Si MOSFET Roadmap for 22nm and beyond

December 14, 2009

CODEC 2009@Hyatt Regency, Kolkata, India

Hiroshi Iwai
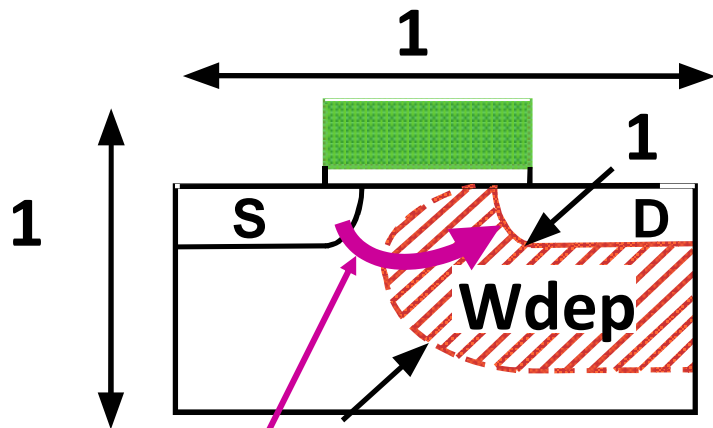
Tokyo Institute of Technology

Outline

1. Scaling

2. ITRS Roadmap

3. Voltage Scaling/ Low Power and Leakage

4. SRAM Cell Scaling

5.Roadmap for further future
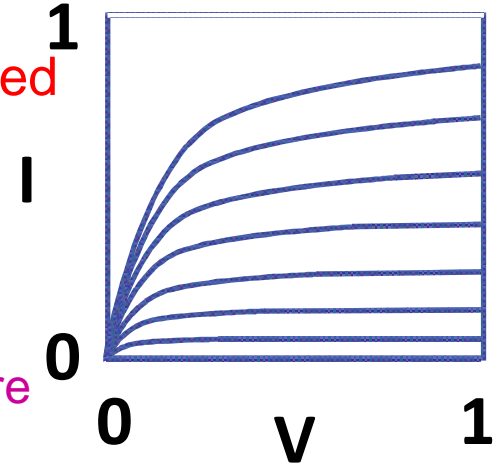
1. Scaling

Scaling Method: by R. Dennard in 1974

Wdep: Space Charge Region (or Depletion Region) Width

1

S   D

Wdep

Leakage current

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.
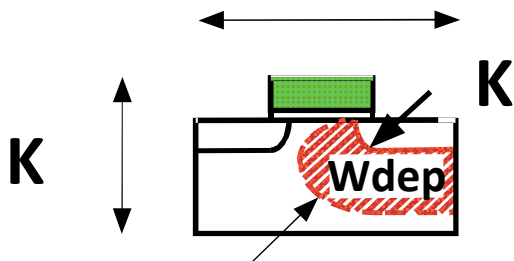
I

0   V   1

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.

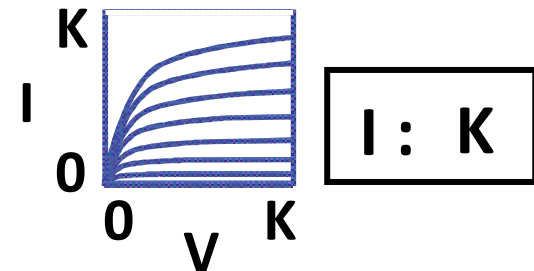$\longrightarrow$ Good scaled I-V characteristics

K

K   Wdep   K

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

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 $\longrightarrow W_g (t_{ox}^{-1})(V_g - V_{th}) = W_g t_{ox}^{-1}(V_g - V_{th}) = KK^{-1}K = K$
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	$C_g = \varepsilon_o \varepsilon_{ox} L_g W_g / t_{ox} \longrightarrow KK/K = K$
Switching speed	τ	K	$\tau = C_g V_{dd}/I_d \longrightarrow KK/K = K$
Clock frequency	f	1/K	$f = 1/\tau = 1/K$
Chip area	A_{chip}	α	**α: Scaling factor** $\longrightarrow$ **In the past, $\alpha>1$ for most cases**
Integration (# of Tr)	N	α/K^2	$N \longrightarrow \alpha/K^2$ **= 1/K^2 , when $\alpha=1$**
Power per chip	P	α	$fNCV^2/2 \longrightarrow K^{-1}(\alpha K^{-2})K(K^1)^2 = \alpha$ **= 1, when $\alpha=1$**

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

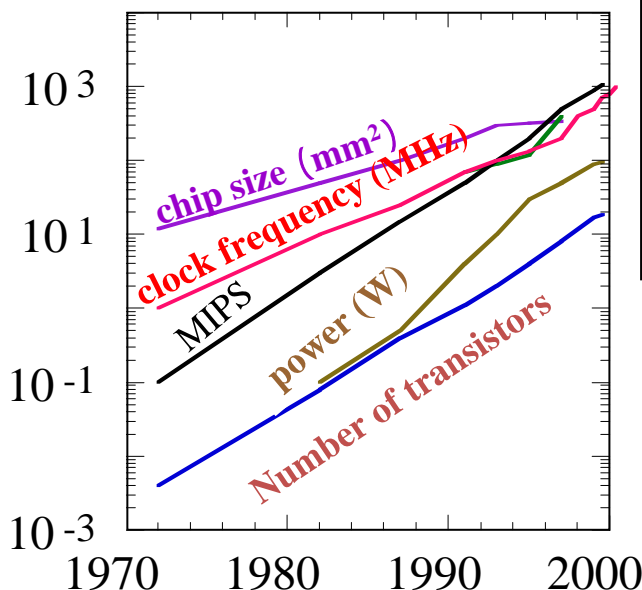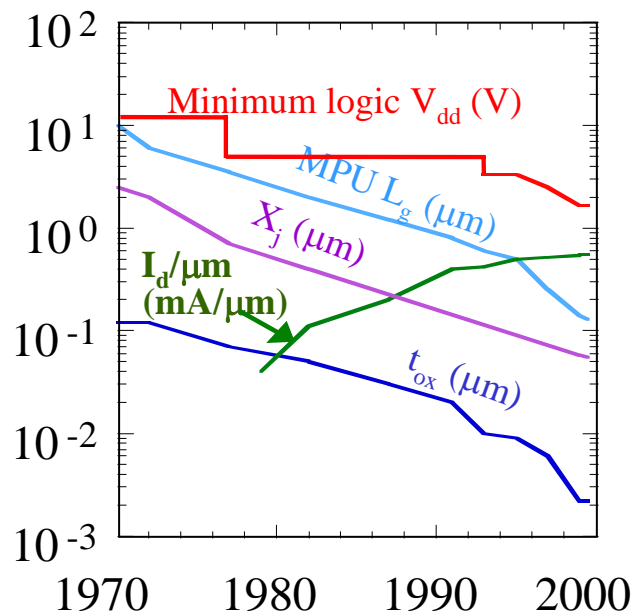- Effort for the down-scaling has to be continued by all means.

---

*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 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

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}						
V_{dd}	$K(10^{-2})$	10^{-1}	$I_d/\mu m$	1	10^1	P	$\alpha(10^1)$	10^5
A_{chip}	α	10^1	N	$\alpha/K^2(10^5)$	10^4	$= f\alpha NCV^2$		

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

8

- 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)

11
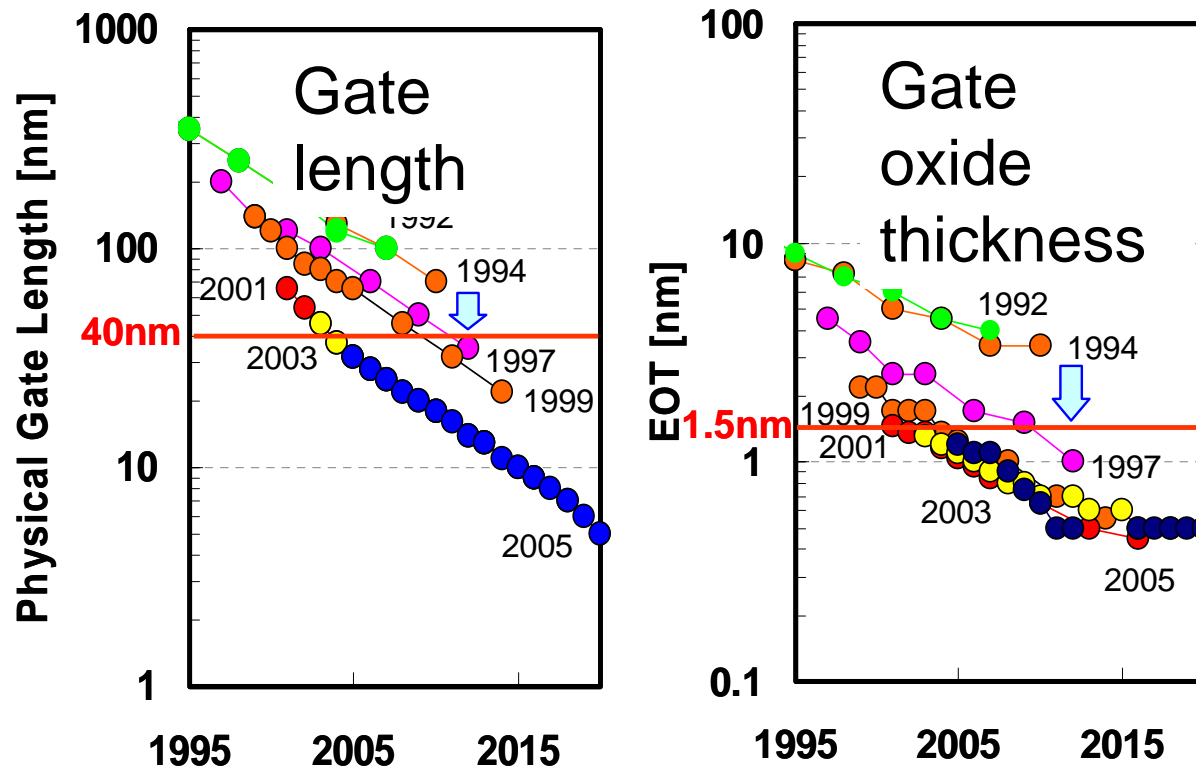
1992 -1997:NTRS (National Technology Roadmap)
1998 -   : ITRS (International Technology Roadmap)

