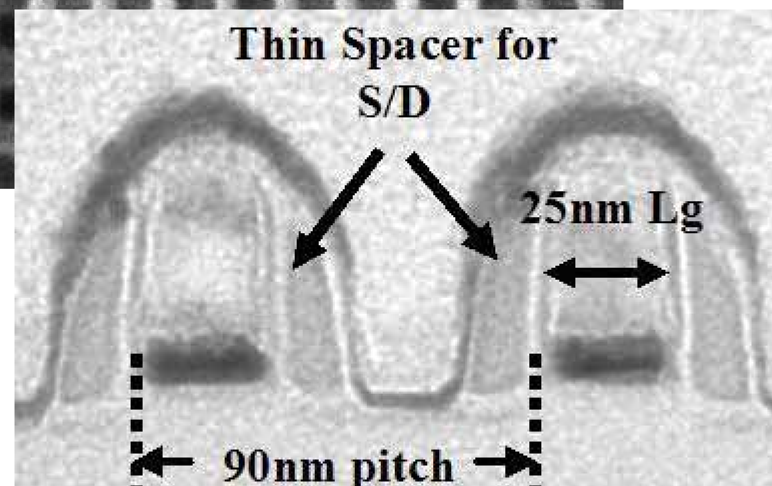
Announced on Aug 18, 2008

Consortium: IBM (NYSE) , AMD, Freescale, STMicroelectronics, Toshiba and the College of Nanoscale Science and Engineering (CNSE)

0.1μm cell size is almost on the down-scaling trend

Static noise margin of 220 mV at 0.9 V





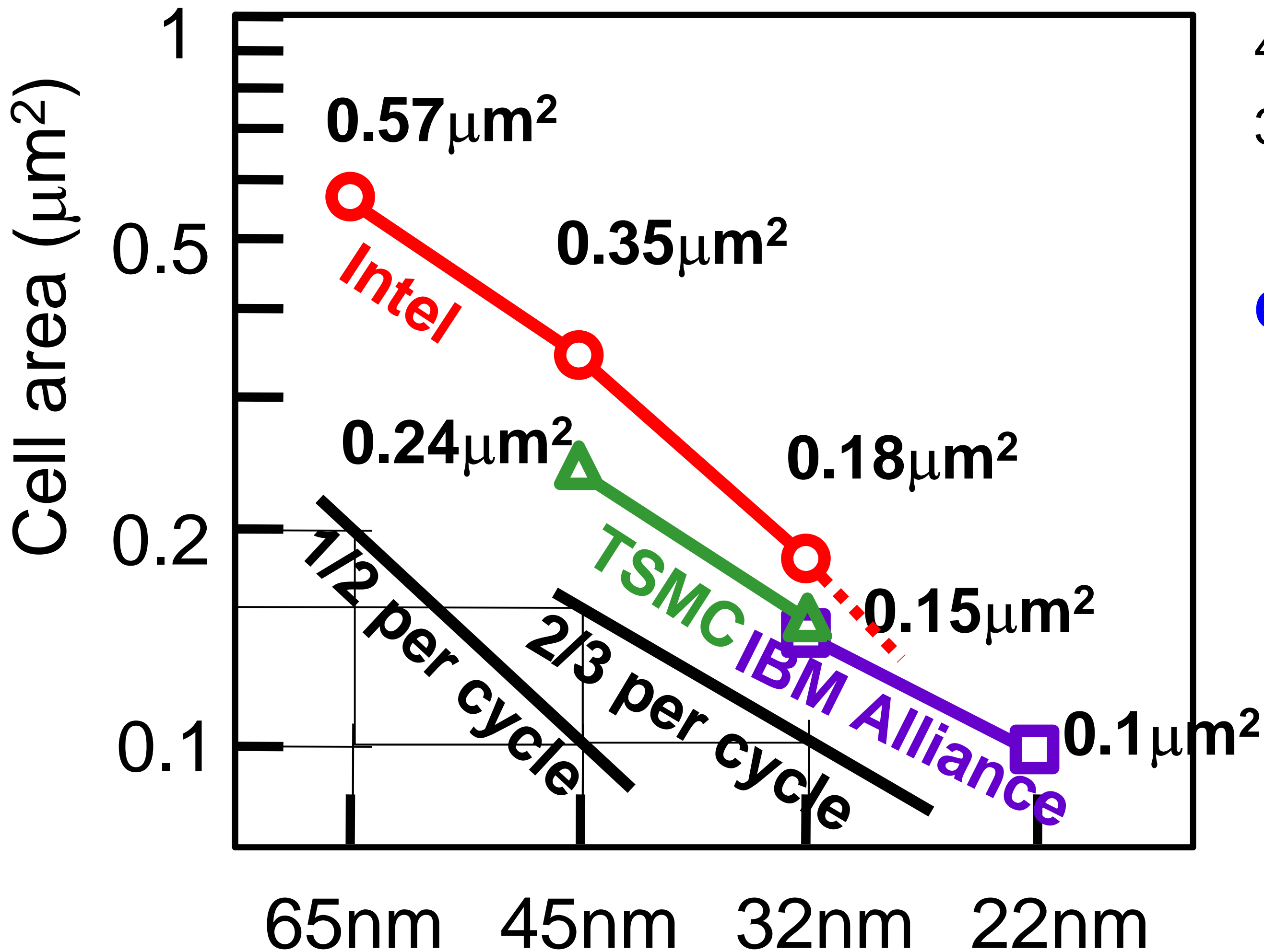
New technologies introduced

- High-NA immersion lithography
- High-K metal gate stacks
- **25 nm gate lengths**
- Thin composite oxide-nitride spacers
- Advanced activation techniques
- Extremely thin silicide
- Damascene copper contacts

Source: IEDM2008 Pre-conference Publicity http://www.btbmarketing.com/iedm/

Cell size reduction trends

1/2 or 2/3 per cycle?



Intel

Functional Si

65nm Apr.2004

45nm Jan.2006

32nm Sep.2007

TSMC

Conference (IEDM)

45nm Dec.2007

32nm Dec.2007

IBM Alliance (Consortium)

Conference (IEDM)

