# Logic LSI Technology Roadmap for 22nm and beyond

July 8, 2009

IPFA 2009@Suzhou, China

Hiroshi Iwai

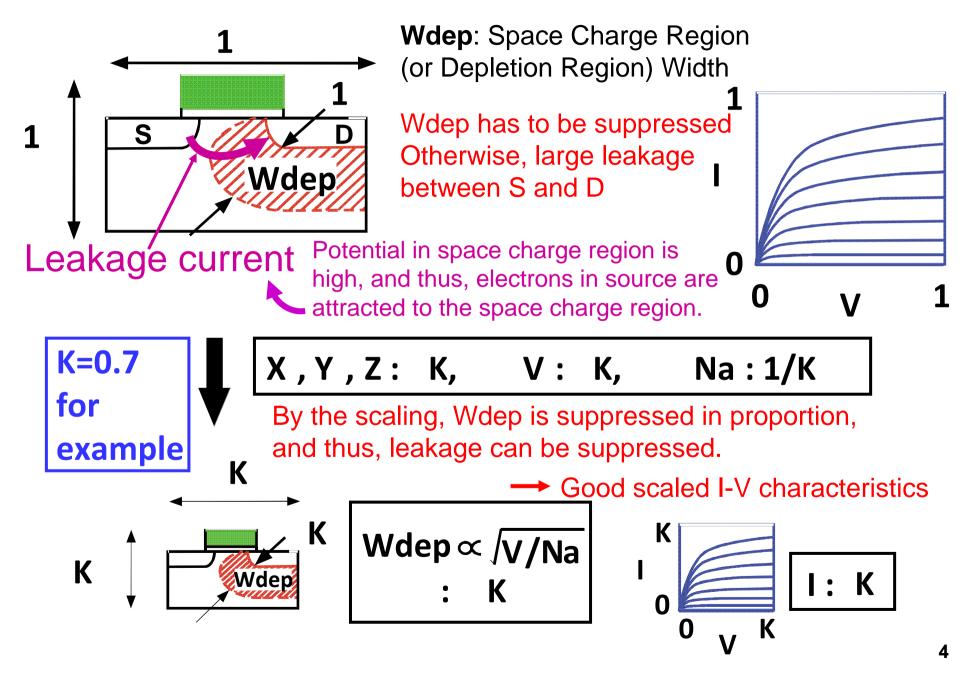
Tokyo Institute of Technology

## 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**



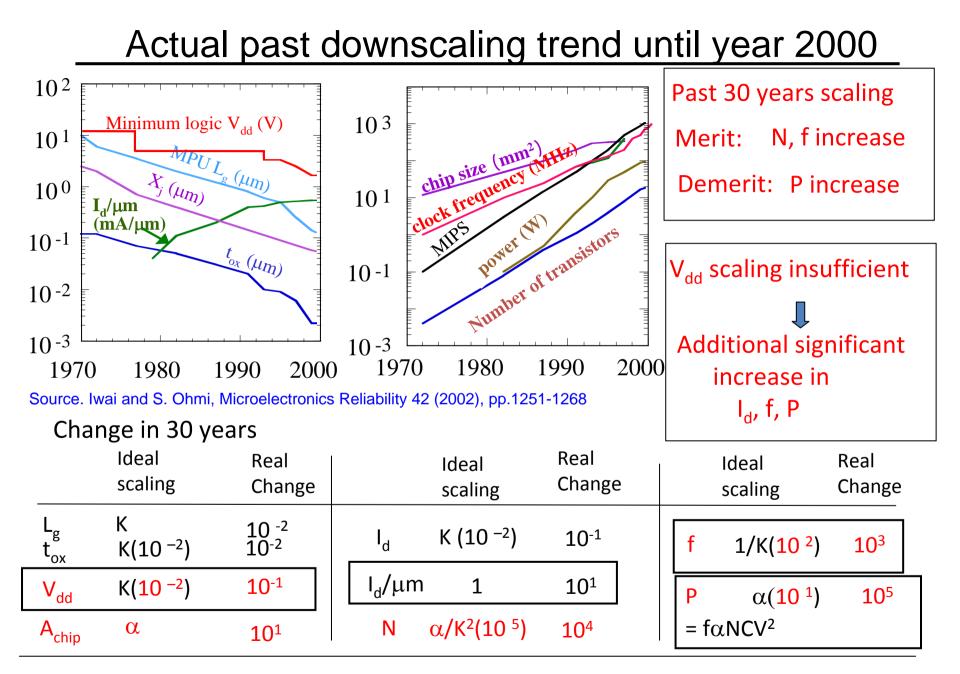
## Downscaling merit: Beautiful!

Geometry & Supply voltage	L <sub>g</sub> , W <sub>g</sub> T <sub>ox,</sub> V <sub>dd</sub>	К	Scaling K: K=0.7 for example				
Drive current in saturation	I <sub>d</sub>	К	$I_{d} = v_{sat}W_{g}C_{o}(V_{g}-V_{th}) \qquad C_{o}: \text{ gate C per unit area}$ $\longrightarrow W_{g}(t_{ox}^{-1})(V_{g}-V_{th}) = W_{g}t_{ox}^{-1}(V_{g}-V_{th}) = KK^{-1}K = K$				
I <sub>d</sub> per unit W <sub>g</sub>	I <sub>d</sub> /μm	1	$I_d$ per unit $W_g = I_d / W_g = 1$				
Gate capacitance	Cg	К	$C_g = \varepsilon_o \varepsilon_{ox} L_g W_g / t_{ox} \longrightarrow KK / K = K$				
Switching speed	τ	К	$\tau = C_g V_{dd} / I_d \longrightarrow KK / K = K$				
Clock frequency	f	1/K	$f = 1/\tau = 1/K$				
Chip area	A <sub>chip</sub>	α	$\alpha$ : Scaling factor $\longrightarrow$ In the past, $\alpha > 1$ for most cases				
Integration (# of Tr)	N	$\alpha/K^2$	N $\rightarrow \alpha/K^2 = 1/K^2$ , when $\alpha = 1$				
Power per chip	Р	α	fNCV <sup>2</sup> /2 $\rightarrow$ K <sup>-1</sup> ( $\alpha$ K <sup>-2</sup> )K(K <sup>1</sup> ) <sup>2</sup> = $\alpha$ = 1, when $\alpha$ =1				

2 Generations	k= 0.7 <sup>2</sup> =0.5 and $\alpha$ =1
Single MOFET	
9	$Vdd \rightarrow 0.5$
	$Lg \rightarrow 0.5$
	$Id \rightarrow 0.5$
	$Cg \rightarrow 0.5$
	P (Power)/Clock
	$\rightarrow 0.5^3 = 0.125$
	$\tau$ (Switching time) $\rightarrow 0.5$
Chip	
	N (# of Tr) $\rightarrow 1/0.5^2 = 4$
	f (Clock) $\rightarrow$ 1/0.5 = 2
	$P(Power) \rightarrow 1$

- 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 vs. Rule of military)



Vd scaling insufficient,  $\alpha$  increased  $\rightarrow$  N, Id, f, P increased significantly

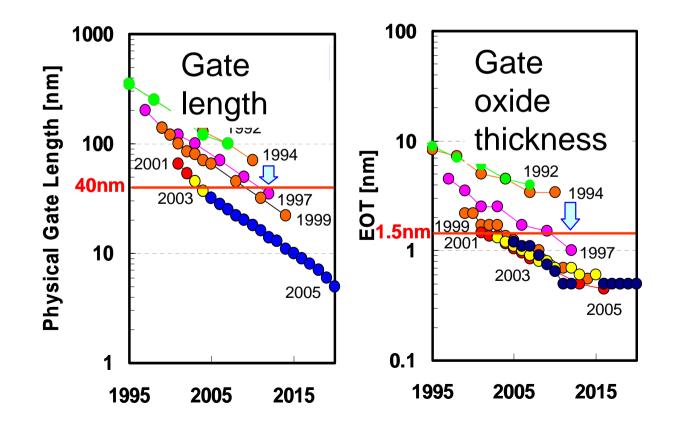
- Now, power and/or heat generation are the limiting factors of the down-scaling
- Supply voltage reduction is becoming difficult, because Vth cannot be decreased any more, as described later.
- Growth rate in clock frequency and chip area becomes smaller.

## 2. ITRS Roadmap (for 22 nm CMOS logic)

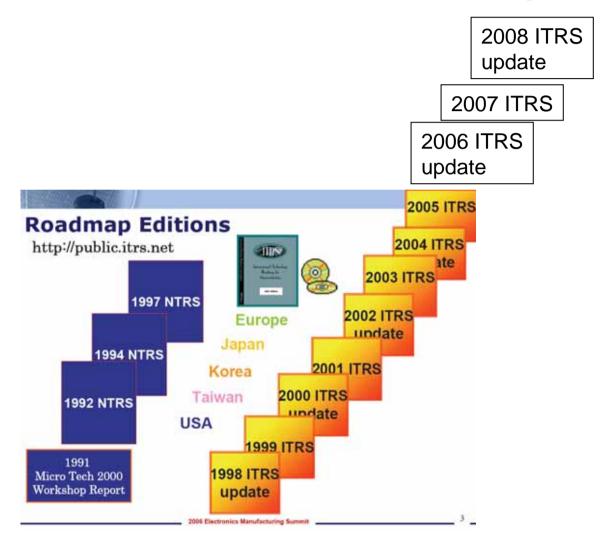
What is a roadmap? What is ITRS?