2008 ITRS update

2007 ITRS

2006 ITRS update
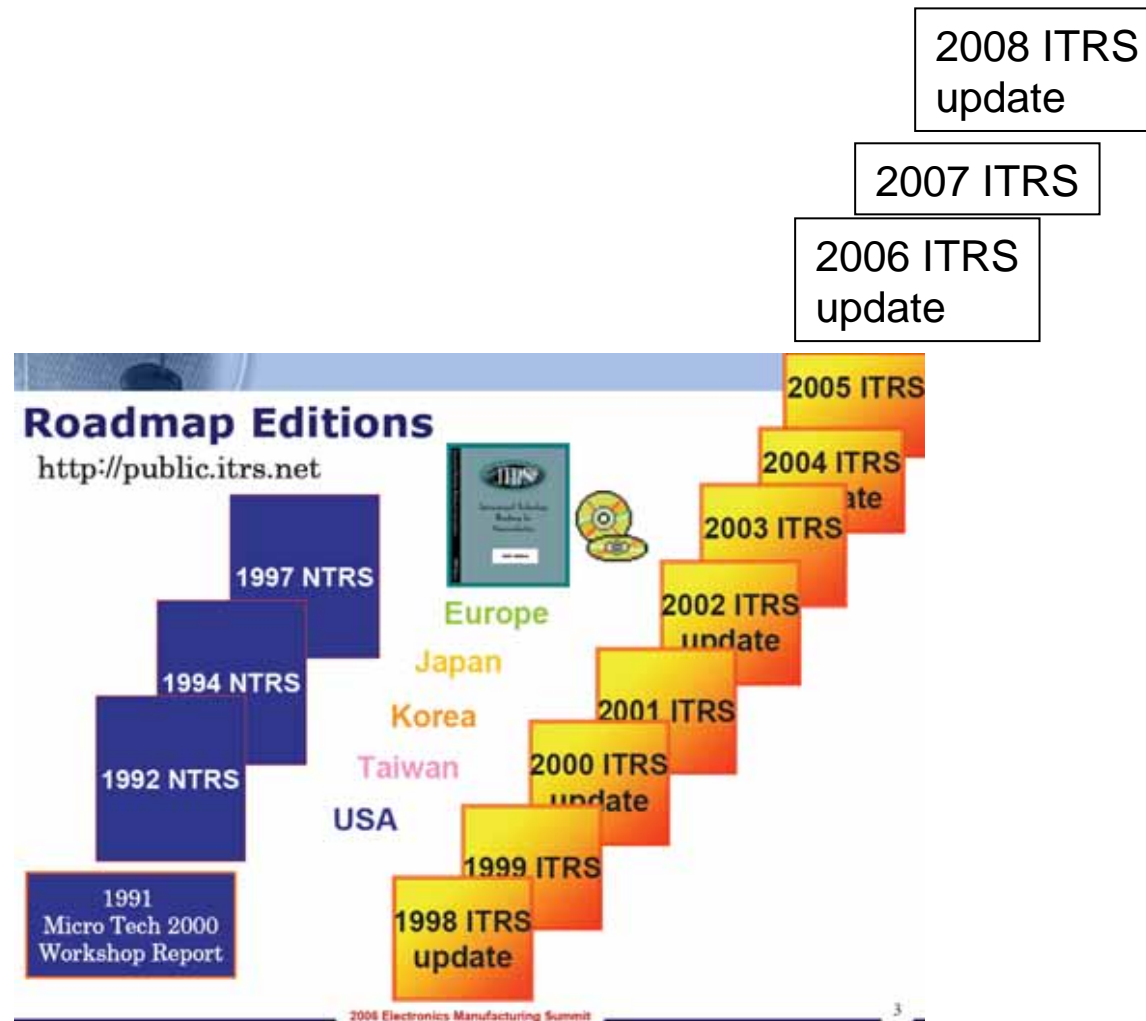
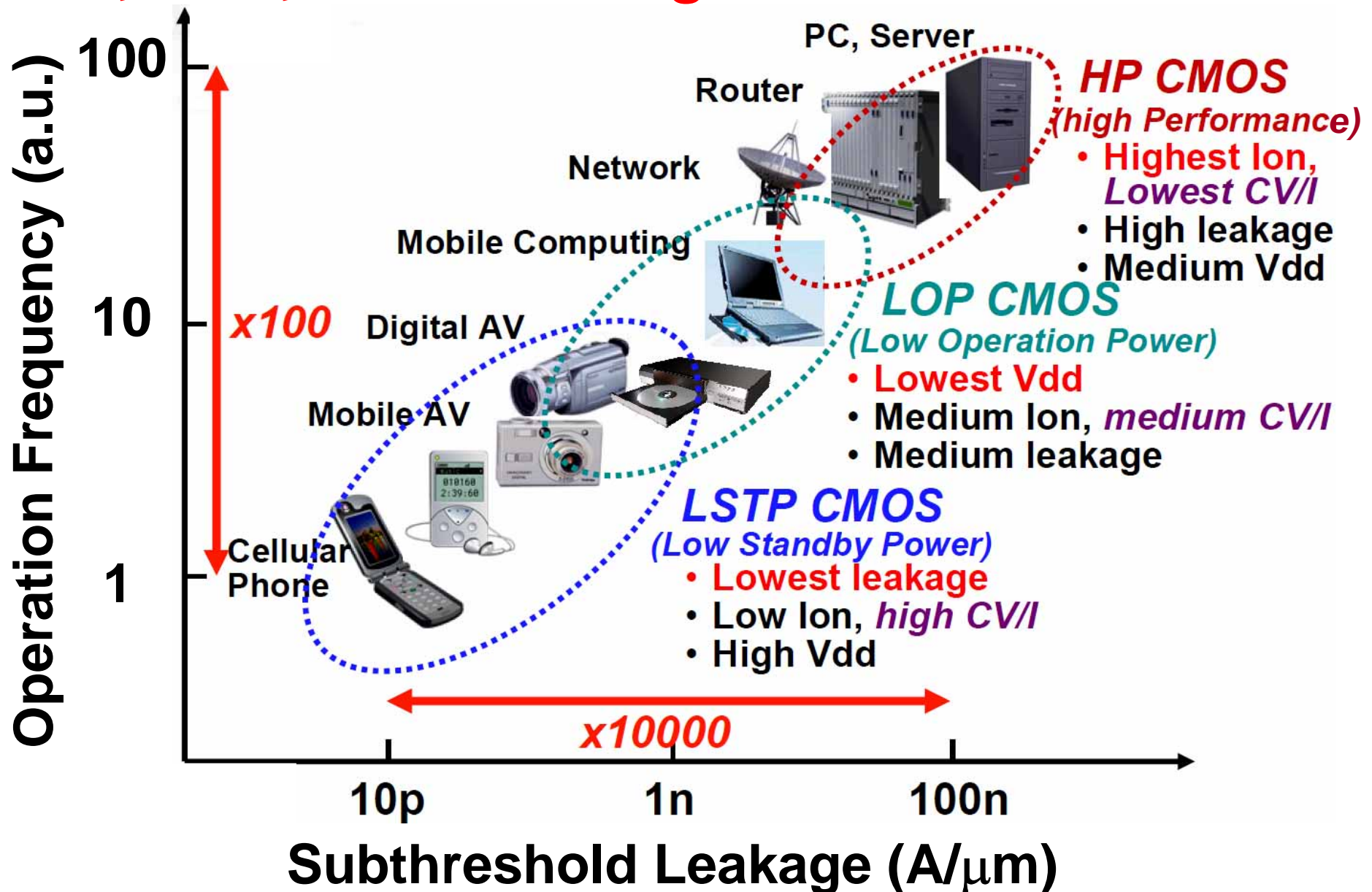# ITRS Roadmap does change every year!