32nm Dec.2007

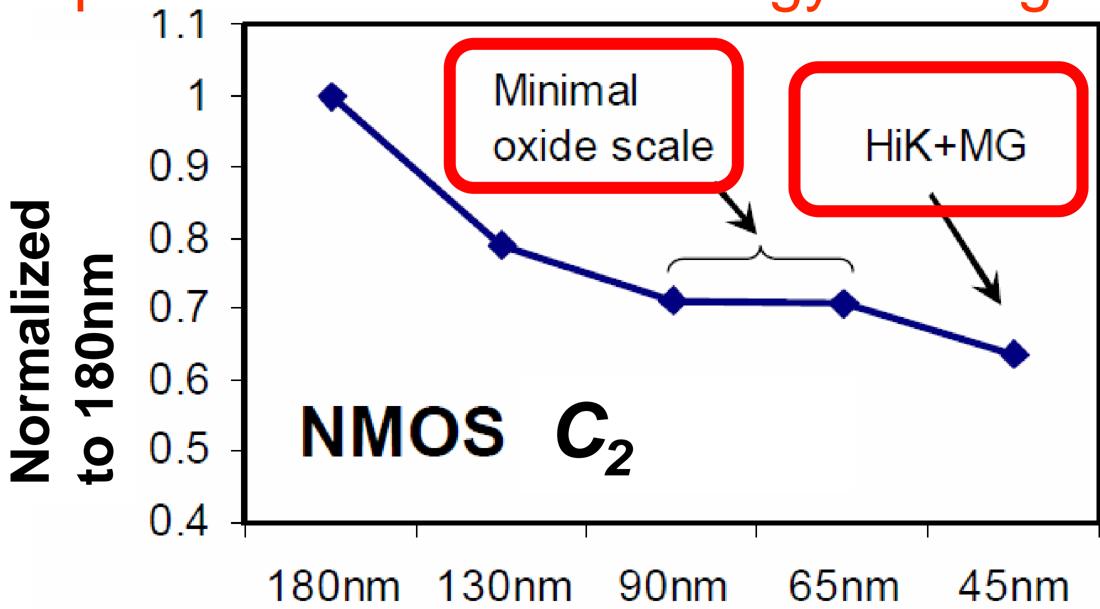
Press release

22nm Aug.2008

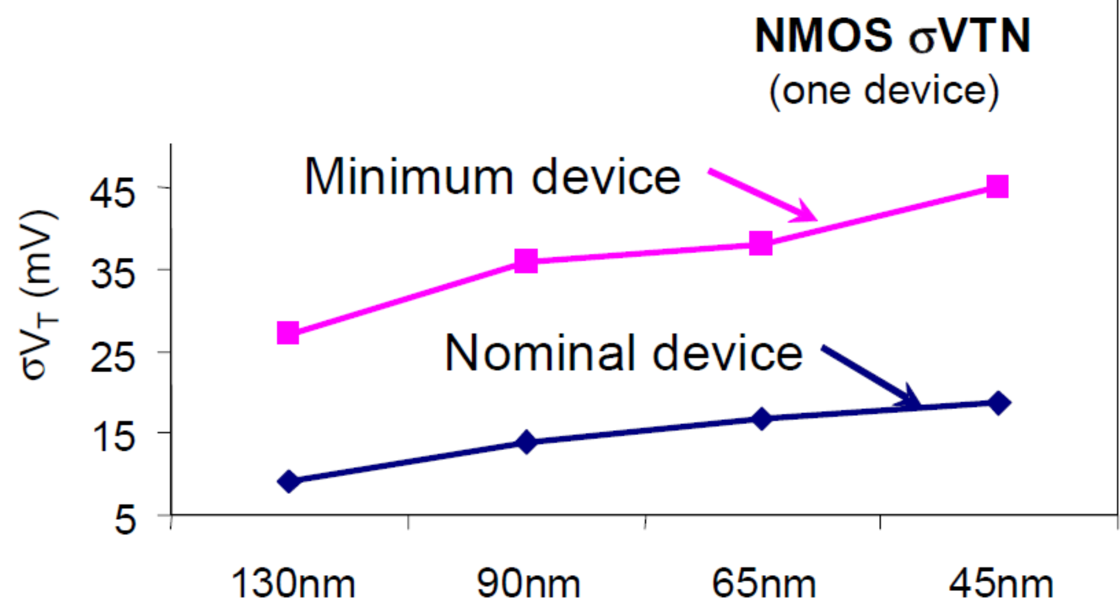
NMOS Mismatch Coefficient (C_2) improvement with technology scaling

$$\sigma V_{Tran} = \left(\frac{\sqrt[4]{4q^3\varepsilon_{si}\phi_B}}{2}\right)\cdot\frac{T_{ox}}{\varepsilon_{ox}}\cdot\left(\frac{\sqrt[4]{N}}{\sqrt{Weff\cdot Leff}}\right)$$

$$= \frac{1}{\sqrt{2}}\left(\frac{c_2}{\sqrt{Weff\cdot Leff}}\right)$$



σV_{Tran}

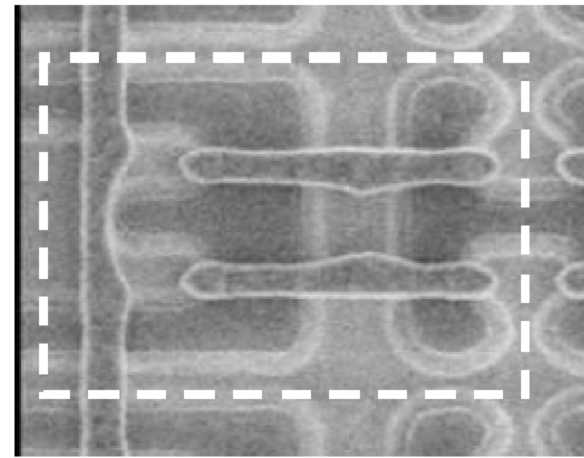


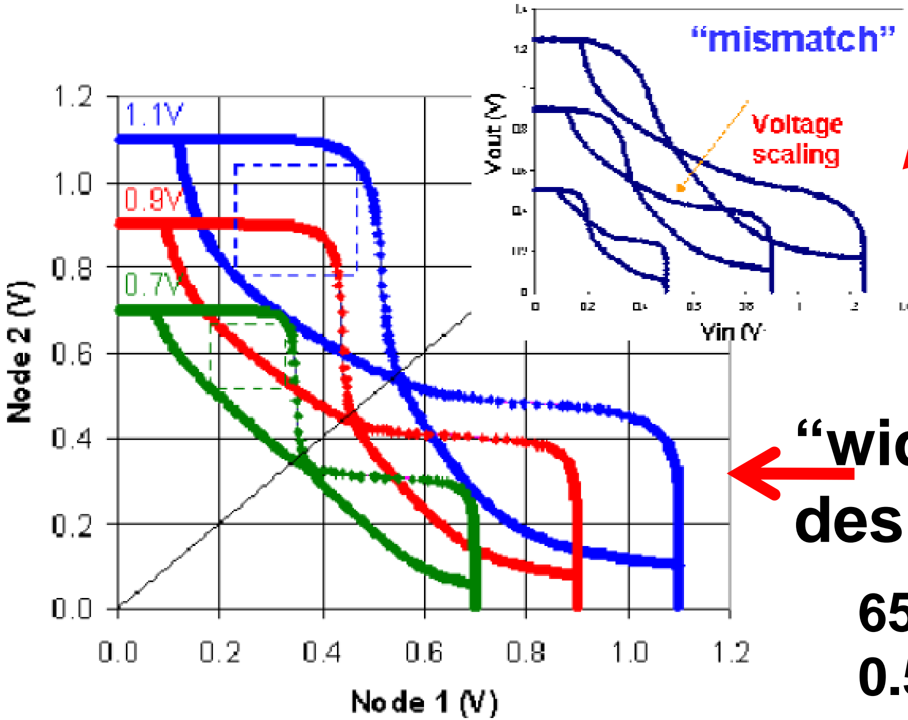
Source: K.J.Kuhn IEDM 2007

Mismatch improvement by layout (Intel)