Roadmap: Prediction of future technologies

ITRS: International Technology Roadmap for Semiconductors made by SIA (Semiconductor Industry Association with Collaboration with Japan, Europe, Korea and Taiwan)



## 1992 -1997:NTRS (National Technology Roadmap) 1998 - : ITRS (International Technology Roadmap)



#### ITRS Roadmap does change every year!

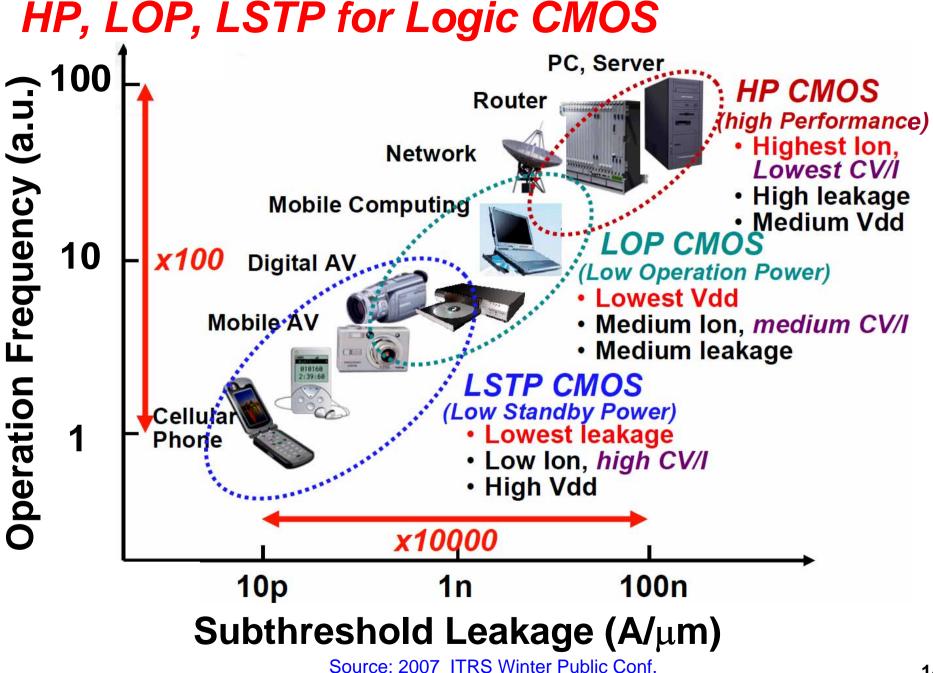
 2007 Edition
 2003 Edition

 2006 Update
 2002 Update

 2005 Edition
 2001 Edition

 2004 Update
 2000 Update

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



### What does '45 nm' mean in 45 nm CMOS Logic?

#### **'XX nm CMOS Technology**

**Commercial Logic CMOS products** 

#### ITRS 2008 Update

for High Performance Logic

Technology name	Starting Year		Year	Half Pitch (1 <sup>st</sup> Metal)	Physical Gate Length
45 nm	2007	<b>←→</b>	2007	68 nm	32 nm
			2008	59 nm	29 nm
32 nm	2009?	$\longrightarrow$	2009	52 nm	27 nm
			2010	45 nm	<u>24 nm</u>

#### '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.

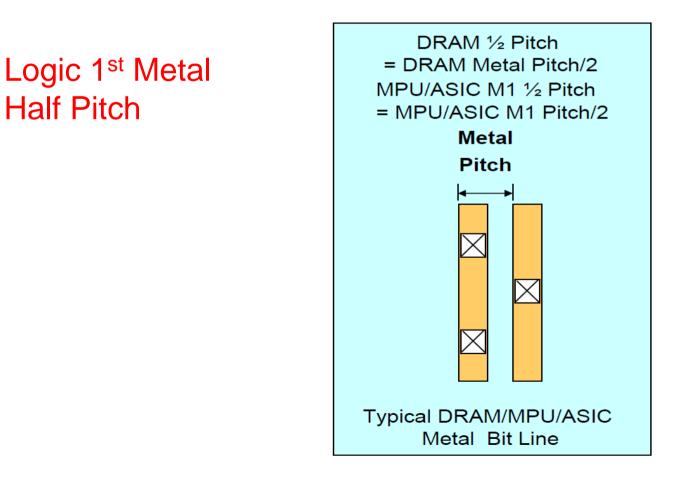
#### What does '45 nm' mean in 45 nm CMOS Logic?

#### $8\mu m \rightarrow 6\mu m \rightarrow 4\mu m \rightarrow 3\mu m \rightarrow 2\mu m \rightarrow 1.2\mu m \rightarrow 0.8\mu m \rightarrow 0.5\mu m$

- Originally, 'XX' means lithography resolution.

Half Pitch

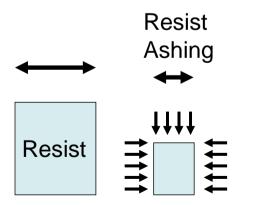
- Thus, 'XX' was the gate length, and half pitch of lines
- 'XX' had shrunk 0.7 in 3 years in average (0.5 in 6 years) those days.



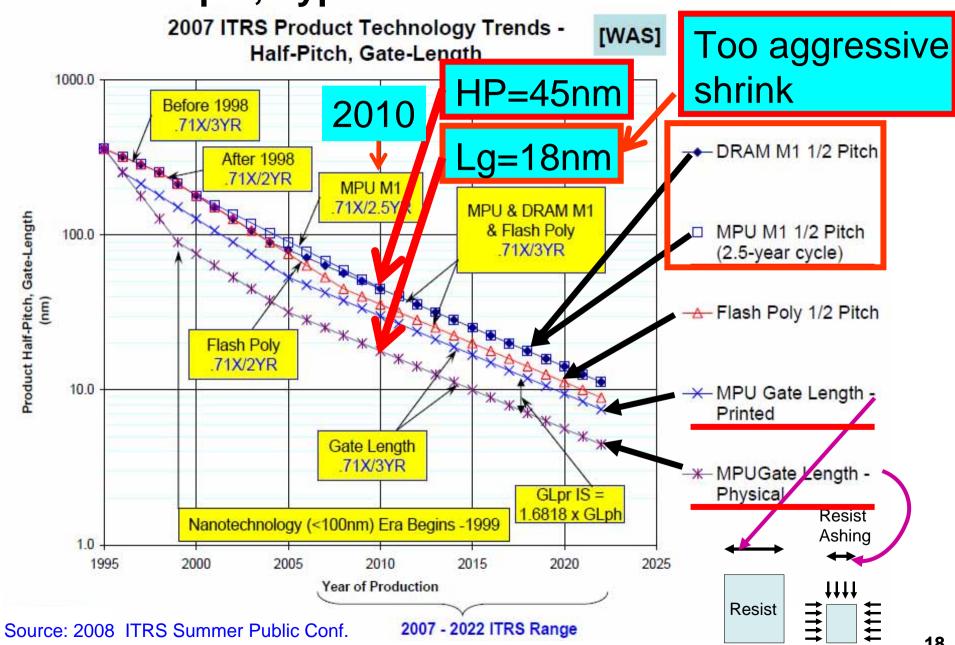
### What does '45 nm' mean in 45 nm CMOS Logic?

#### $\rightarrow$ 350nm $\rightarrow$ 250nm $\rightarrow$ 180nm $\rightarrow$ 130nm $\rightarrow$ 90nm $\rightarrow$ 65nm $\rightarrow$ 45nm