2007 Edition   2003 Edition

2006 Update   2002 Update

2005 Edition   2001 Edition

2004 Update   2000 Update

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

HP, LOP, LSTP for Logic CMOS

Operation Frequency (a.u.) vs **Subthreshold Leakage (A/μm)**

x100 (vertical axis range)

x10000 (horizontal axis range)

Devices: PC, Server; Router; Network; Mobile Computing; Digital AV; Mobile AV; Cellular Phone

HP CMOS (high Performance)
- Highest Ion, Lowest CV/I
- High leakage
- Medium Vdd

LOP CMOS (Low Operation Power)
- Lowest Vdd
- Medium Ion, medium CV/I
- Medium leakage

LSTP CMOS (Low Standby Power)
- Lowest leakage
- Low Ion, high CV/I
- High Vdd

Axis values: 100, 10, 1 (vertical); 10p, 1n, 100n (horizontal)

14

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

'XX nm CMOS Technology
Commercial Logic CMOS products

Technology name	Starting Year
45 nm	2007
32 nm	2009?
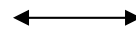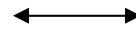
ITRS 2008 Update
for High Performance Logic

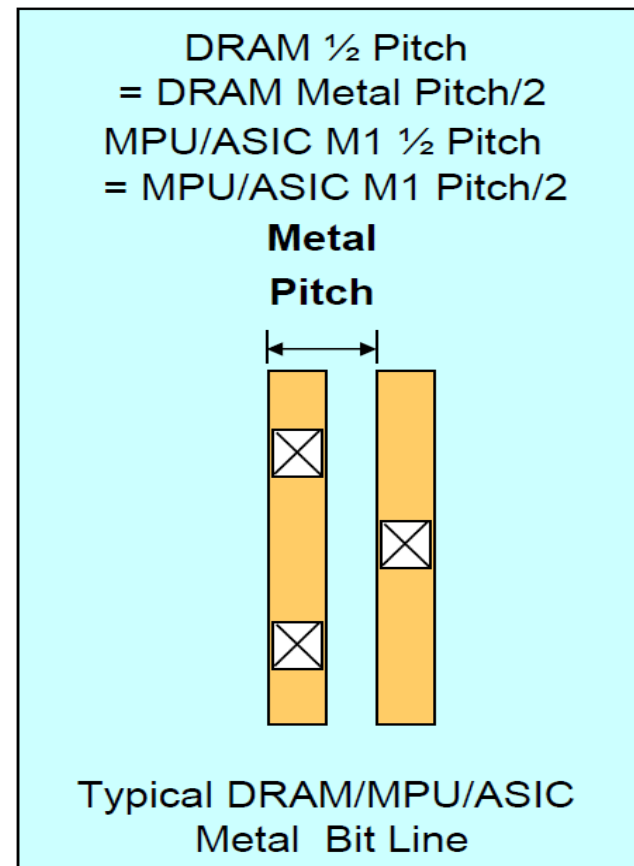| Year	Half Pitch (1st Metal)	Physical Gate Length
2007	68 nm	32 nm
2008	59 nm	29 nm
2009	52 nm	27 nm
2010	45 nm	24 nm

'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?

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.

Logic 1st Metal
Half Pitch

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

Metal

Pitch

Typical DRAM/MPU/ASIC
Metal Bit Line

16

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

→ **350nm → 250nm → 180nm → 130nm → 90nm → 65nm → 45nm**
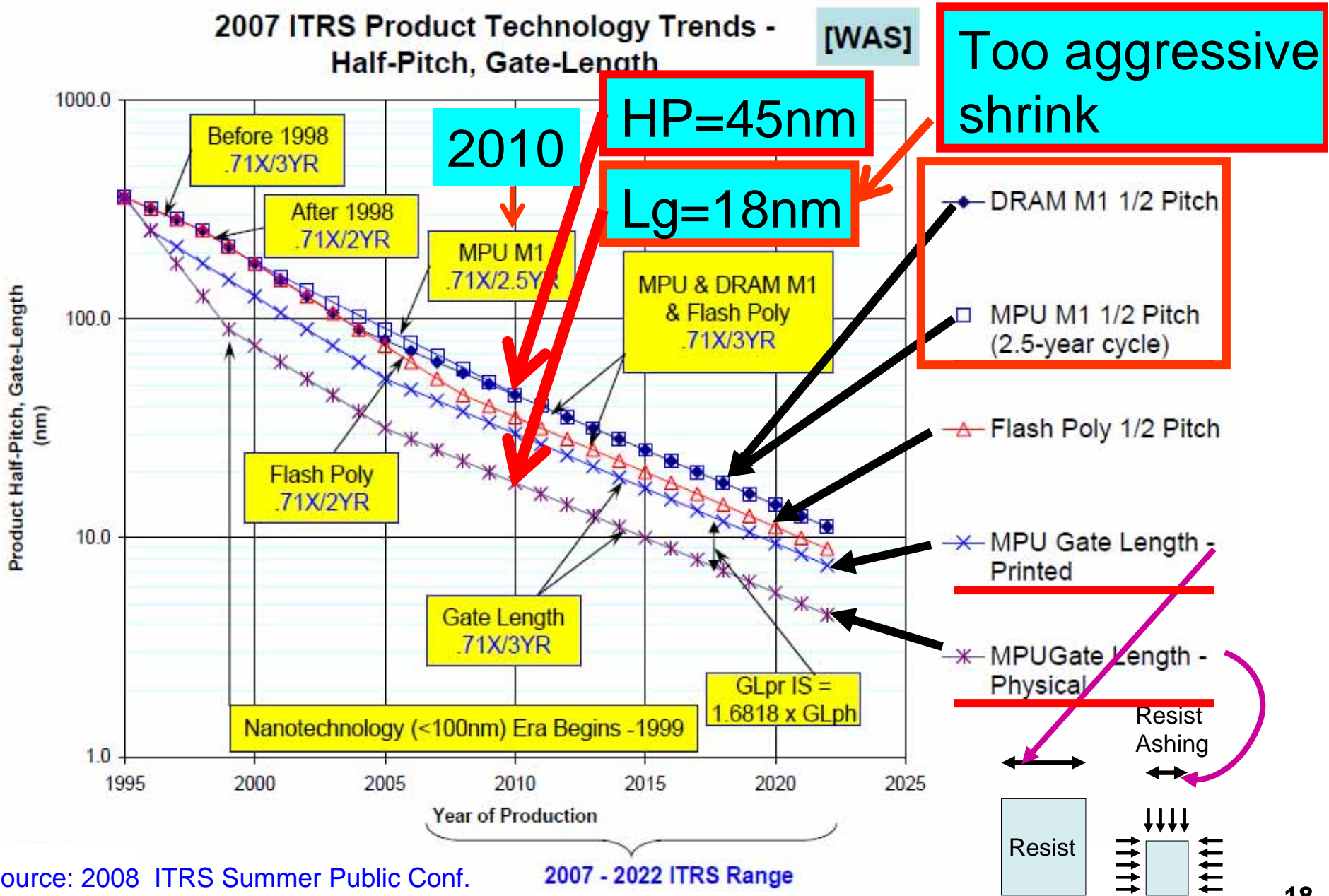
- 'XX' values were established by NTRS* and ITRS with the term of 'Technology Node**' and 'Cycle***' using typical 'half pitch value'.

- 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.

Resist
Ashing

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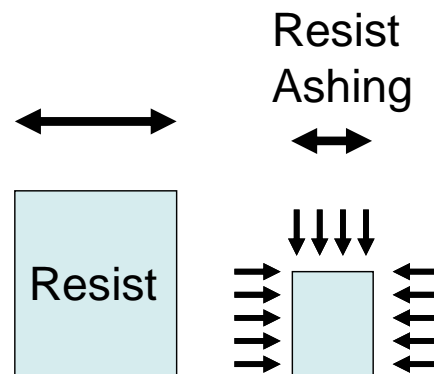
- Memory still keeps the half pitch as the value of 'XX'

For example, Typical Half Pitches at ITRS 2007

2007 ITRS Product Technology Trends - Half-Pitch, Gate-Length [WAS]

2010

HP=45nm

Lg=18nm

Too aggressive shrink

Before 1998 .71X/3YR

After 1998 .71X/2YR

MPU M1 .71X/2.5YR

MPU & DRAM M1 & Flash Poly .71X/3YR

Flash Poly .71X/2YR

Gate Length .71X/3YR

Nanotechnology (<100nm) Era Begins -1999

GLpr IS = 1.6818 x GLph

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

Resist Ashing

Resist

Year of Production

2007 - 2022 ITRS Range

Product Half-Pitch, Gate-Length (nm)