"tall" design

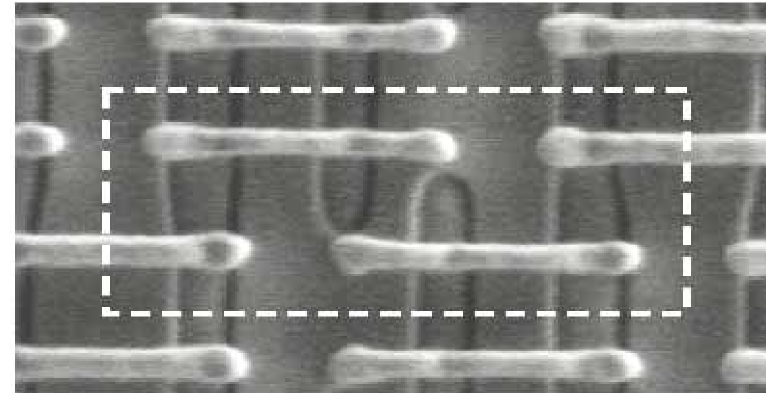
90nm :1.0 μm²





"wide" design

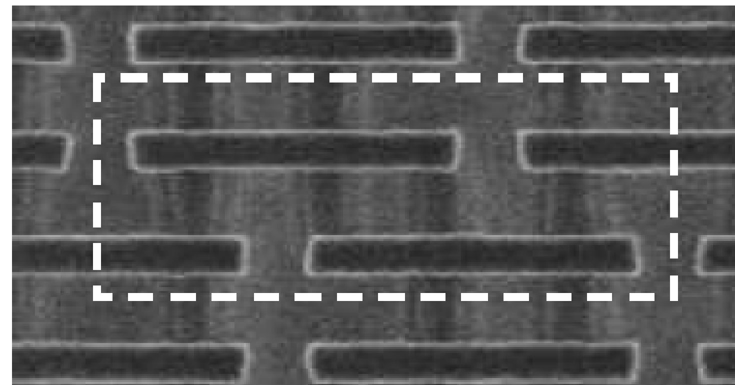
65nm : 0.57 μm²



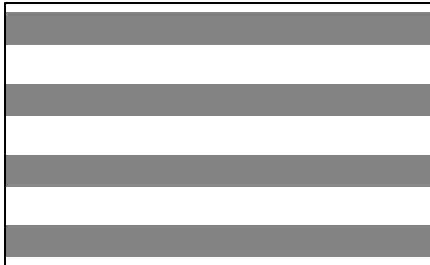
Source: K. J. Kuhn
IEDM2007 Tech. Dig. pp.471

"wide" design (**Square endcaps**)