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

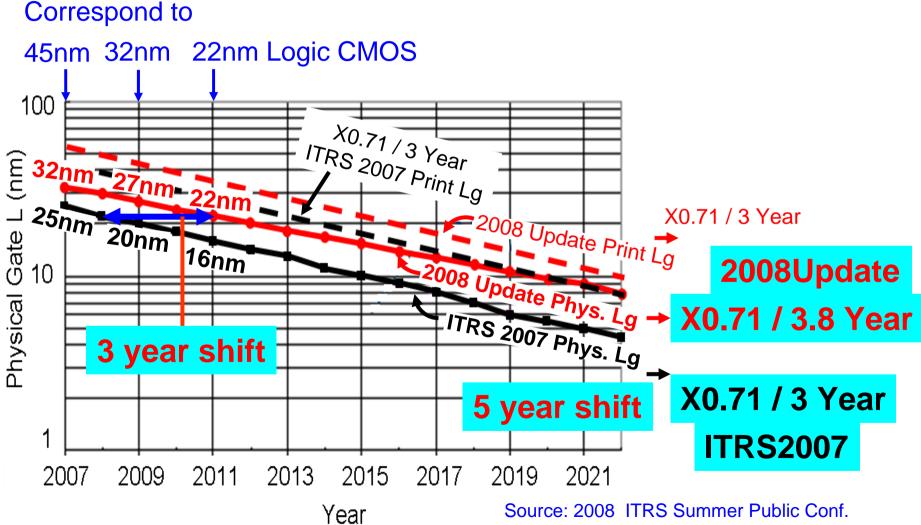


## For example, Typical Half Pitches at ITRS 2007

#### 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.



## EOT and Xj shift backward, corresponding to Lg shift

#### EOT: 0.55 nm $\rightarrow$ 0.88 nm, Xj: 8 nm $\rightarrow$ 11 nm @ 22nm CMOS

**Correspond to 22nm** Source: 2008/ ITRS Summer Public Conf. Likely in 2008 Update Year of Production 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017 2018 2019 2020 2021 2022 2007 MPU/ASIC Lg (nn 25 7 5.6 23 20 18 16 14 13 11 10 9 8 6.3 5 4.5 2008 MPU/ASIC Lg (nm) 32 24 22 20 18 15 29 27 17 14.0 12.8 9.7 8.9 8.1 11.7 10.7 2005 2009 2010 2012 Shift/Interpolate Formua intro intro intro intro intro intro intro intrp intro intro intro intro EOT w/3E20 poly, bulk 1.2 0.54 0.71 0.41 MPU (nm) EOT w/3E20 poly, bulk 1.3 1.2 1.2 1 Likely in 2008 Update MPU (nm) 0.55 EOT w/metal gate, bulk 0.9 0.75 0.65 0.50 MPU (nm) EOT w/metal gate, bulk 0.75 1.0 0.95 0.88 Likely in 2008 Update MPU (nm) Drain Ext. X; bulk MPU (nm) 9 12.5 11 10 8 Drain Ext. X; bulk MPU (nm) 11 11 11 11 9 8.5 7.7 7 11 Likely in 2008 Update filled in for metal gate EOT for 2009/10 non-steady trend based on latest conference presentations corrected

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

#### **'XX nm CMOS Technology**

Commercial Logic CMOS products

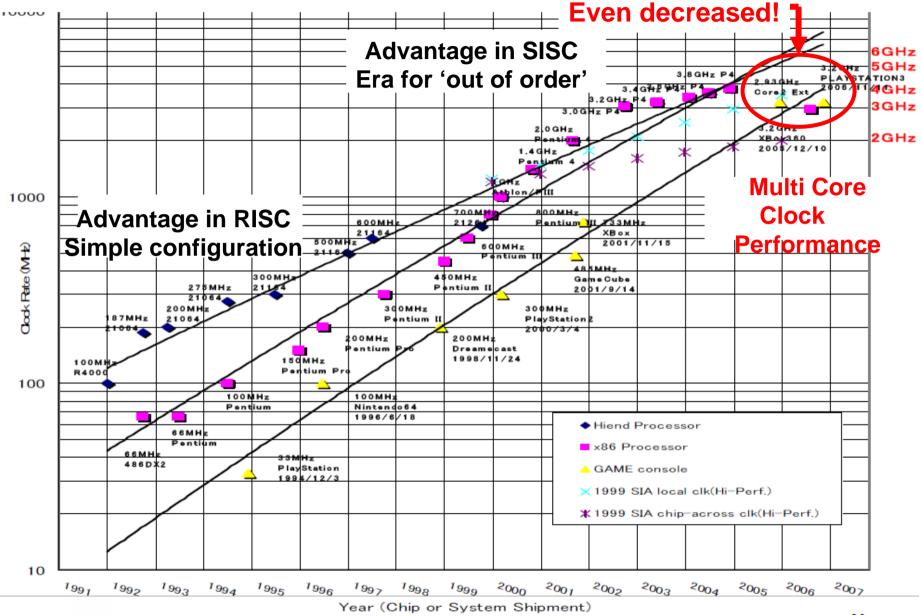
#### ITRS (Likely in 2008 Update)

for High Performance Logic

Starting Year		Year	Half Pitch (1 <sup>st</sup> Metal)	Physical Gate Length
	$\longrightarrow$	2007	68 nm	32 nm
_001		2008	59 nm	29 nm
2009?	<b>←→</b>	2009	52 nm	27 nm
		2010	45 nm	24 nm
2011?~	←→	2011	40 nm	22 nm
2012?		2012	36 nm	20 nm
2013?~		2013	32 nm	18 nm
2014?	_	2014	29 nm	16 nm
	Year 2007 2009? 2011?~ 2012? 2013?~	Year 2007 2009? 2011?~ 2012? 2013?~	Year $2007$ $2007$ 2009? $2009$ $2009$ 2011?~ $2010$ $2010$ 2012? $2013$ $2013$	Year(1st Metal) $2007$ $\leftarrow$ $2007$ $68 \text{ nm}$ $2009?$ $\leftarrow$ $2008$ $59 \text{ nm}$ $2009?$ $\leftarrow$ $2009$ $52 \text{ nm}$ $2011?~$ $\leftarrow$ $2010$ $45 \text{ nm}$ $2012?$ $\leftarrow$ $2012$ $36 \text{ nm}$ $2013?~$ $\leftarrow$ $2013$ $32 \text{ nm}$

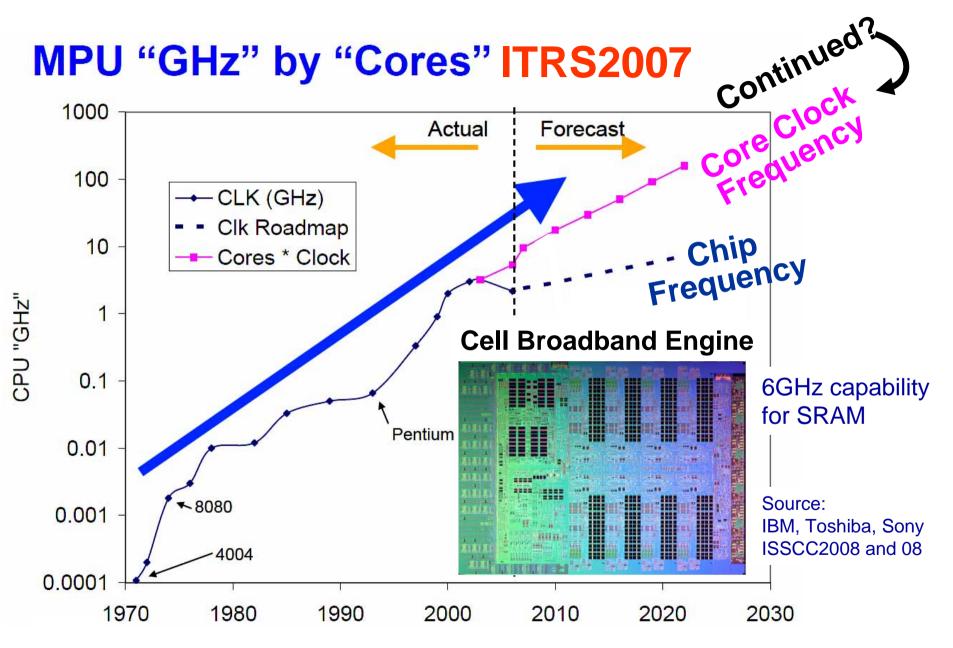
Source: 2008 ITRS Summer Public Conf.