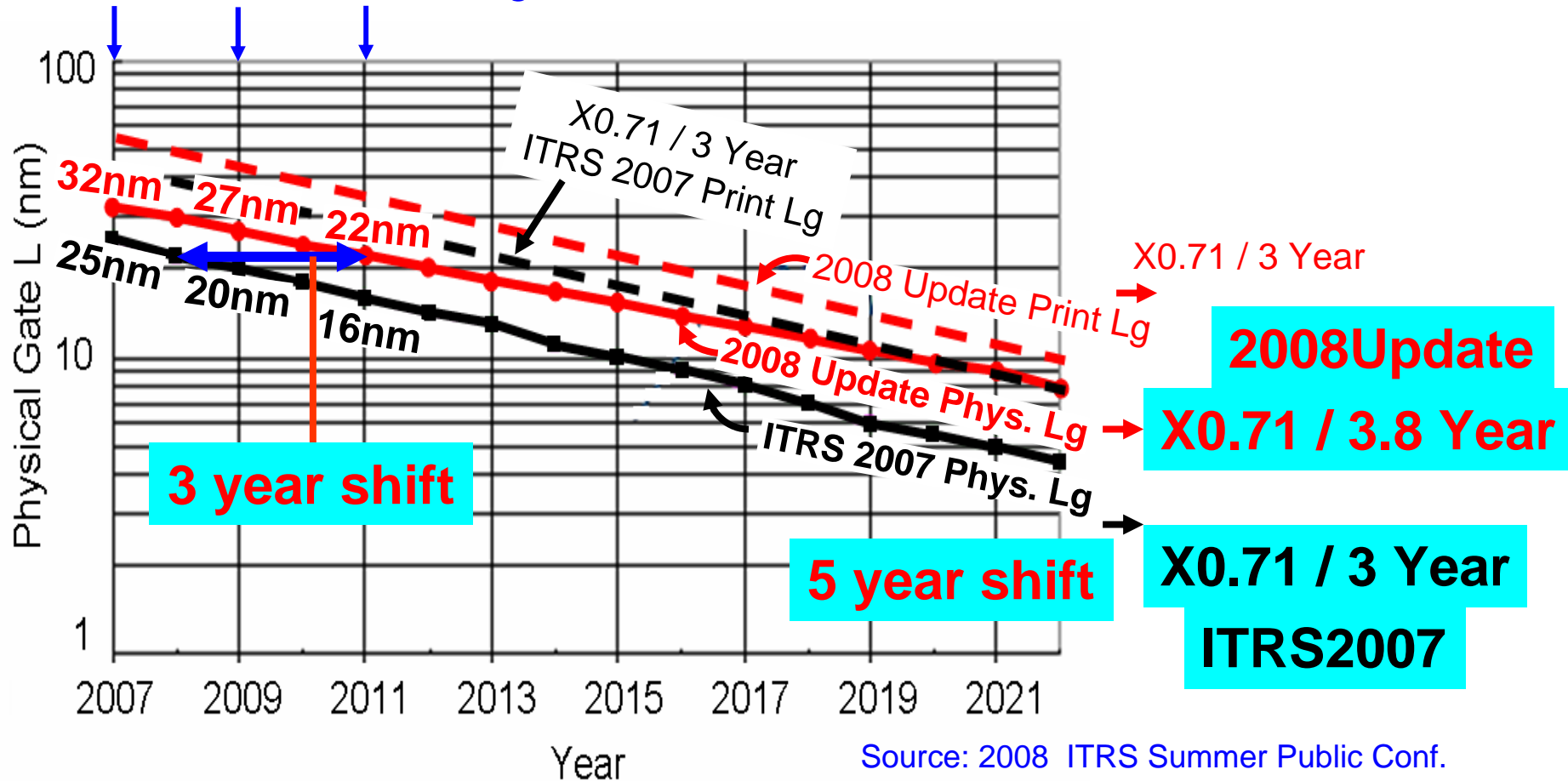Source: 2008 ITRS Summer Public Conf.

18

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.

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 (EOT w/3E20 poly row)

Likely in 2008 Update (EOT w/metal gate row)

Likely in 2008 Update (Drain Ext. Xj row)

non-steady trend corrected

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

20

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.