45nm 0.346 μm²

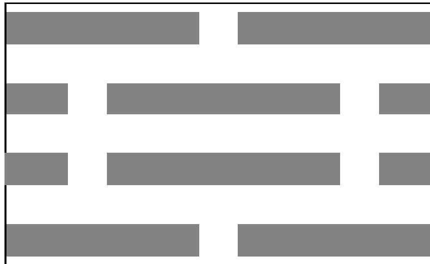


Double patterning for square endcap

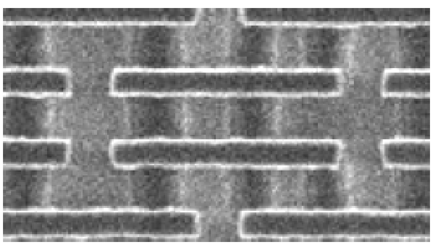


a) Pattern gate lines/spaces

b) Pattern cut mask



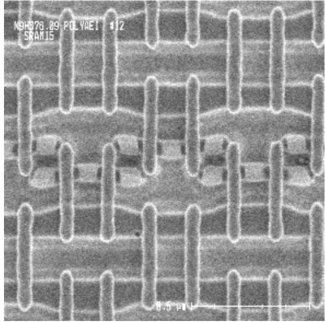
c) Final gate pattern



d) Intel 45nm SRAM cell

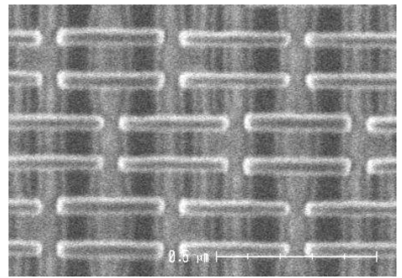
Source: M. Bohr, ICSICT2008

Cell evolution is similar



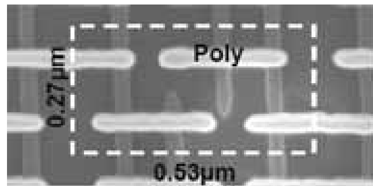
TSMC 45nm

IEDM 2007



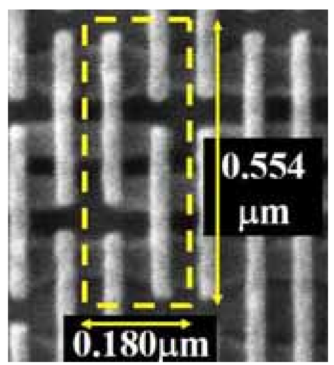
TSMC 32nm

IEDM 2007



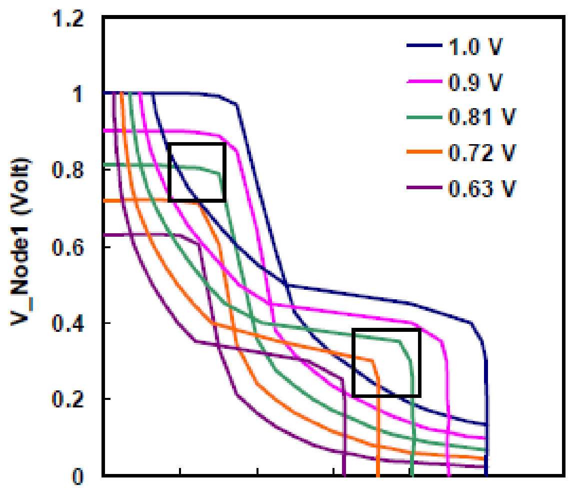
IBM Alliance 32nm

IEDM 2004

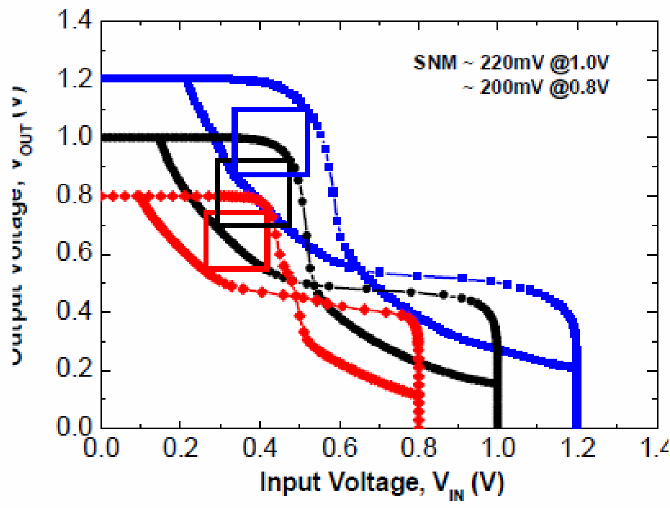


IBM Alliance 22nm

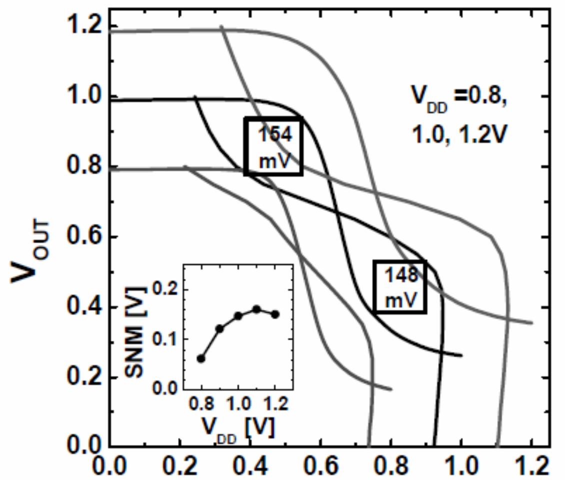
IEDM 2008



TSMC 45nm



TSMC 32nm



IBM Gr. 32nm

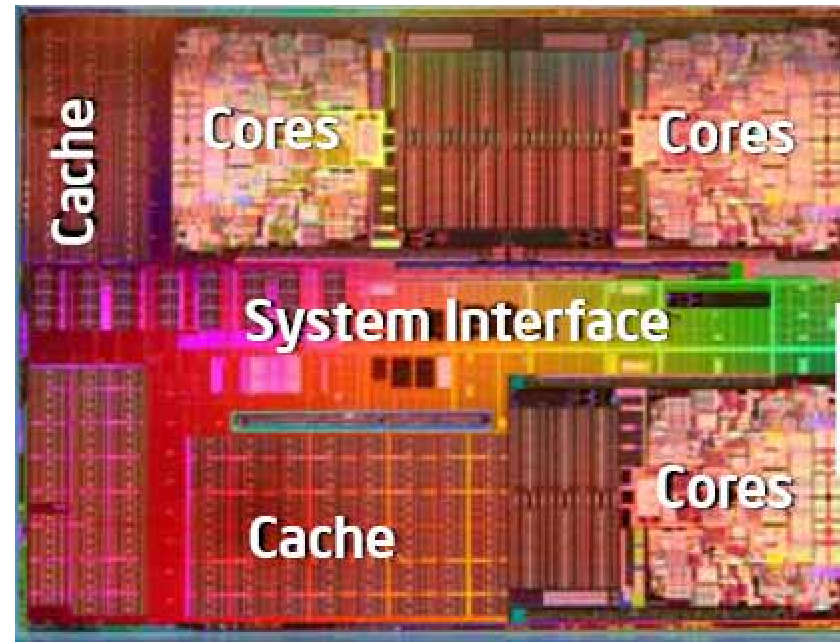
Most Difficult part of SRAM down-scaling is Vdd down-scaling

Density of on-chip cache SRAM memory is high and thus, Vth cannot be down-scaled too much because of large Isd-leak

Also, under low Vdd, read- and write margin degrades, data retention degrade.

Thus, Vdd down-scaling is more severe in SRAM than logic part of the circuits

Intel® Xeon® 7400 Series (Dunnington)



45 nm high-k6 cores

16MB shared L3 cache

Source: Intel Developer Forum 2008

Cache occupies huge area

→ Cell size of SRAM should be minimized

→ Isd-leak should be minimized

→ Vth are often designed to be higher than Min. logic Vth

→ Lg are often designed to be larger than Min. logic Lg

Future Directions For Improving Vmin

• Application
– Improvement in voltage and temperature tolerance
• Package
– Separated array / logic voltage to minimize logic noise effect on SRAM
• Design
– Higher array VDD and improved on-chip supply robustness
– Increased redundancy
– Improved timings
– Cells per BL hierarchical BL structure
– Write/Read assist and sense-amp design
• Cell and Process
– Improved bit cell optimization
• NFET/PFET centering and Beta/Gamma control
• Minimize device fluctuation by limiting device-geometry scaling
larger cell
• Lpoly, Weff, LER
– Leakage / defect mechanisms

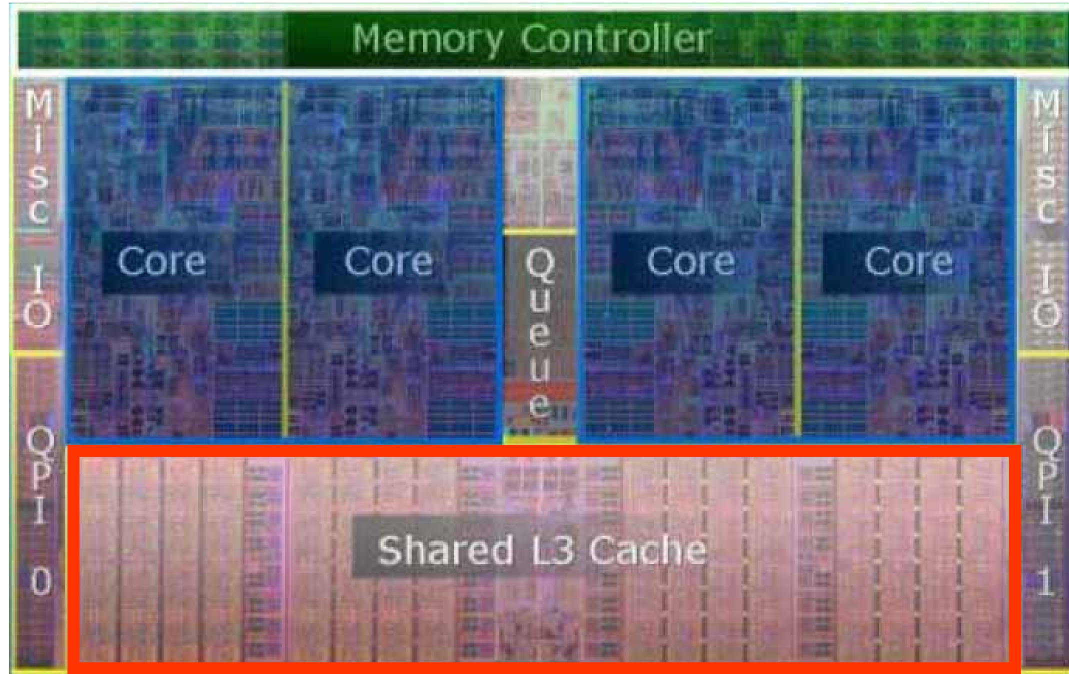
Source: Harold Pilo IEDM2006 Short Course