From ITRS2008 Update, maybe XX nm stands for the physical Gate length

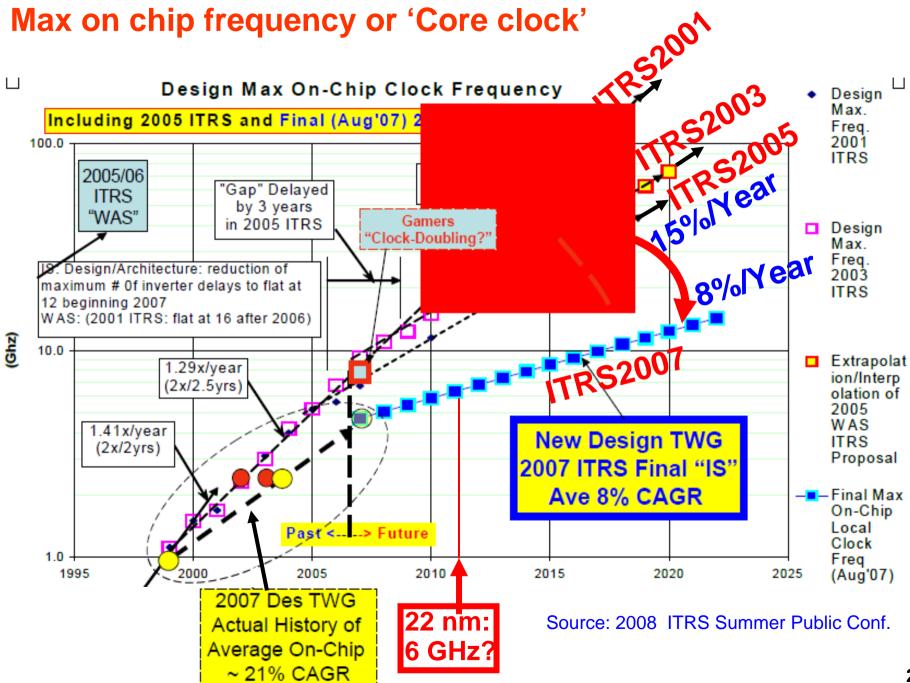


#### **<u>Clock frequency does not increase aggressively anymore.</u>**

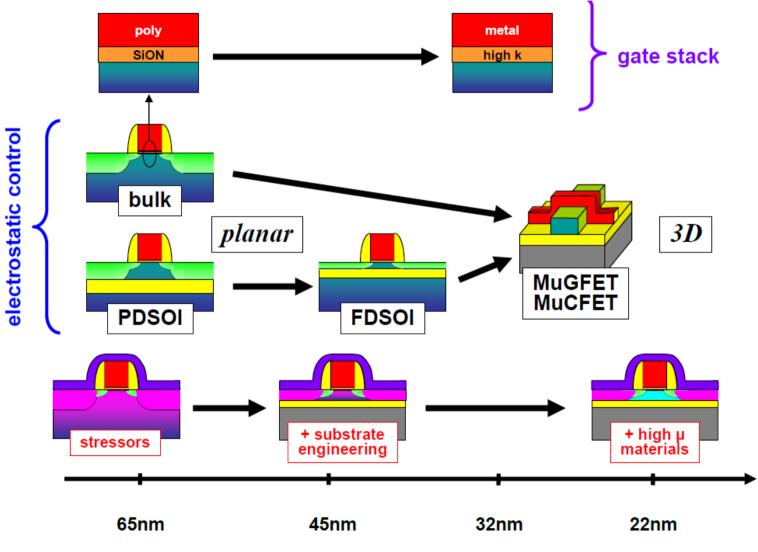
Source: Mitsuo Saito, Toshiba



Source: 2007 ITRS Winter Public Conf.



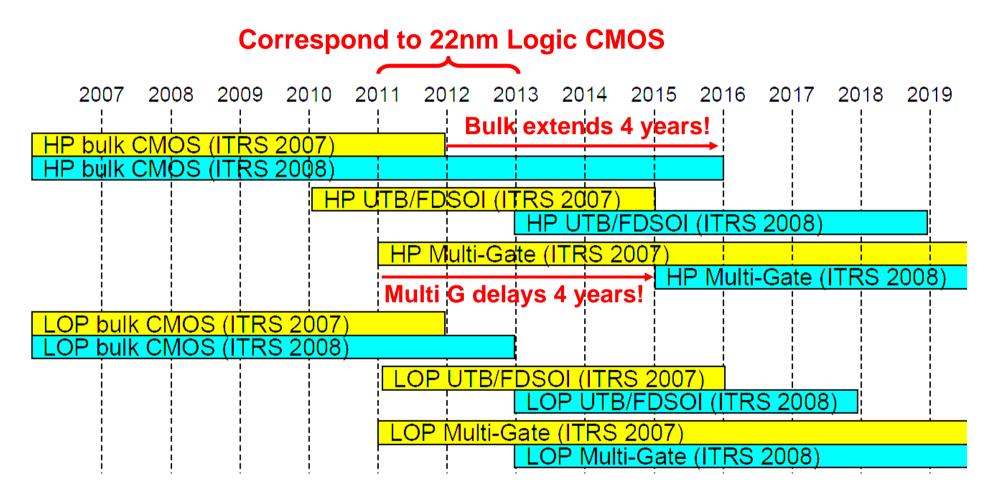
Structure and technology innovation (ITRS 2007)



Source: 2008 ITRS Summer Public Conf.

Timing of CMOS innovations shifts backward.

#### **Bulk CMOS has longer life now!**



Source: 2008 ITRS Summer Public Conf.

## Wafer size (ITRS 2007)

Correspo	nd	to	<b>22nm</b>
	-		

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 <sup>2</sup> )	310	246	195	310	246	195	310	246	195
MPU High-Performance Active Transistor Area(mm <sup>2</sup> )	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??

### ITRS2008 Low-k Roadmap Update

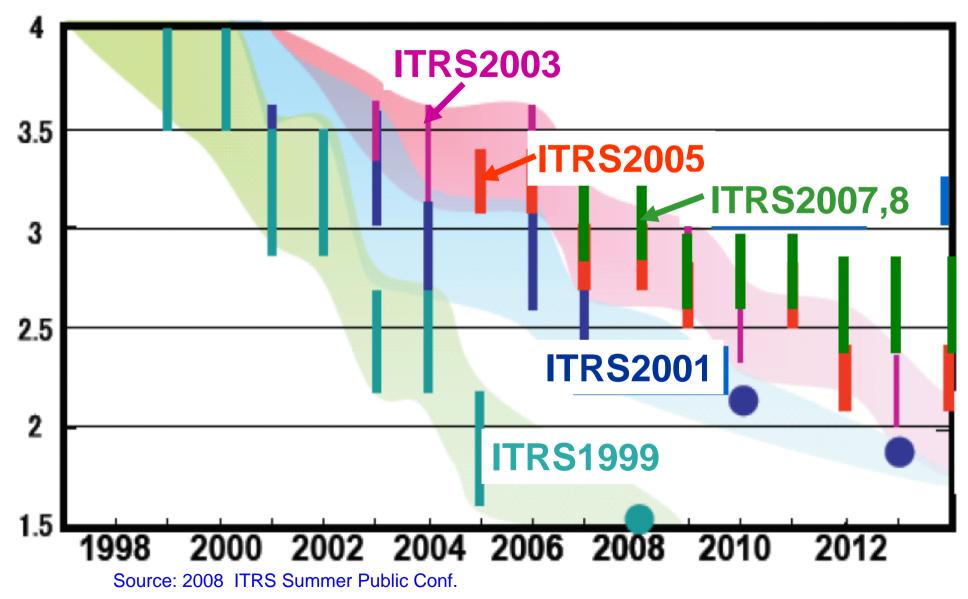
#### Near-term Year of Production 2008 2009 2010 2011 2012 2013 ITRS Interlevel metal insulator - effective dielectric 2007 2.7-3.0 2.5-2.8 2.5-2.8 2.5-2.8 2.1-2.4 2.1-2.4 constant (ĸ) Interlevel metal insulator - effective dielectric Update 2.9-3.3 2.6-2.9 2.6-2.9 2.6-2.9 2.4-2.8 2.4-2.8 2008 constant (x) ITRS Interlevel metal insulator - bulk dielectric 2.3-2.7 2.1-2.4 2.1-2.4 2.1-2.4 1.8-2.1 1.8-2.1 2007 constant (ĸ) ........... Interlevel metal insulator - bulk dielectric Update 2.3-2.6 2.3-2.6 2.5-<u>2.8</u> 2.3-2.6 2.1-2.4 2007 constant (ĸ) ......

#### **Correspond to 22nm Logic**

Source: 2008 ITRS Summer Public Conf.

#### k value increases by $0.1 \sim 0.3$

### Historical Transition of ITRS Low-k Roadmap

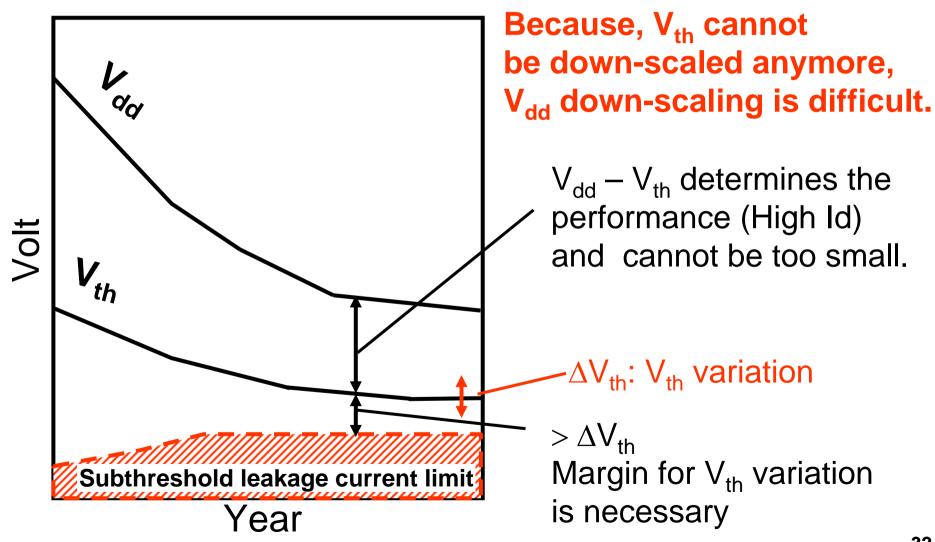


## Roadmap towards 22nm technology and beyond

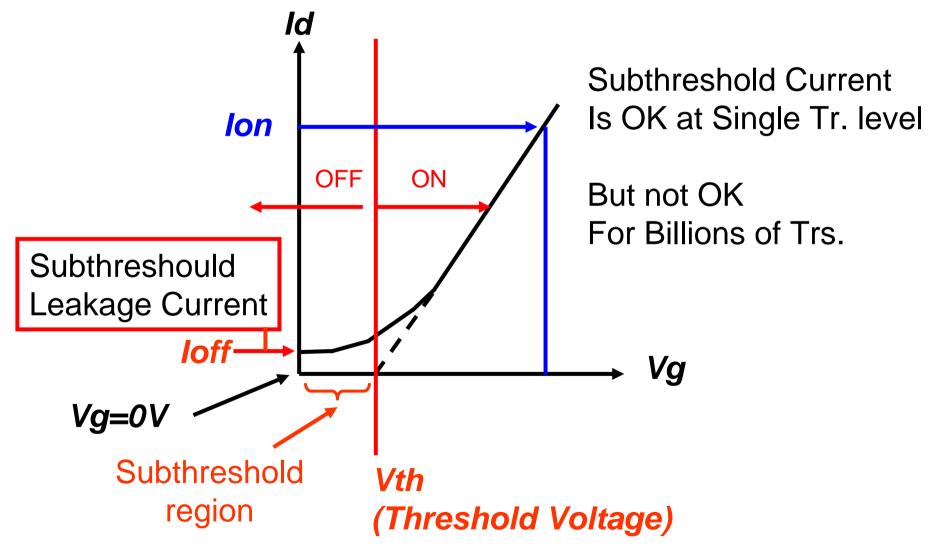
- Physical gate length downsizing rate will be less aggressive.
- Corresponding to the above, performance increase would slow down Clock frequency, etc.
- Introduction of innovative structures UTB SOI and DG delayed, and bulk CMOS has longer life than predicted by previous ITRS roadmaps.