From ITRS2008 Update, maybe XX nm stands for the physical Gate length
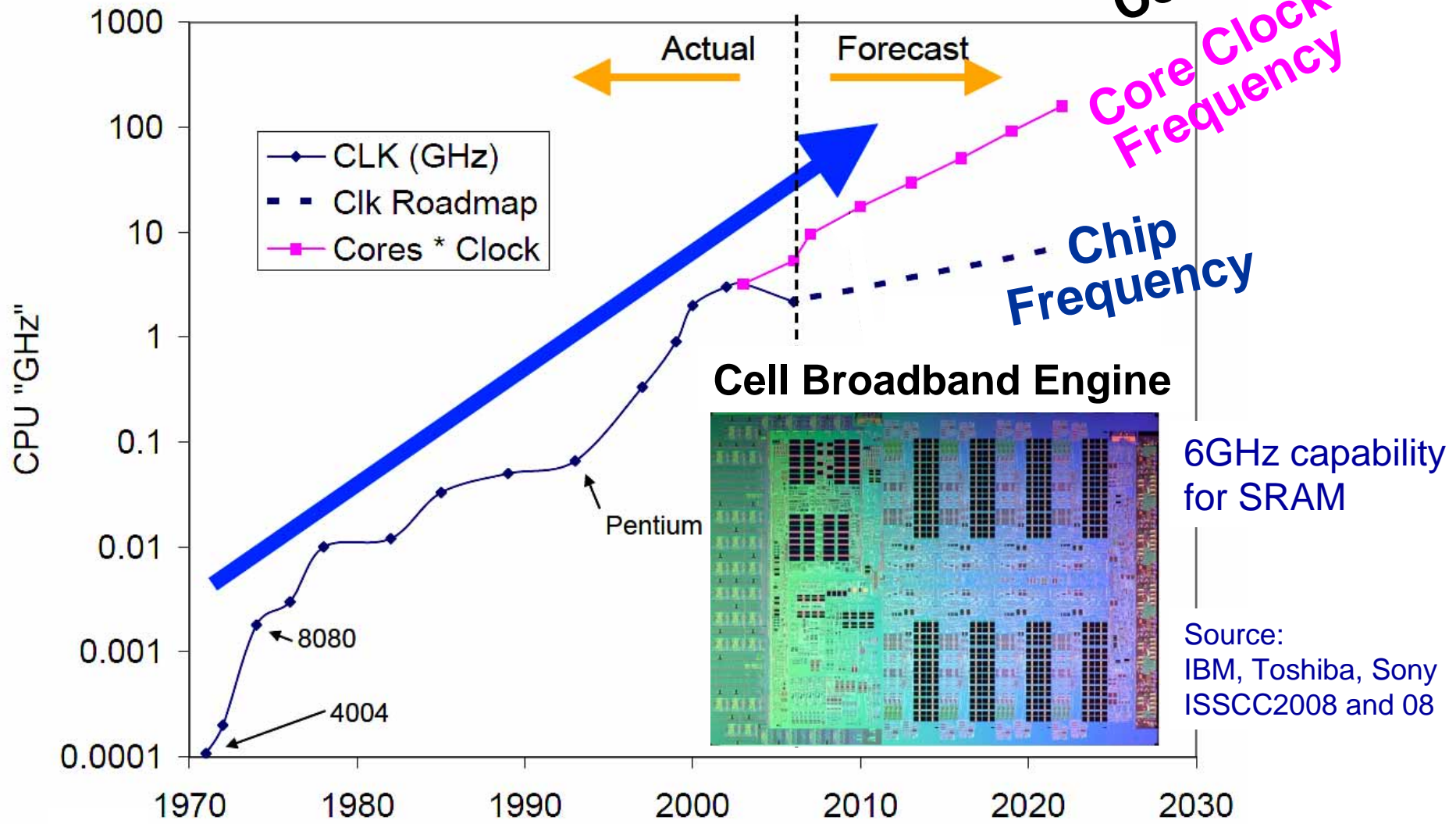
Clock frequency does not increase aggressively anymore.
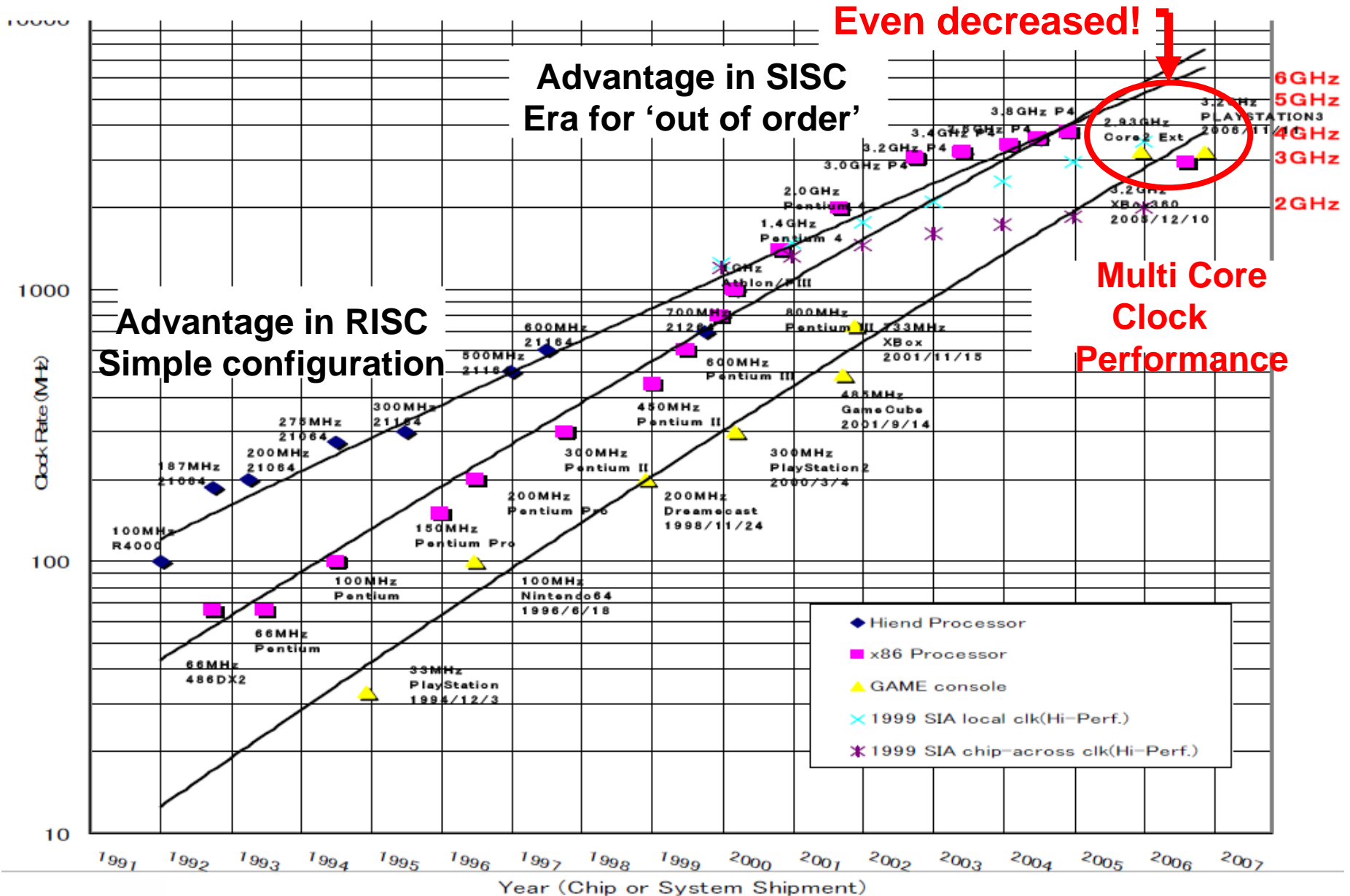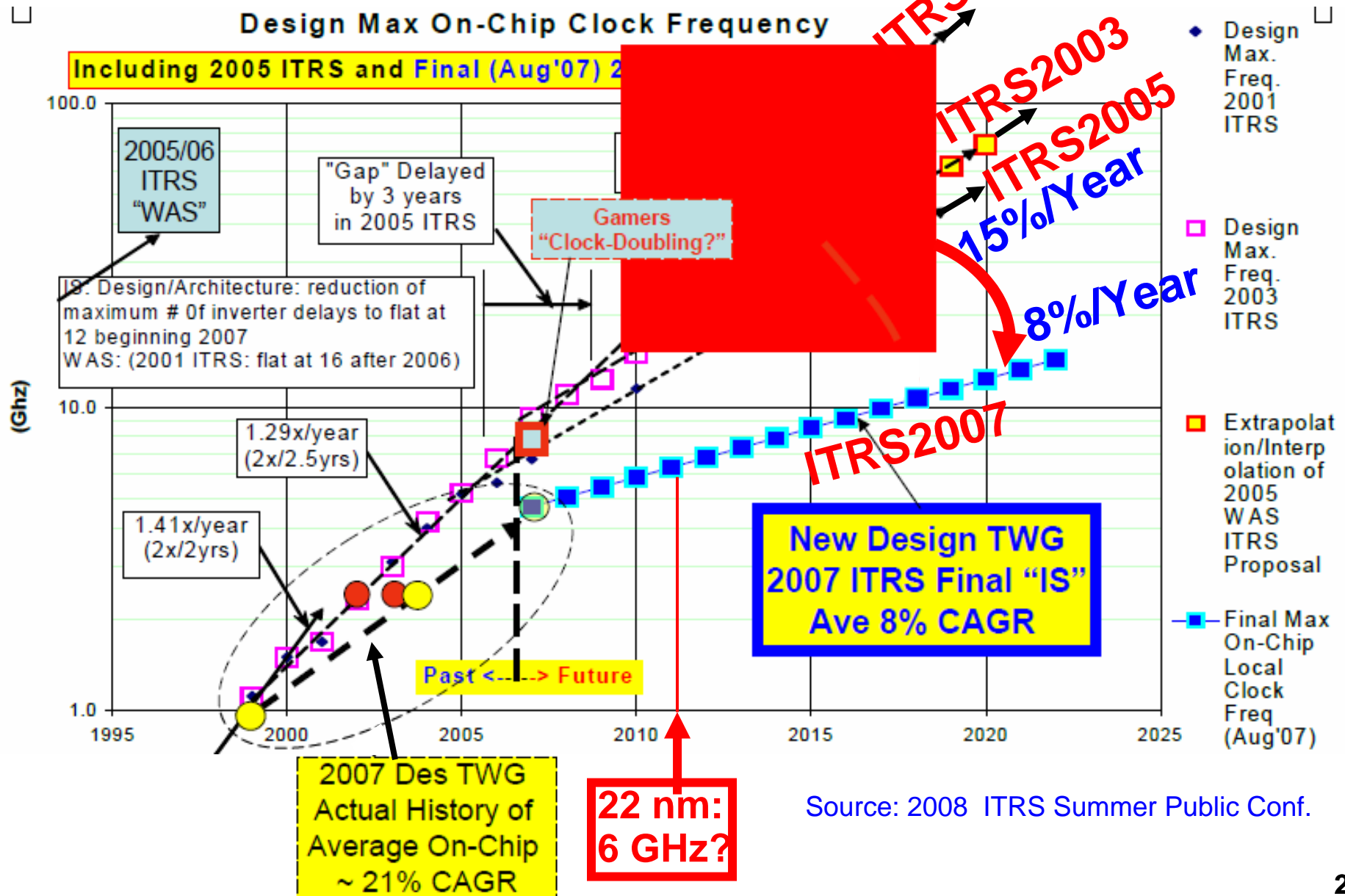

Source: Mitsuo Saito, Toshiba

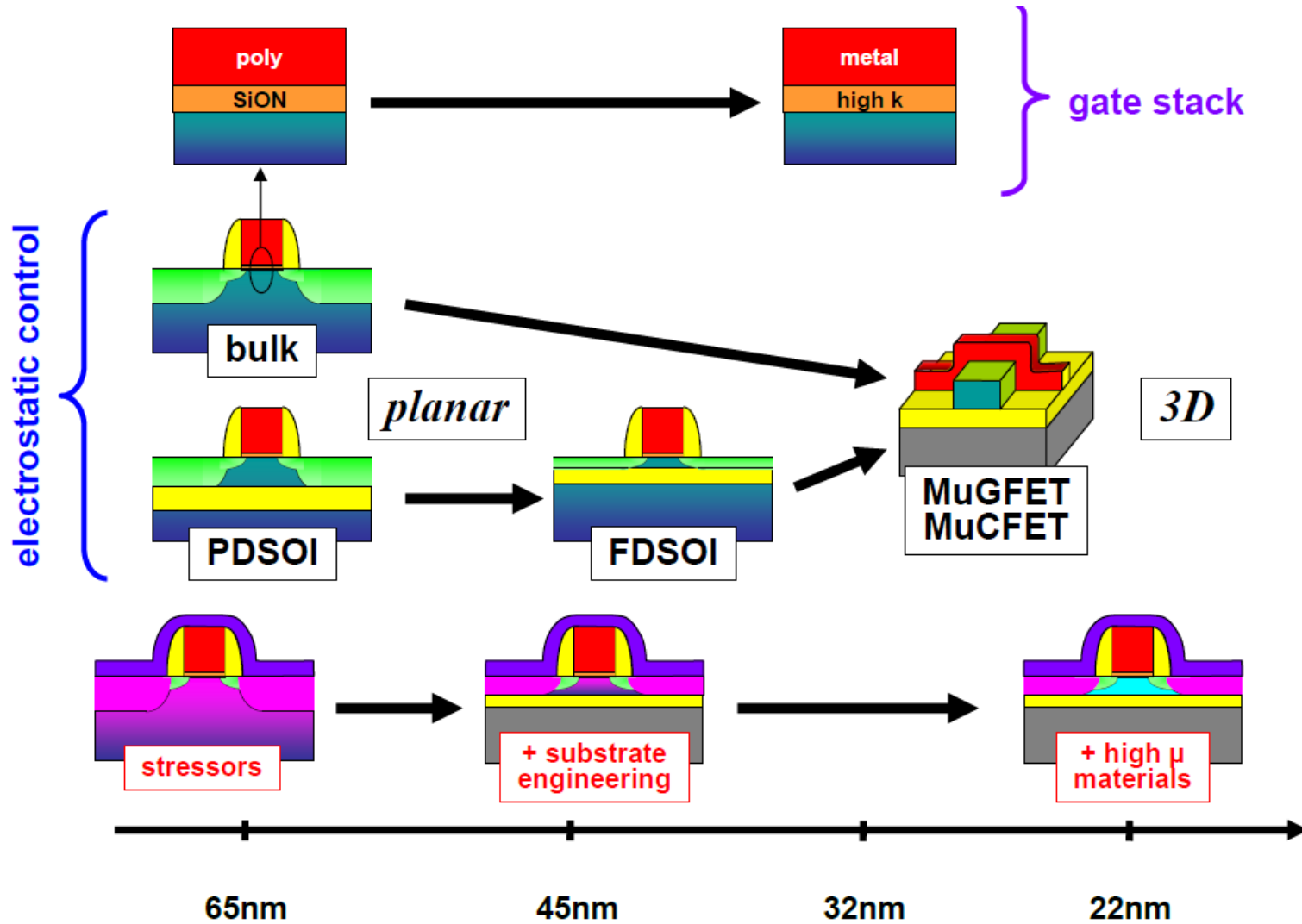# MPU "GHz" by "Cores" ITRS2007

Continued?