Nehalem(Intel) 2,4 or 8 Cores

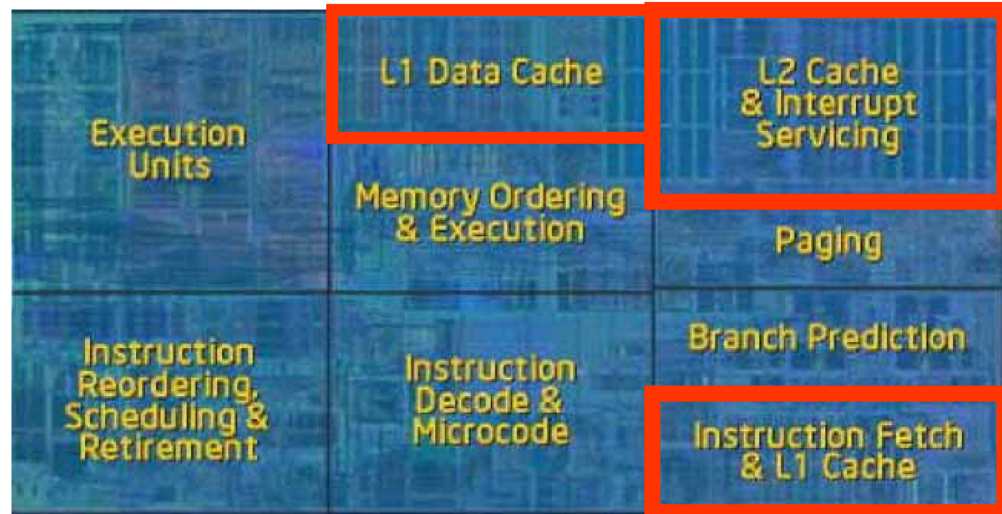
Voltage/Frequency Partitioning

- DDR Vcc
- Core Vcc
- Uncore Vcc

Dynamic Power Management



Chip



Core

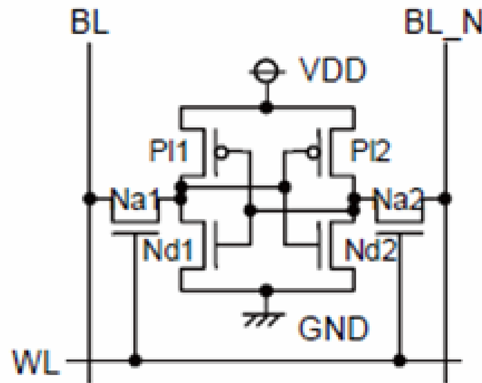
8T SRAMCell
32kB L1 I -cache
32kB L1 D-cache
256kB L2 -cache

6T SRAMCell
8 MB L3 cache

Source: Intel Developer Forum 2008

6T and 8T Cell

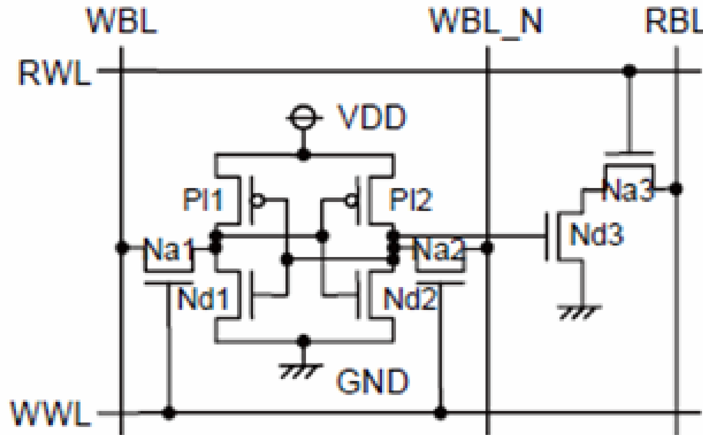
6T Cell



Cell size is small

For high density use

8T Cell



Add separate read function

Cell size increase 30%

For low voltage use

Source: Morita et. al, Symp. on VLSI Circ. 2007

5. Roadmap for further future as a Personal View

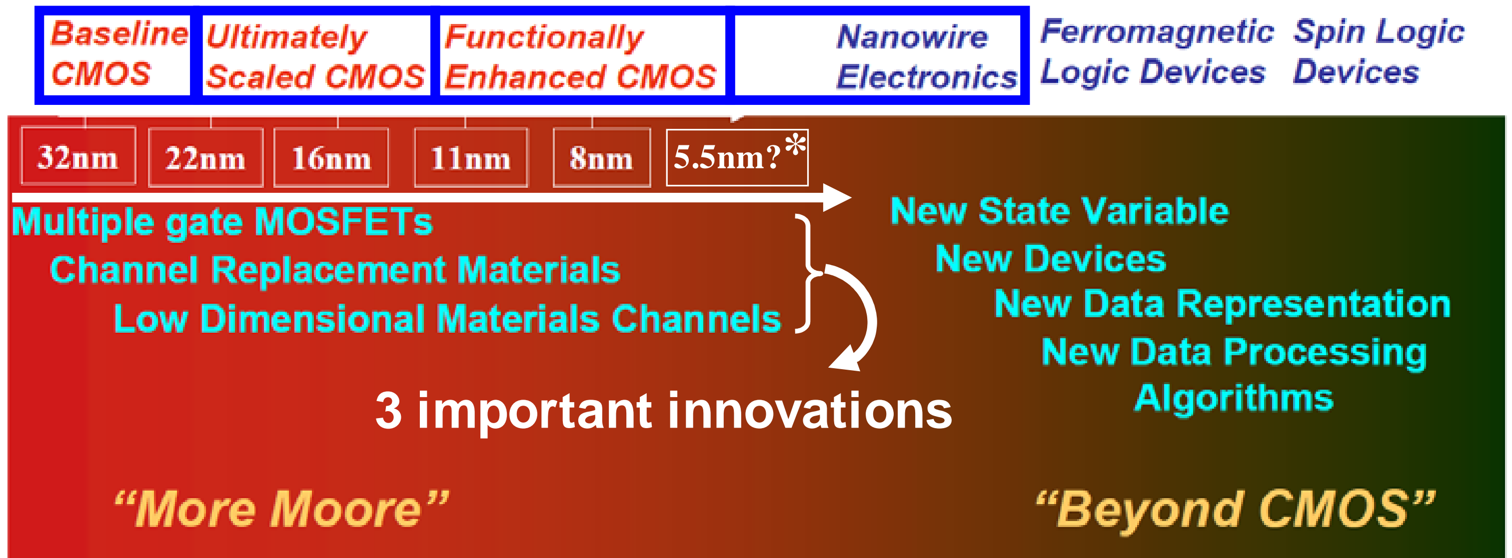
-There will be still 4~6 cycles (or technology generations) left until we reach 11 ~ 5.5 nm technologies, at which we will reach down-scaling limit, in some year between 2020-30 (H. Iwai, IWJT2008).

-Even After reaching the down-scaling limit, we could still continue R & D, seeking sufficiently higher Id-sat under low Vdd.

-Two candidates have emerged for R & D

1. Nanowire/tube MOSFETs
2. Alternative channel MOSFETs (III-V, Ge)

- Other Beyond CMOS devices are still in the cloud.