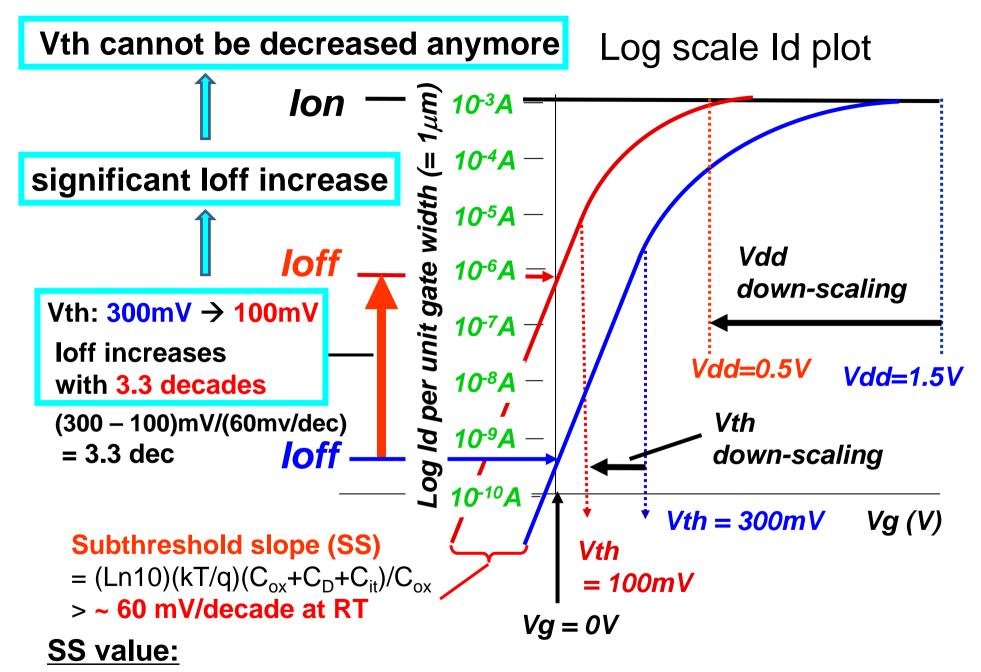
# 3. Voltage Scaling/ Low Power and Leakage