Core Clock Frequency

Chip Frequency

Cell Broadband Engine

6GHz capability for SRAM

Source:
IBM, Toshiba, Sony
ISSCC2008 and 08

Legend:
- CLK (GHz)
- Clk Roadmap
- Cores * Clock

Actual | Forecast

Pentium

8080

4004

CPU "GHz" axis: 1000, 100, 10, 1, 0.1, 0.01, 0.001, 0.0001

Year axis: 1970, 1980, 1990, 2000, 2010, 2020, 2030

Source: 2007 ITRS Winter Public Conf.

Max on chip frequency or 'Core clock'

Design Max On-Chip Clock Frequency

Including 2005 ITRS and Final (Aug'07) 2

ITRS2001

ITRS2003

ITRS2005

ITRS2007

15%/Year

8%/Year

100.0

2005/06 ITRS "WAS"

"Gap" Delayed by 3 years in 2005 ITRS

Gamers "Clock-Doubling?"

IS: Design/Architecture: reduction of maximum # 0f inverter delays to flat at 12 beginning 2007
WAS: (2001 ITRS: flat at 16 after 2006)

(Ghz)

10.0

1.29x/year (2x/2.5yrs)

1.41x/year (2x/2yrs)

New Design TWG 2007 ITRS Final "IS" Ave 8% CAGR

Past <---> Future

1.0

1995   2000   2005   2010   2015   2020   2025

Design Max. Freq. 2001 ITRS

Design Max. Freq. 2003 ITRS

Extrapolation/Interpolation of 2005 WAS ITRS Proposal

Final Max On-Chip Local Clock Freq (Aug'07)

2007 Des TWG Actual History of Average On-Chip ~ 21% CAGR

22 nm: 6 GHz?

Source: 2008 ITRS Summer Public Conf.

24

Structure and technology innovation (ITRS 2007)

Source: 2008 ITRS Summer Public Conf.

Timing of CMOS innovations shifts backward.

Bulk CMOS has longer life now!

Correspond to 22nm Logic CMOS

2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017 2018 2019

Bulk extends 4 years!

HP bulk CMOS (ITRS 2007)
HP bulk CMOS (ITRS 2008)
HP UTB/FDSOI (ITRS 2007)
HP UTB/FDSOI (ITRS 2008)
HP Multi-Gate (ITRS 2007)
HP Multi-Gate (ITRS 2008)

Multi G delays 4 years!

LOP bulk CMOS (ITRS 2007)
LOP bulk CMOS (ITRS 2008)
LOP UTB/FDSOI (ITRS 2007)
LOP UTB/FDSOI (ITRS 2008)
LOP Multi-Gate (ITRS 2007)
LOP Multi-Gate (ITRS 2008)
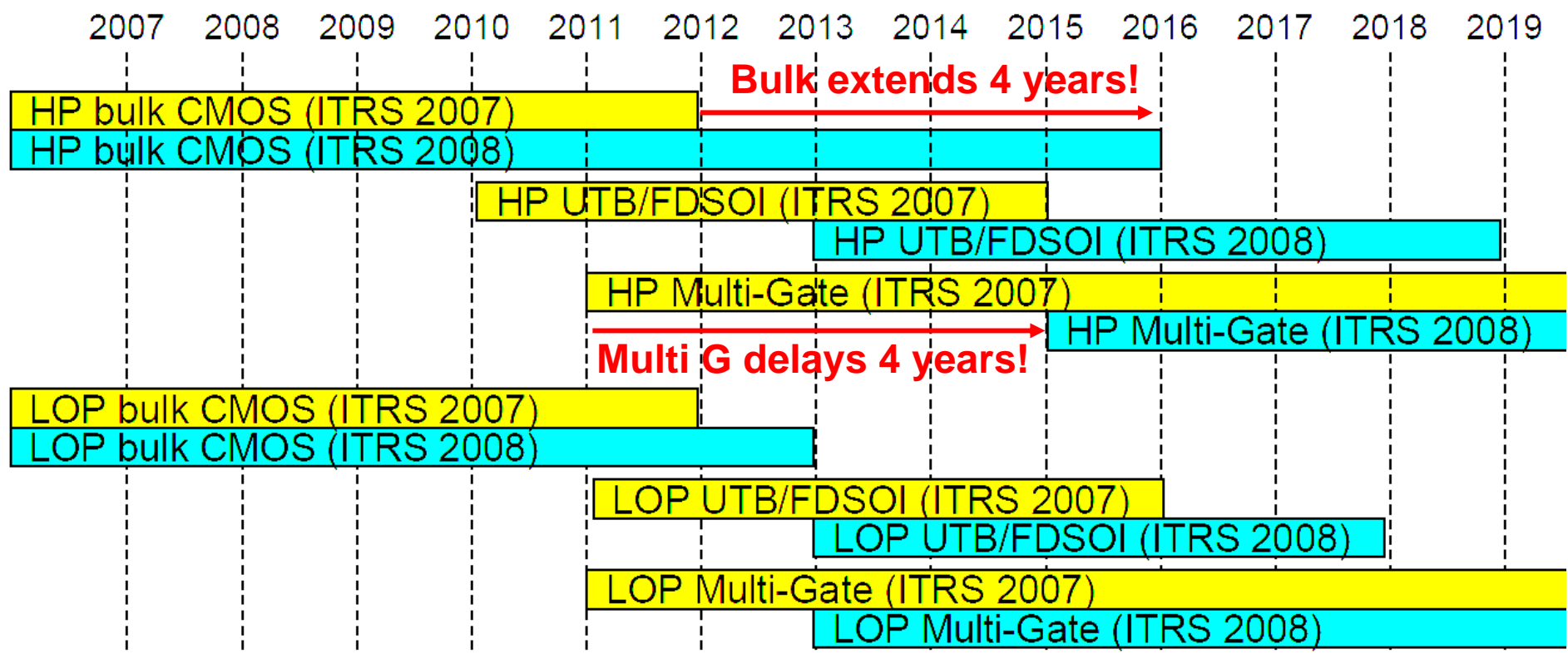
Source: 2008 ITRS Summer Public Conf.

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??

ITRS2008 Low-k Roadmap Update

Correspond to 22nm Logic

	Year of Production	Near-term	2009	2010	2011	2012	2013
		2008					
ITRS 2007	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
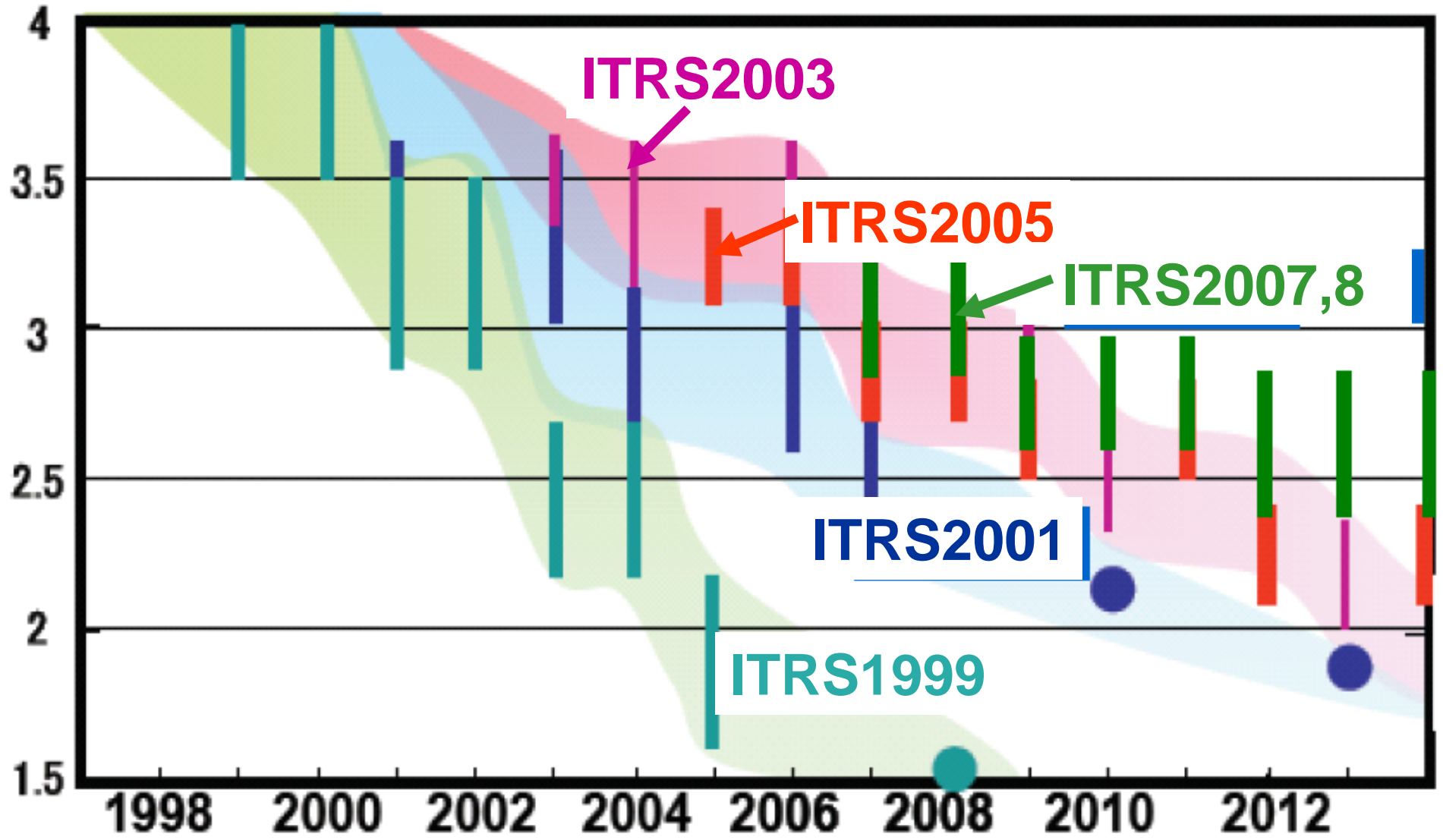
Source: 2008 ITRS Summer Public Conf.