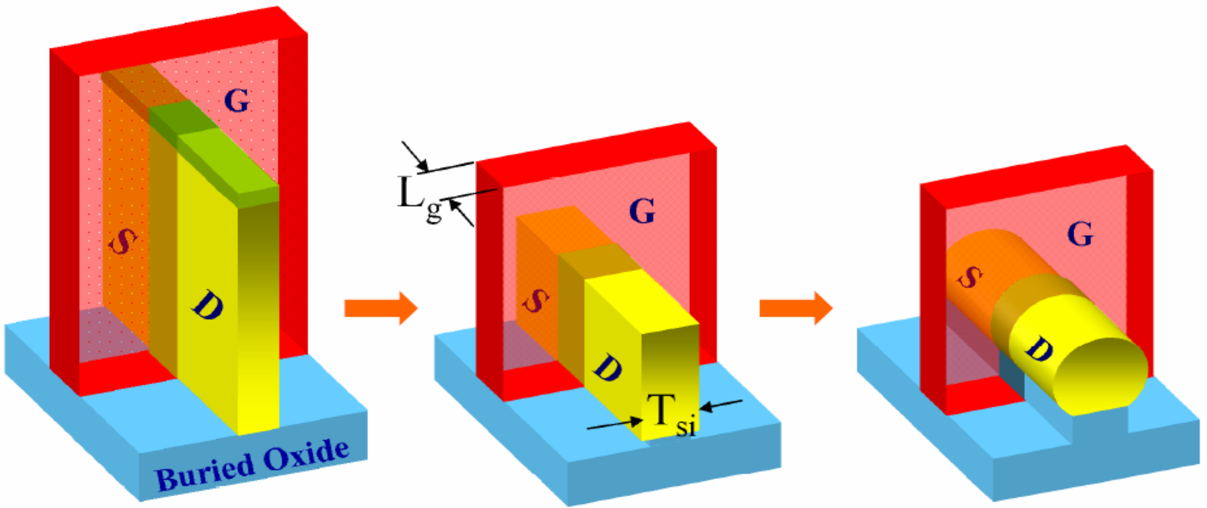


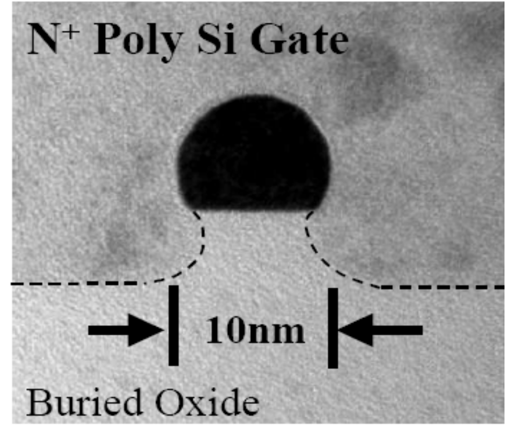
ITRS figure edited by Iwai

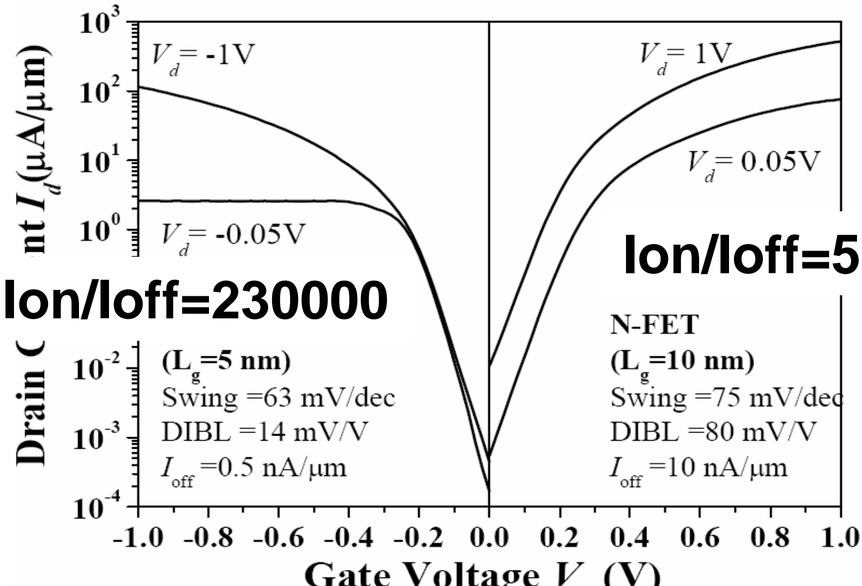
Source: 2008 ITRS Summer Public Conf.

*5.5nm? was added by Iwai

FinFET to Nanowire







Ion/Ioff=230000

Ion/Ioff=52200

Channel conductance is well controlled by Gate even at L=5nm

F.-L.Yang, VLSI2004

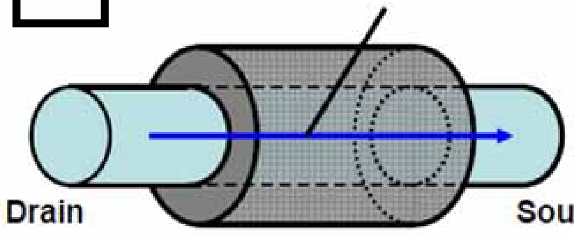
Si nanowire FET with Semi-1D Ballistic Transport

Merit of Si-nanowire

Source: Y. Lee., T. Nagata., K. Kakushima., K. Shiraishi, and H. Iwai, IWDTF 2008, Tokyo, November, 2008

0

Reduction in Ioff (Isd-leak)



Good control of Isd-leak by surrounding gate

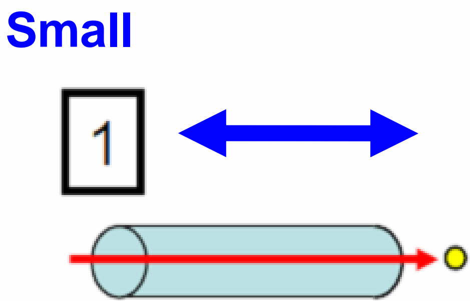
Trade off

Carrier scattering probability

Small **Large**

# of quantum channel

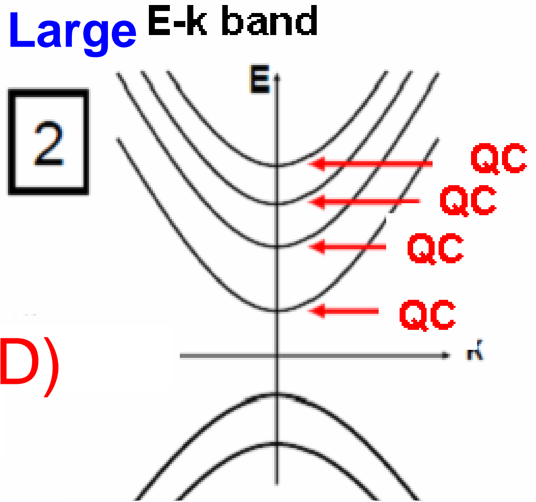
Small **Large**



High Conduction (1D)