#### Difficulty in Down-scaling of Supply Voltage: Vdd



#### Subtheshold leakage current of MOSFET

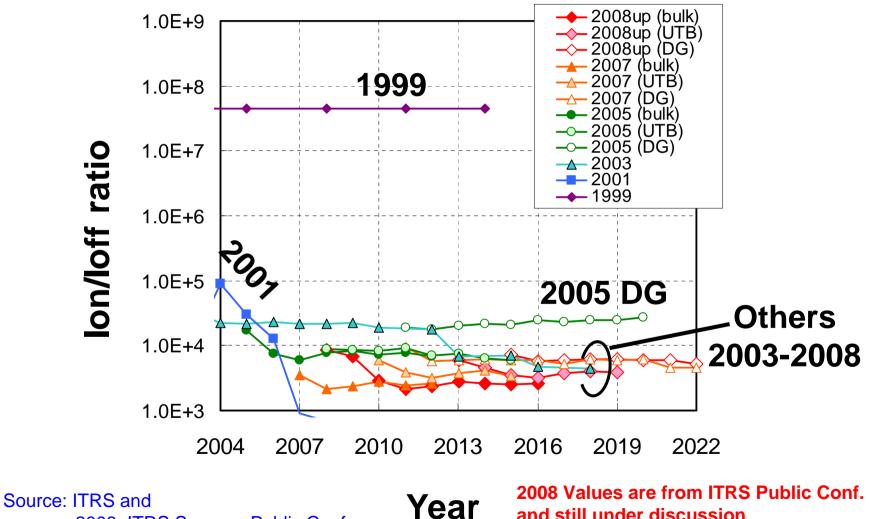




Constant and does not become small with down-scaling

#### **ITRS for HP logic**

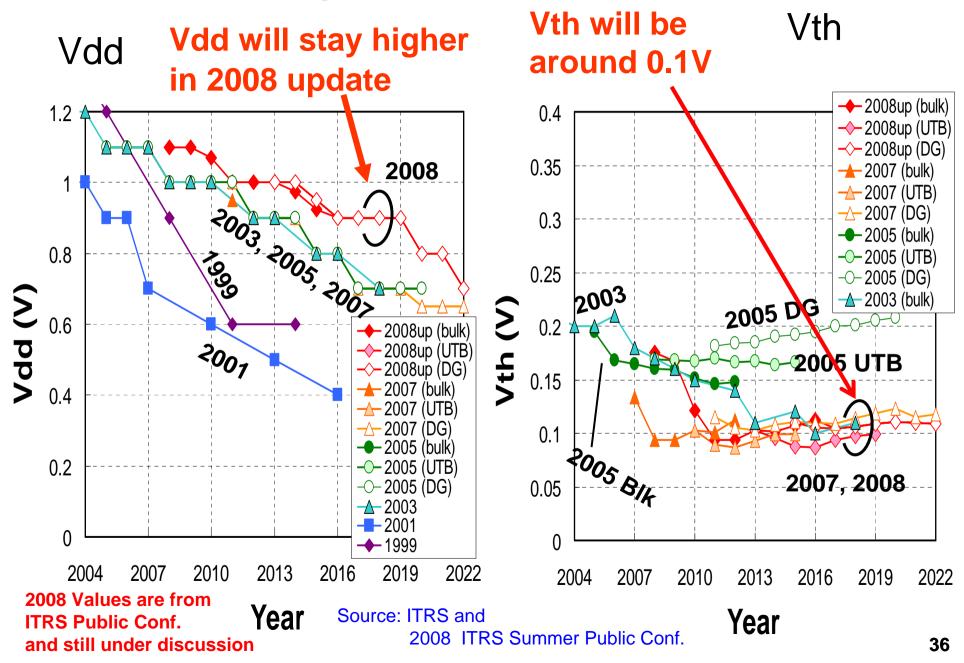
## **Ion/Ioff** ratio

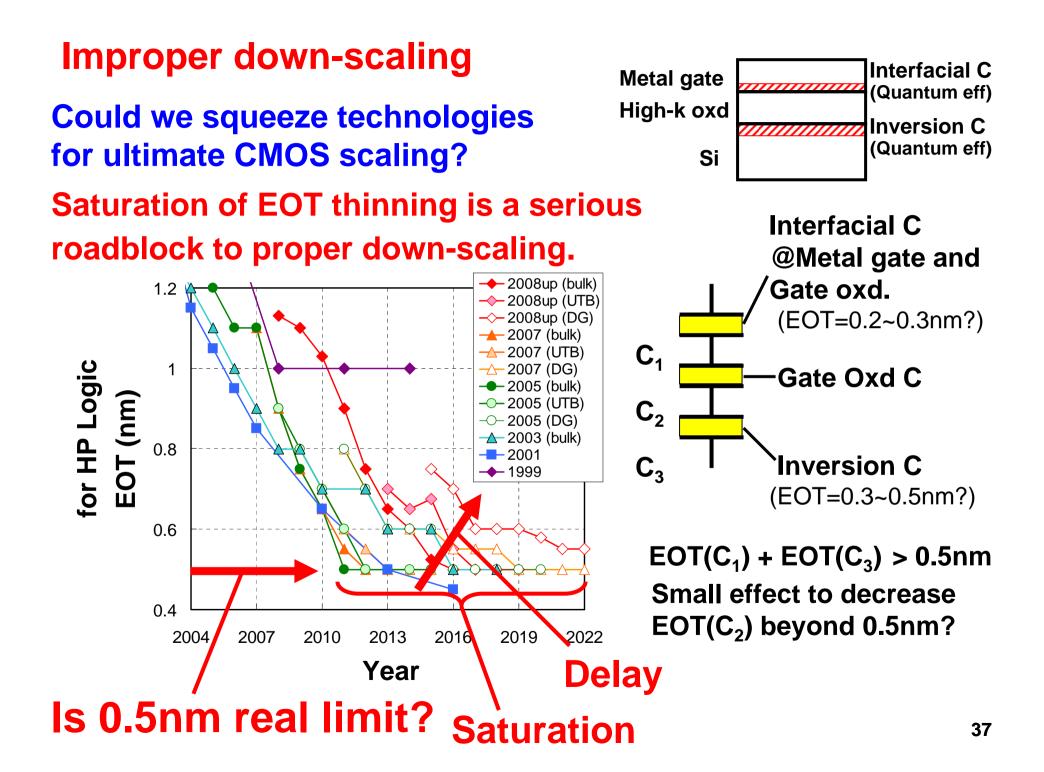


2008 ITRS Summer Public Conf.

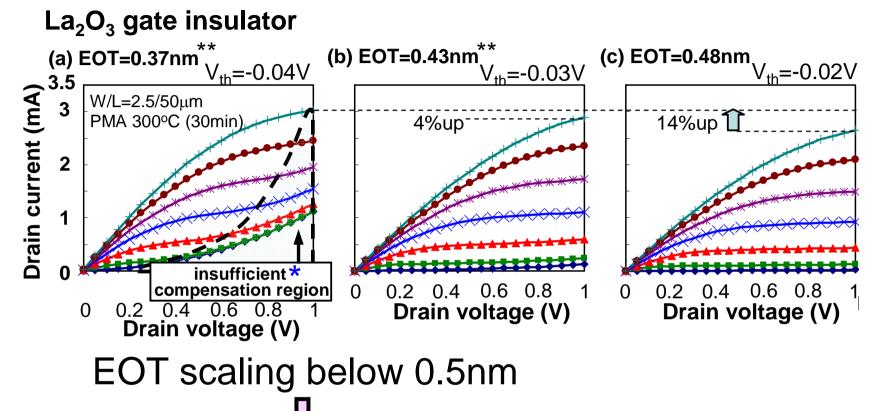
and still under discussion

#### **ITRS for HP logic**





# **EOT<0.5nm with Gain in Drive Current is Possible**



# 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

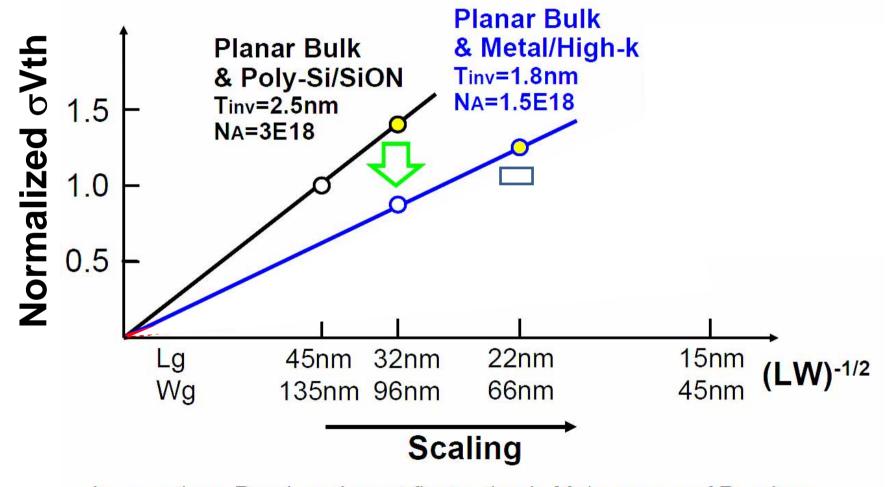
Because Lg is very large (2.5μm), gate leakage is large in case (a). The gate leakage component was subtracted from measured data for case (a). However, if we make small gate length, the gate leakage current should become sufficiently small to be ignored compared with Id as we verified with SiO<sub>2</sub> 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 downscaling possible.

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

# Random Variability Reduction Scenario in ITRS 2007



Assumption: Random dopant fluctuation is Main source of Random Variability: Line width roughness of Lg and Wg is not considered in this Source: 2007 TERS WINTER Public Cont.

# 4. SRAM cell scaling

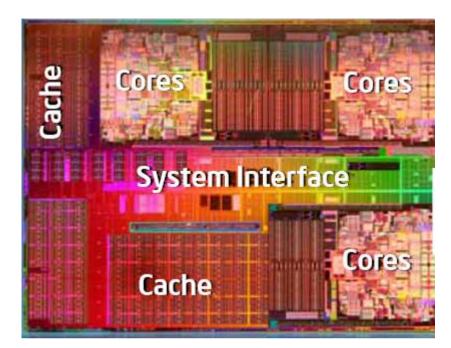
Intel® Xeon® 7400 Series (Dunnington)

45 nm high-k 6 cores 16MB shared L3 cache

Source: Intel Developer Forum 2008

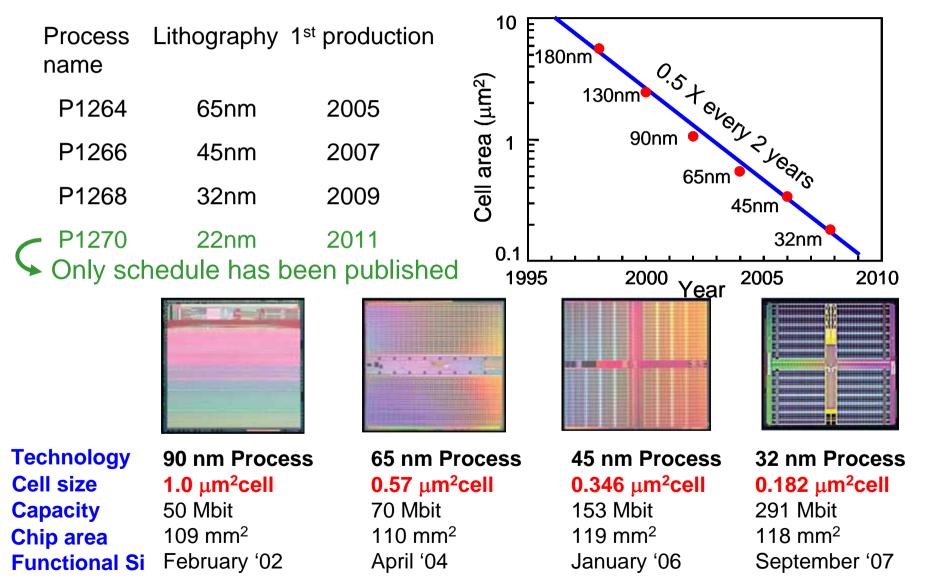
#### Cache occupies huge area

- $\rightarrow$  Cell size of SRAM should be minimized
- $\rightarrow$  However, Isd-leak should be minimized
  - $\rightarrow$  Vth are often designed to be higher than Min. logic Vth
  - $\rightarrow$  Lg are often designed to be larger than Min. logic Lg



# 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



# 22 nm technology 6T SRAM Cell: Size = $0.1 \mu m$

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

Announced on Aug 18, 2008

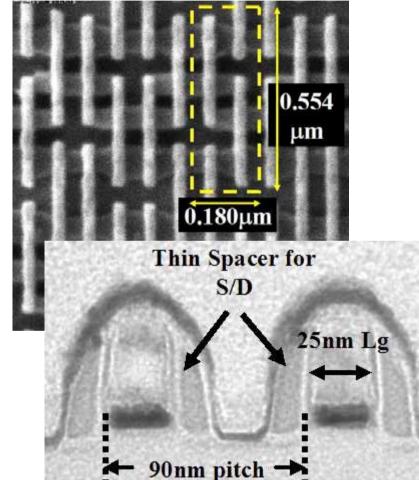
Consortium: IBM (NYSE), AMD, Freescale, STMicroelectronics, Toshiba and the College of Nanoscale Science and Engineering (CNSE)