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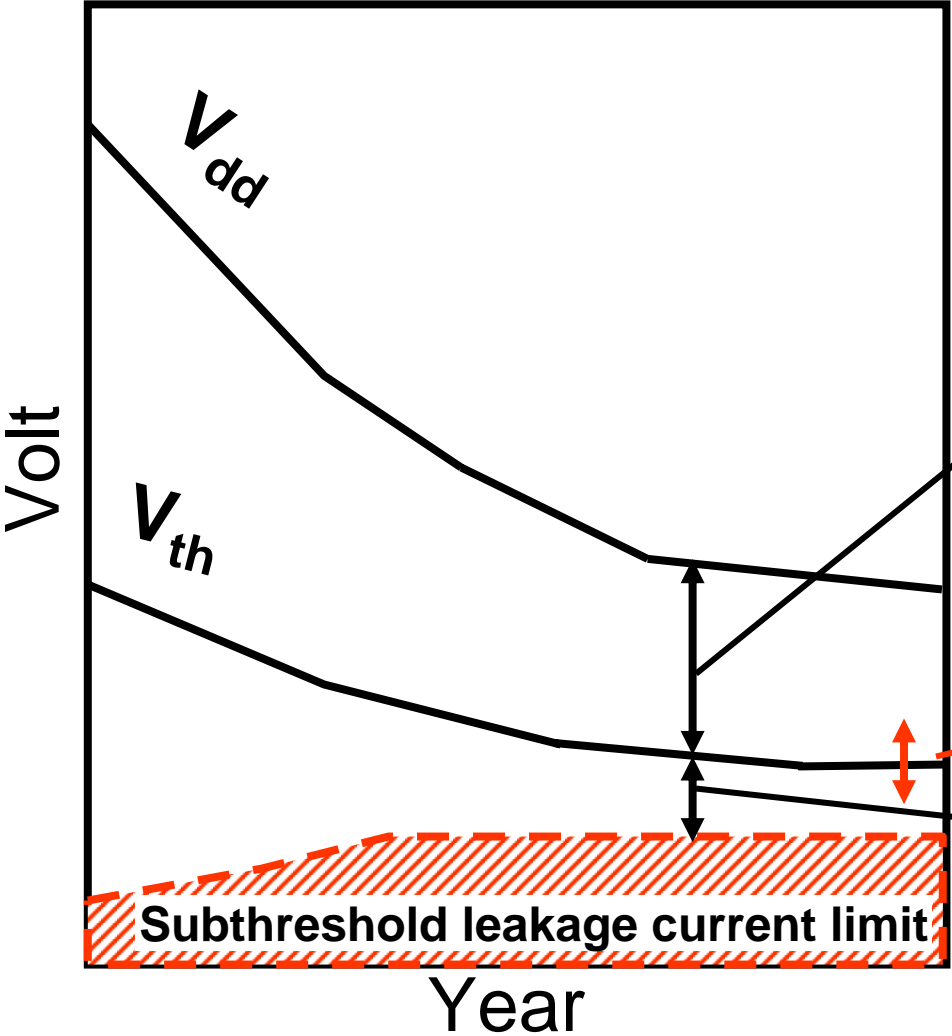
Because, V_{th} cannot be down-scaled anymore, V_{dd} down-scaling is difficult.

$V_{dd} - V_{th}$ 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

Subthreshold leakage current limit

V_{dd}

V_{th}

Volt

Year

Subtheshold leakage current of MOSFET

Id

Ion

OFF   ON

Subthreshould
Leakage Current

Ioff

Vg=0V

Subthreshold
region

**Vth
(Threshold Voltage)**

Subthreshold Current
Is OK at Single Tr. level

But not OK
For Billions of Trs.

Vg
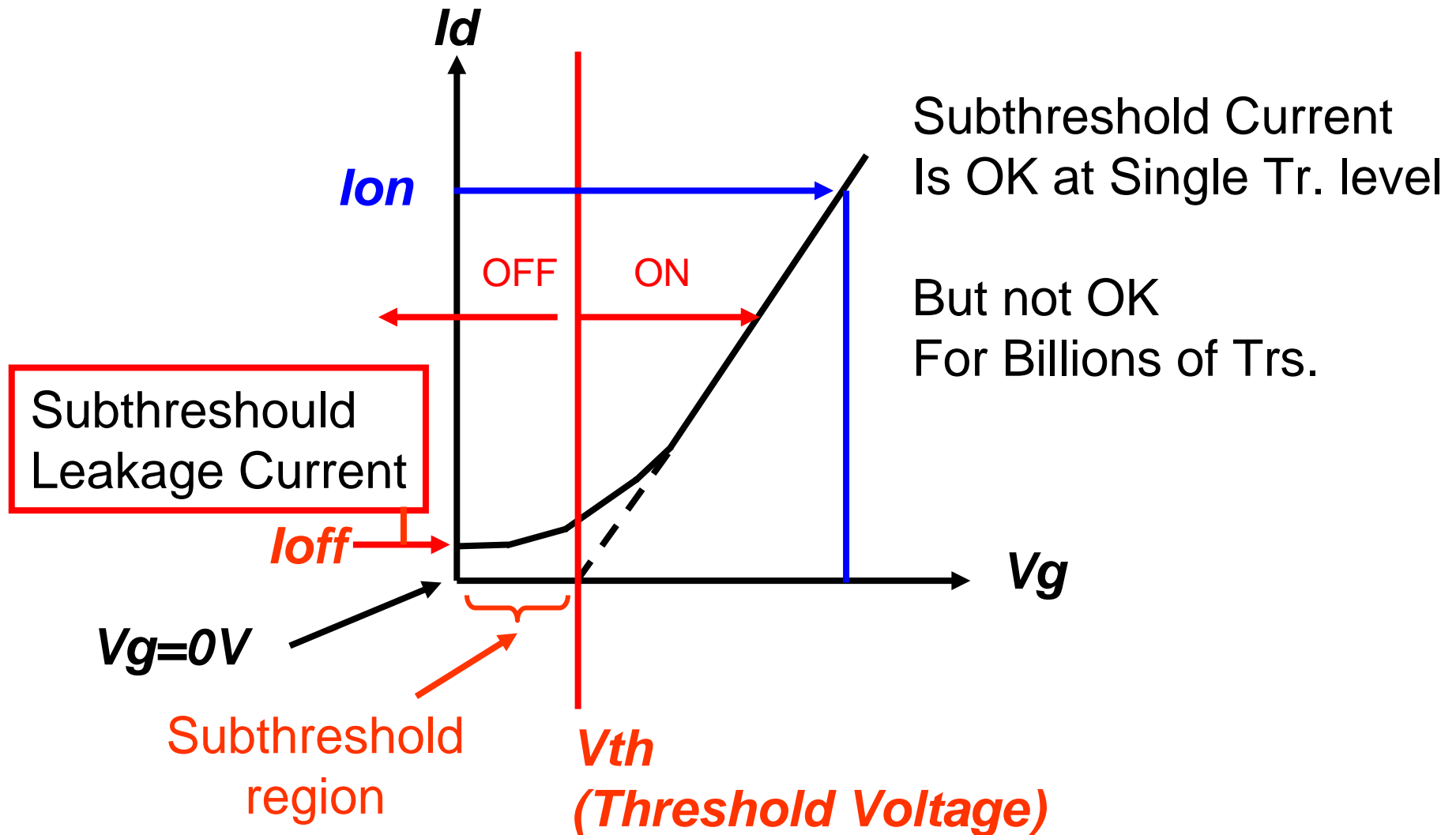
Vth cannot be decreased anymore

Log scale Id plot

Ion —

significant Ioff increase

Ioff

Vth: **300mV** → **100mV**

Ioff increases
with **3.3 decades**

(300 − 100)mV/(60mv/dec)
= 3.3 dec

Ioff

Log Id per unit gate width (= 1μm)