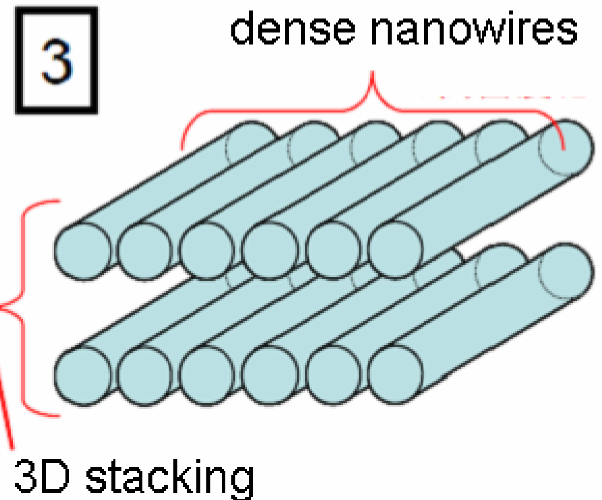
Go=77.8μS/wire

Increase in Ion (Id-sat)



Multiple quantum channel (QC) used for conduction

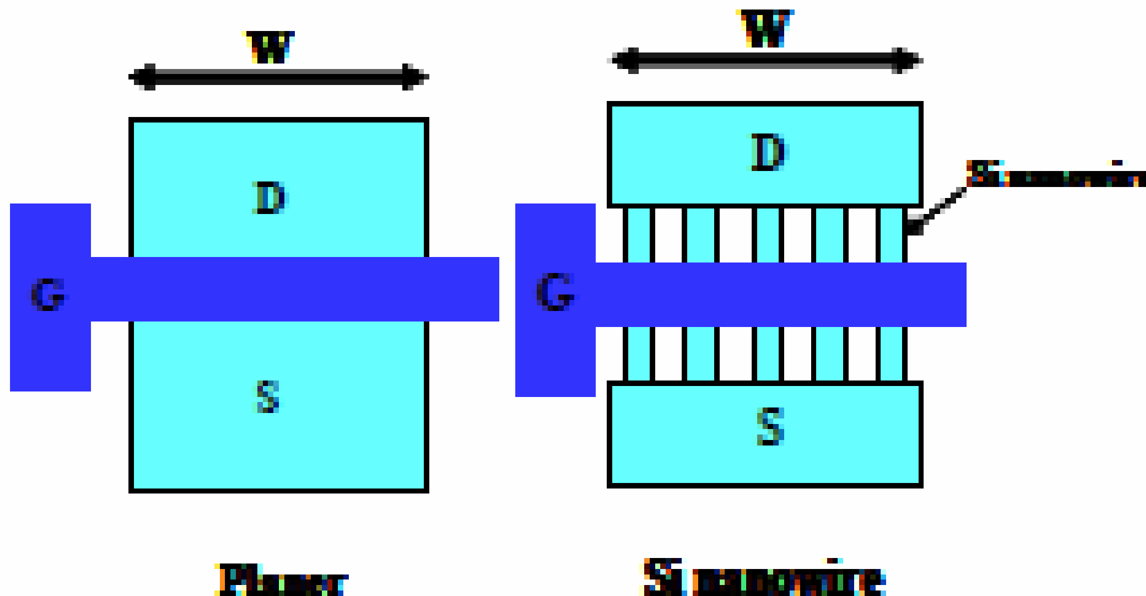
Source: T. Ohno, K. Shiraishi, and T. Ogawa, Phys. Rev. Lett. ,1992



High-density lateral and vertical integration

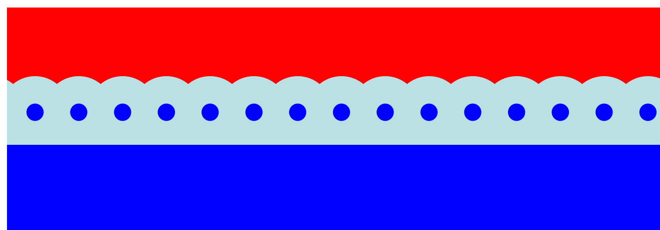
Comparison with planar bulk MOSFET

Does Nanowire FET have really higher drive current per unit area?