# 0.1 $\mu$ m cell size is almost on the down-scaling trend

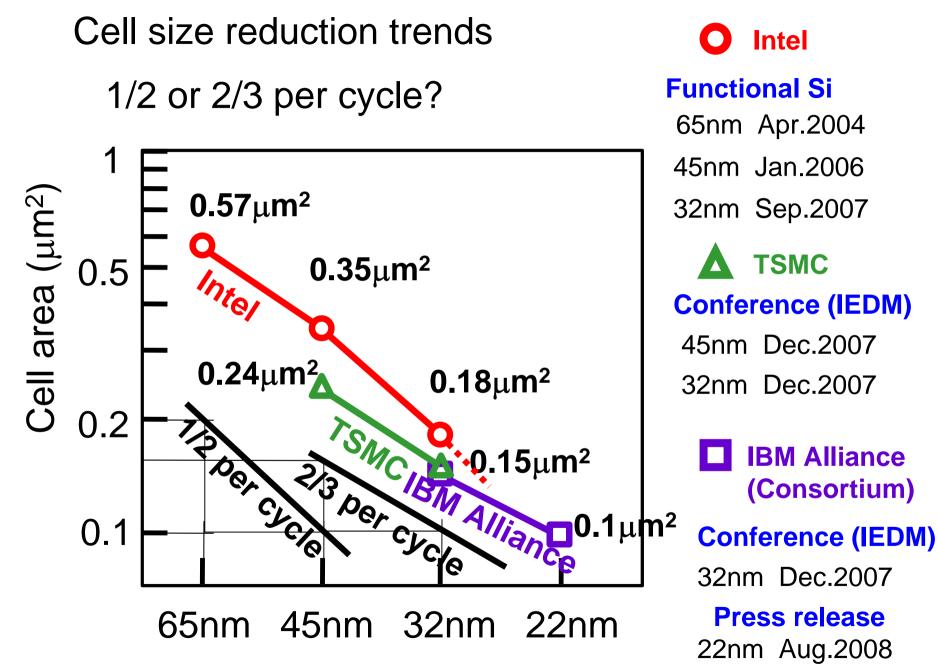
#### New technologies introduced

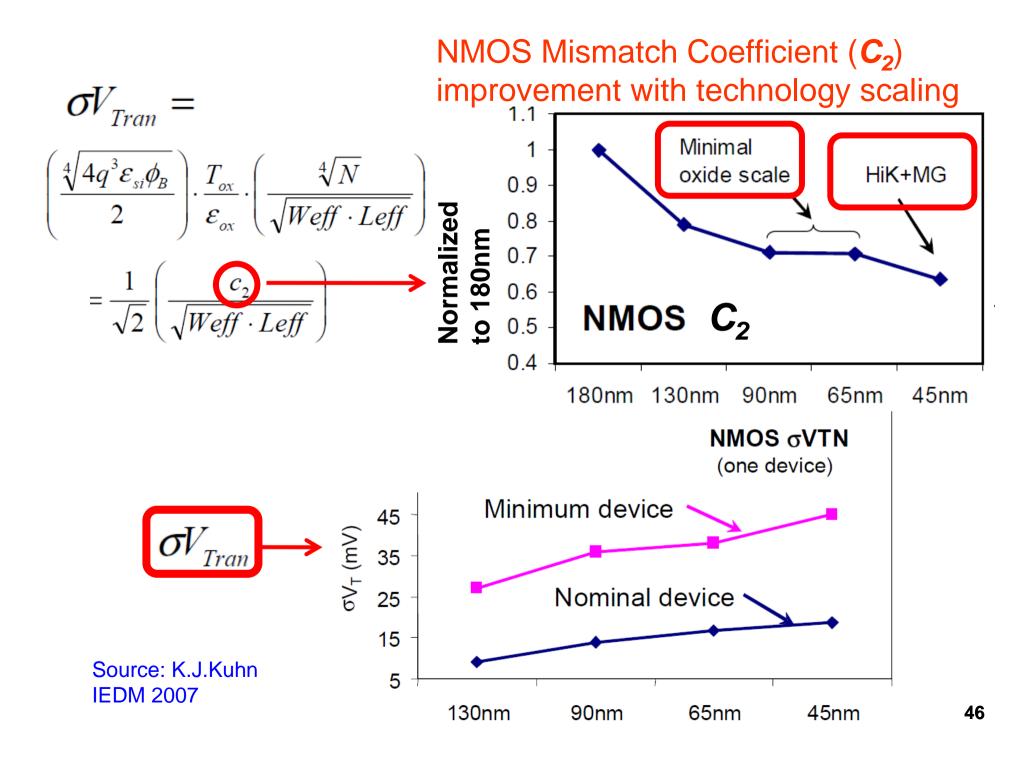
- High-NA immersion lithography
- High-K metal gate stacks
- 25 nm gate lengths
- Thin composite oxide-nitride spacers
- Advanced activation techniques
- Extremely thin silicide
- Damascene copper contacts

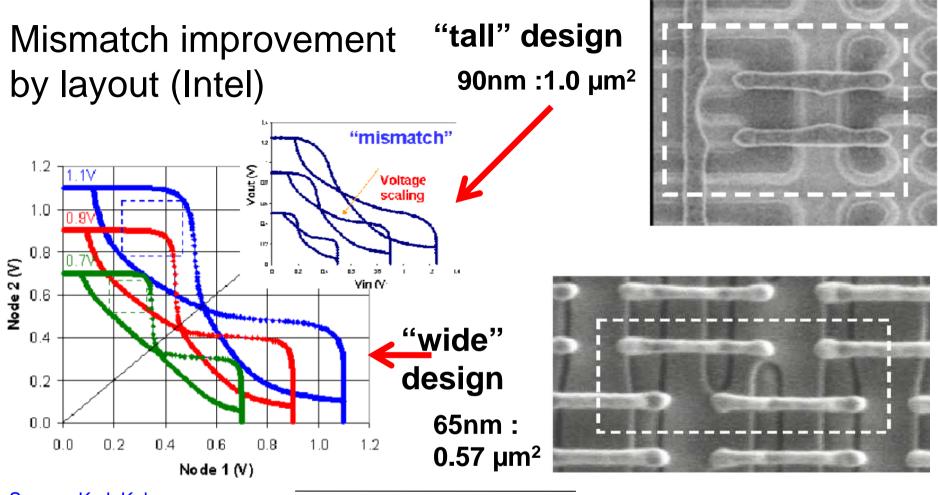
# Static noise margin of 220 mV at 0.9 V



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



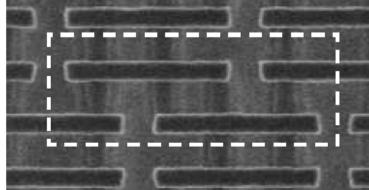


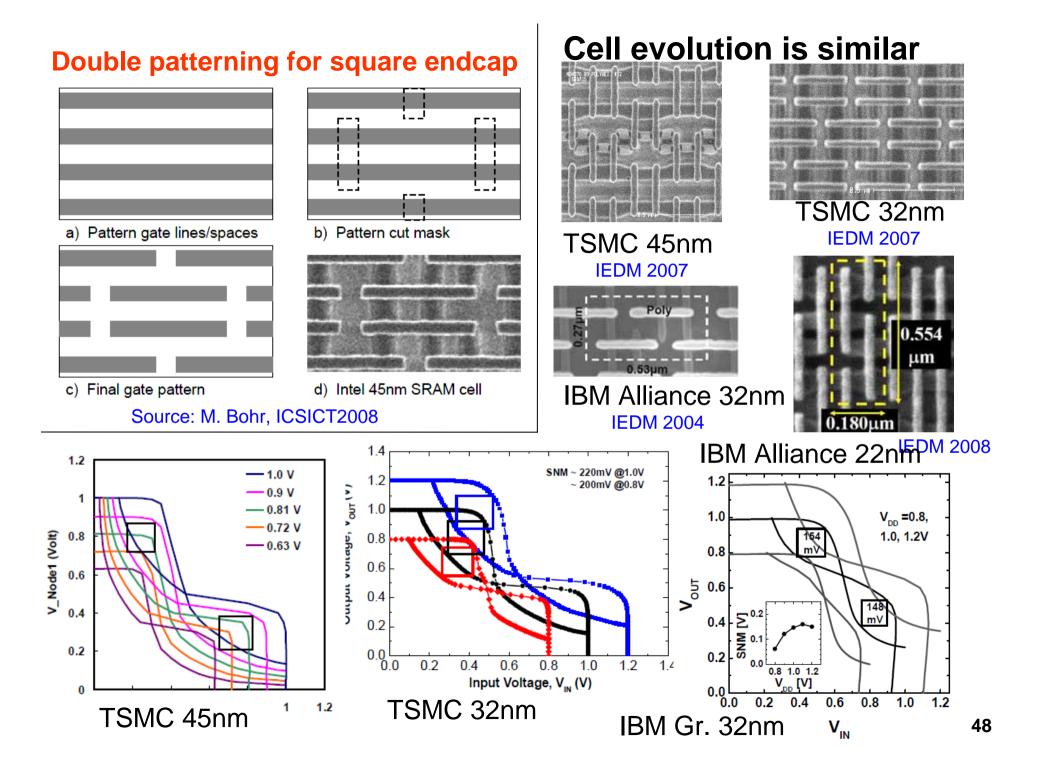


Source: K. J. Kuhn IEDM2007 Tech. Dig. pp.471

> "wide" design (Square endcaps)

> > $45 nm \ 0.346 \ \mu m^2$





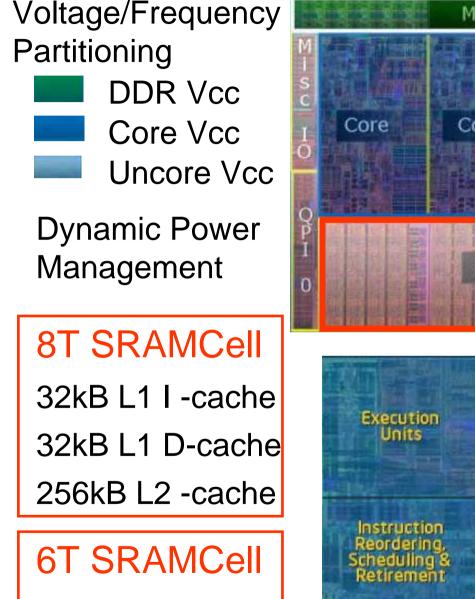
Most Difficult part of SRAM down-scaling is Vdd down-scaling

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

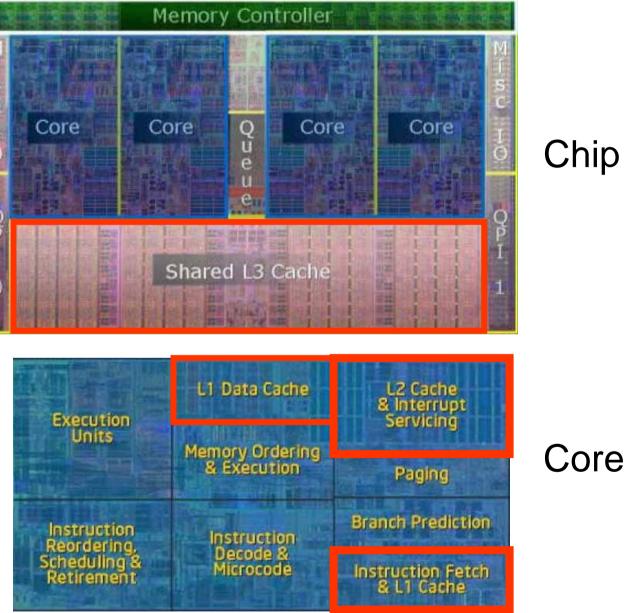
Also, under low Vdd, read- and write margin degrades, data retention degrade.

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

## Nehalem(Intel) 2,4 or 8 Cores



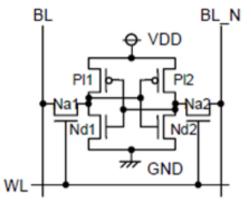
8 MB L3 cache



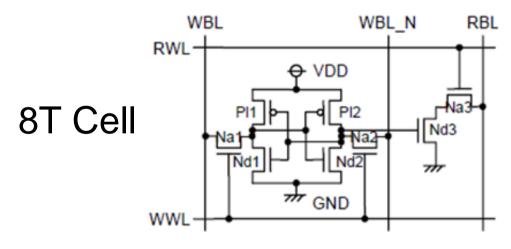
Source: Intel Developer Forum 2008

### 6T and 8T Cell





Cell size is small For high density use



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

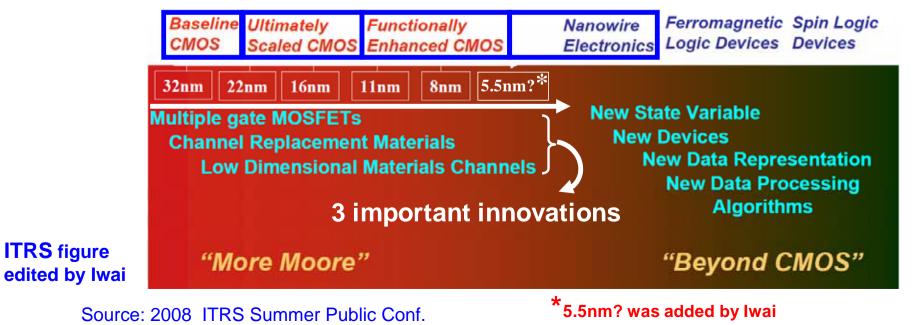
Add separate read function

Cell size increase 30%

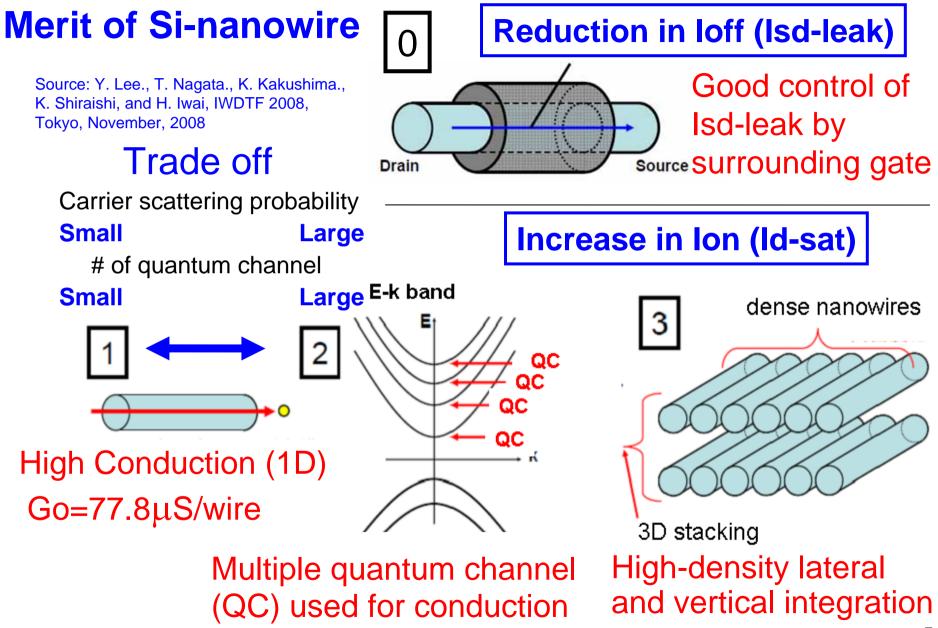
For low voltage use

5. Roadmap for further future as a Personal View

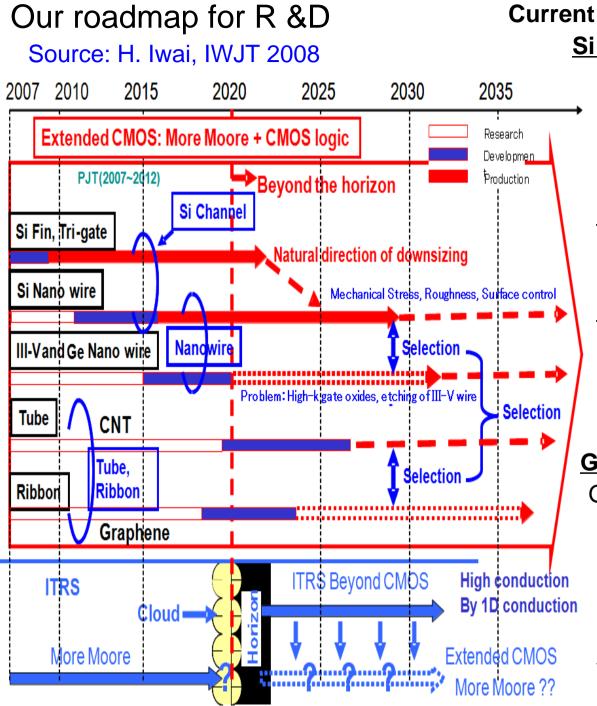
- -There will be still 4~6 cycles (or technology generations) left until we reach 11 ~ 5.5 nm technologies, at which we will reach downscaling limit, in some year between 2020-30 (H. Iwai, IWJT2008).
- -Even After reaching the down-scaling limit, we could still continue R & D, seeking sufficiently higher 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.



# Si nanowire FET with Semi-1D Ballistic Transport



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



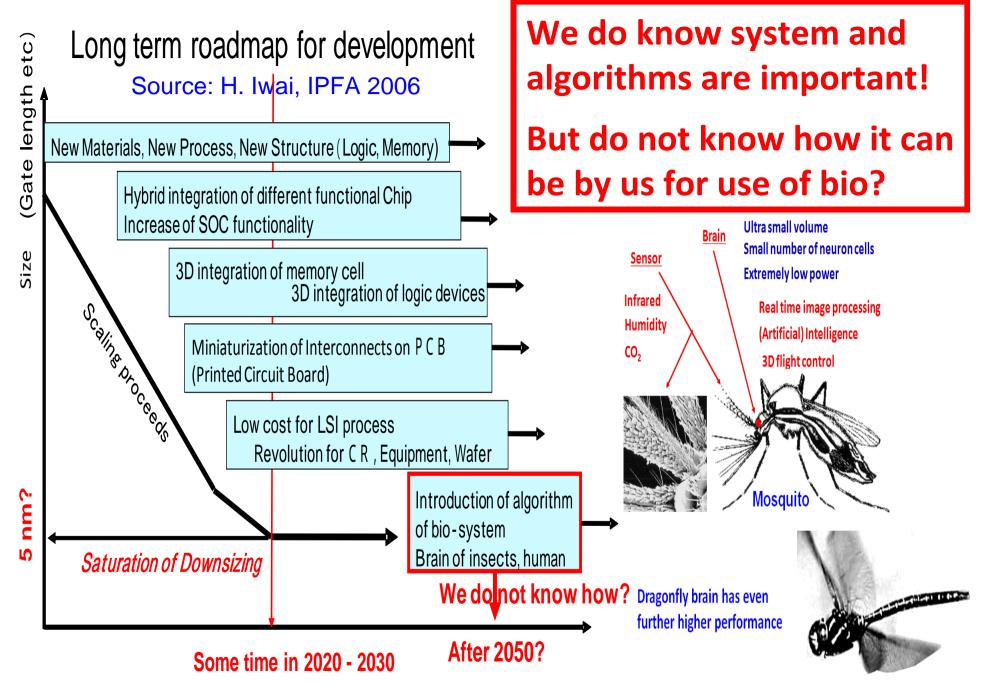
#### Current Issues <u>Si Nanowire</u>

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



## 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!