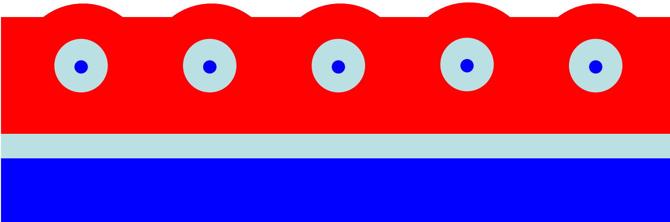


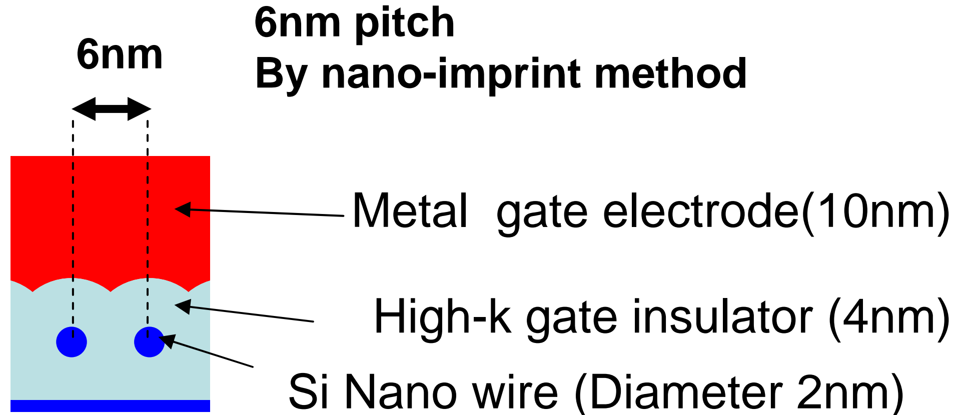
Maximum number of wires per 1 µm

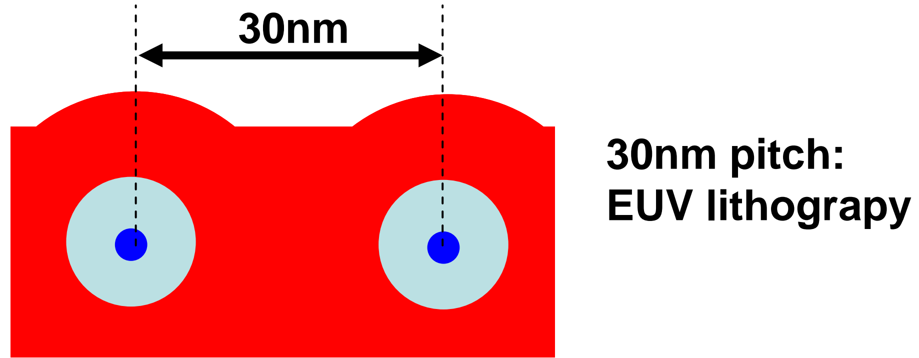
Front gate type MOS 165 wires /µm

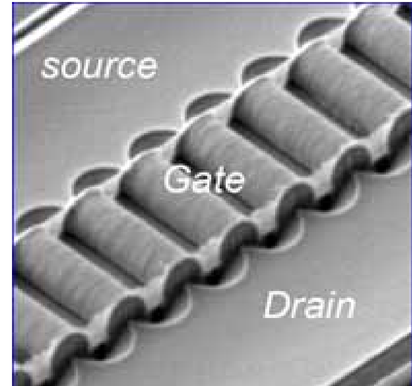


Surrounded gate type MOS 33 wires/µm





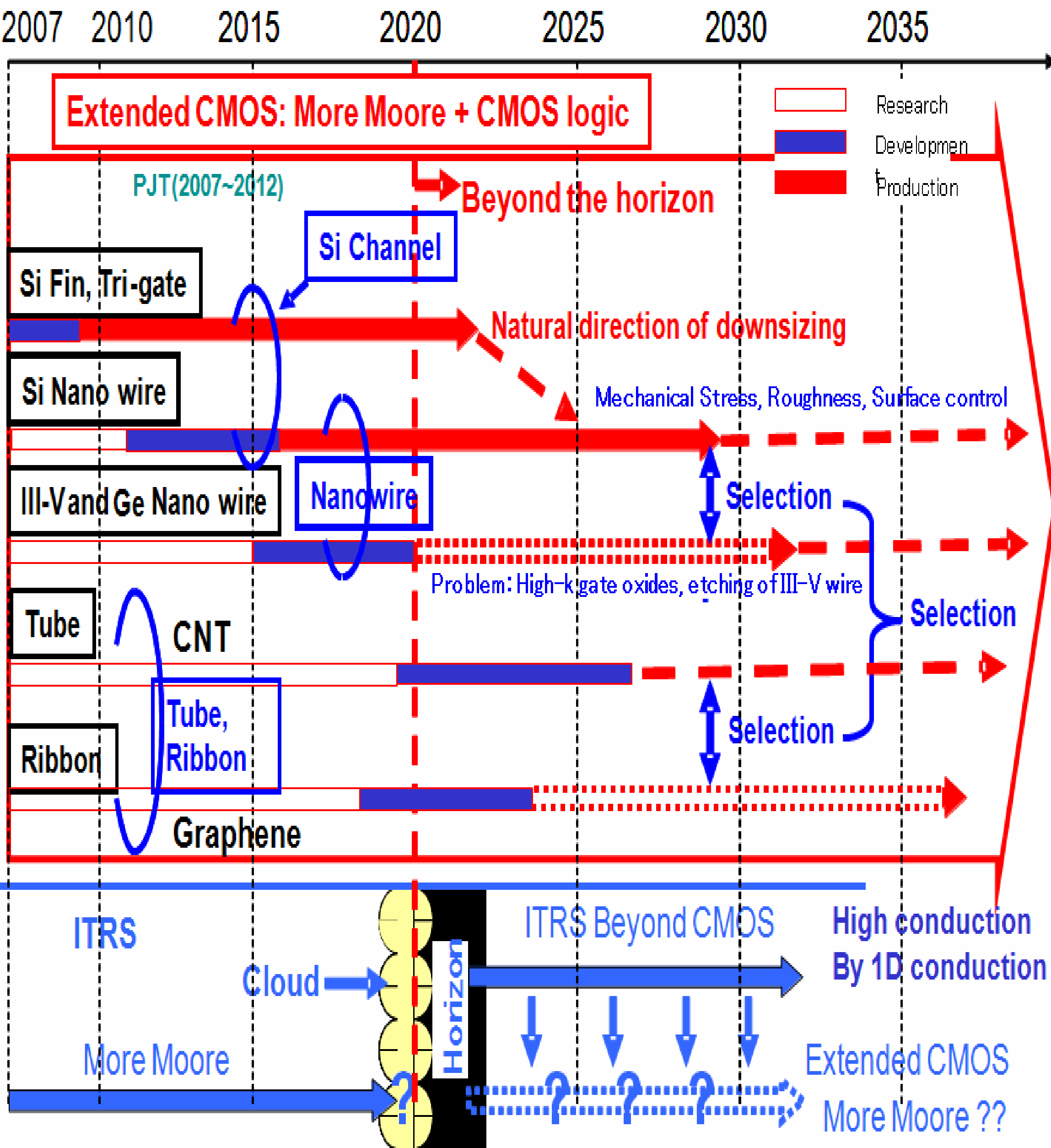




Surrounded gate MOS

Our roadmap for R &D

Source: H. Iwai, IWJT 2008



Current Issues

Si Nanowire

Control of wire surface property
Source Drain contact
Optimization of wire diameter
Compact I-V model

III-V & Ge Nanowire

High-k gate insulator
Wire formation technique

CNT:

Growth and integration of CNT
Width and Chirality control
Chirality determines conduction types: metal or semiconductor

Graphene:

Graphene formation technique
Suppression of off-current

Very small bandgap or no bandgap (semi-metal)

Control of ribbon edge structure which affects bandgap

Long term roadmap for development

Source: H. Iwai, IPFA 2006

Size (Gate length etc)

5 nm?

New Materials, New Process, New Structure (Logic, Memory)

Hybrid integration of different functional Chip
Increase of SOC functionality

3D integration of memory cell
3D integration of logic devices

Miniaturization of Interconnects on PCB
(Printed Circuit Board)

Low cost for LSI process
Revolution for CR, Equipment, Wafer

Introduction of algorithm of bio-system
Brain of insects, human

Scaling proceeds

Saturation of Downsizing

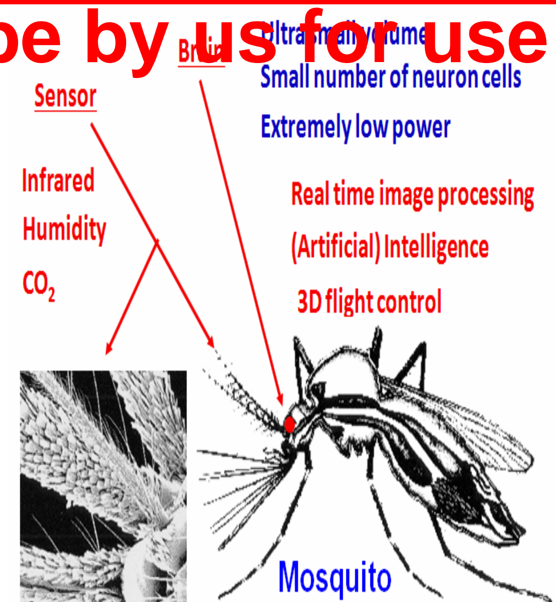
We do not know how?

Some time in 2020 - 2030

After 2050?

We do know system and algorithms are important!

But do not know how it can be by us for use of bio?



Dragonfly brain has even further higher performance

Acknowledgement

I would like to express deep appreciation to the following people for the useful advice and support for material preparation. Special thanks to ITRS committee for the permission to refer roadmap and Public conference.

ITRS Committee: Hidemi Ishiuchi (Toshiba), Paolo Gargini (Intel)

Toshiba Corporation: Mitsuo Saito, Yukihiro Urakawa, Tomoaki Yabe

Tsukuba University: Kenji Shiraishi, Kenji Natori

Intel Corporation: Mark Bohr

IBM Alliance : B.S. Haran et al,

Tokyo Institute of Technology: Kuniyuki Kaukshima, Parhat Ahmet, Takamasa Kawanago, Yeonghun Lee

Thank you

for your attention!