$10^{-3}A$
$10^{-4}A$
$10^{-5}A$
$10^{-6}A$
$10^{-7}A$
$10^{-8}A$
$10^{-9}A$
$10^{-10}A$

Subthreshold slope (SS)
= $(Ln10)(kT/q)(C_{ox}+C_D+C_{it})/C_{ox}$
> ~ **60 mV/decade at RT**
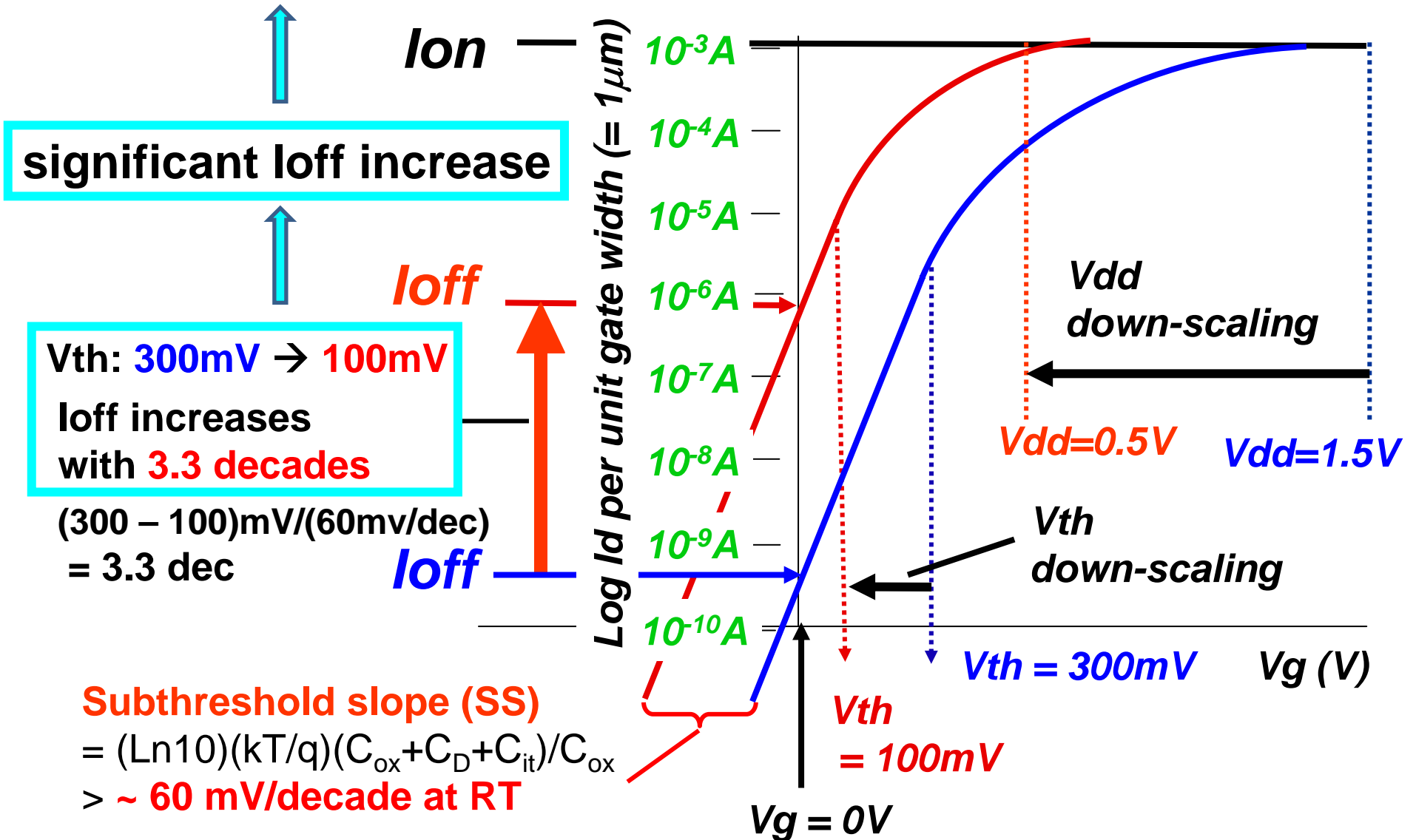
SS value:
Constant and does not become small with down-scaling

**Vdd
down-scaling**

Vdd=0.5V   *Vdd=1.5V*

Vth
down-scaling

Vth = 300mV   *Vg (V)*

*Vth
= 100mV*

Vg = 0V

ITRS for HP logic

Ion/Ioff ratio

ITRS for HP logic

Vdd

Vdd will stay higher in 2008 update

Vth

Vth will be around 0.1V
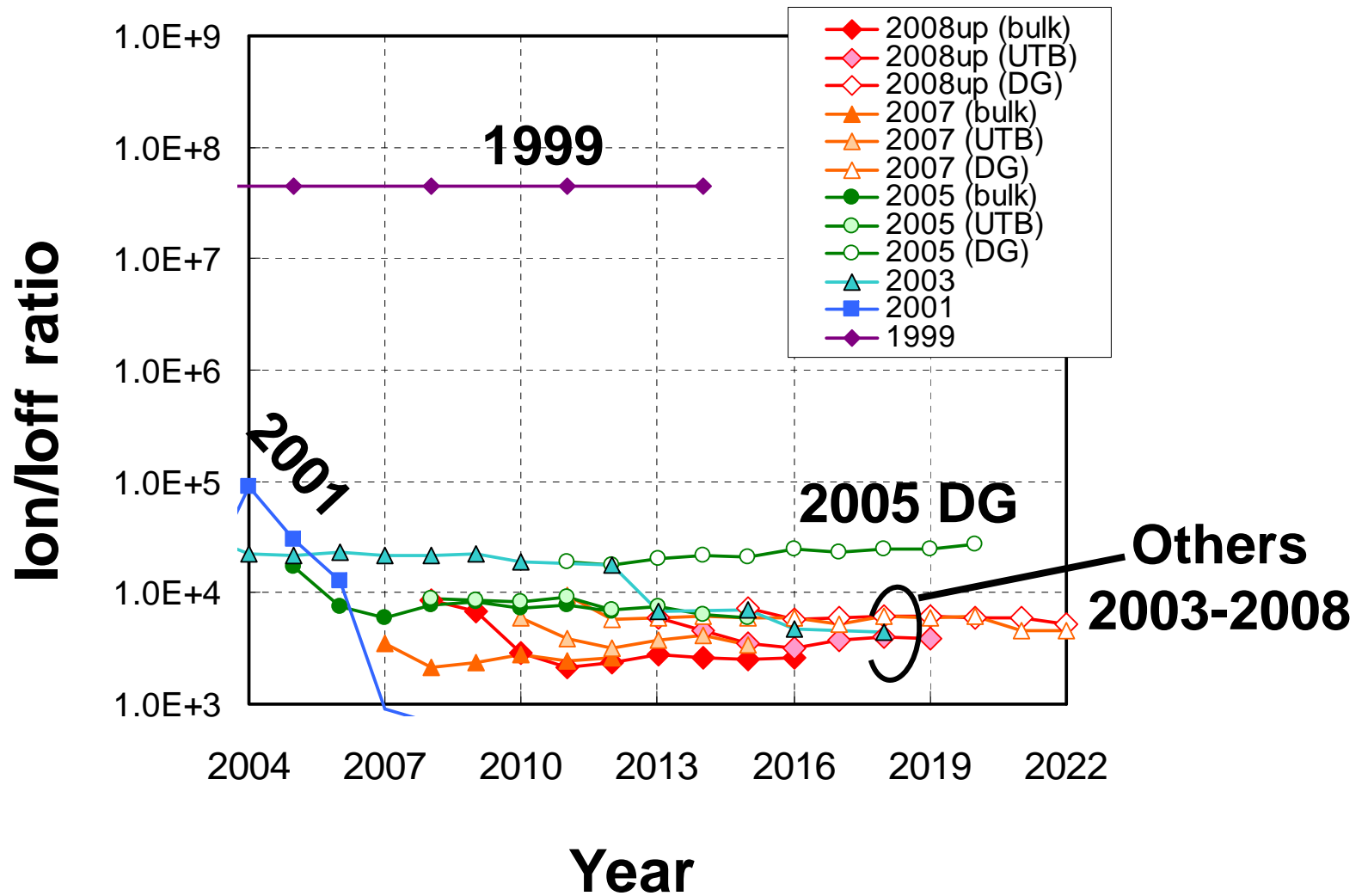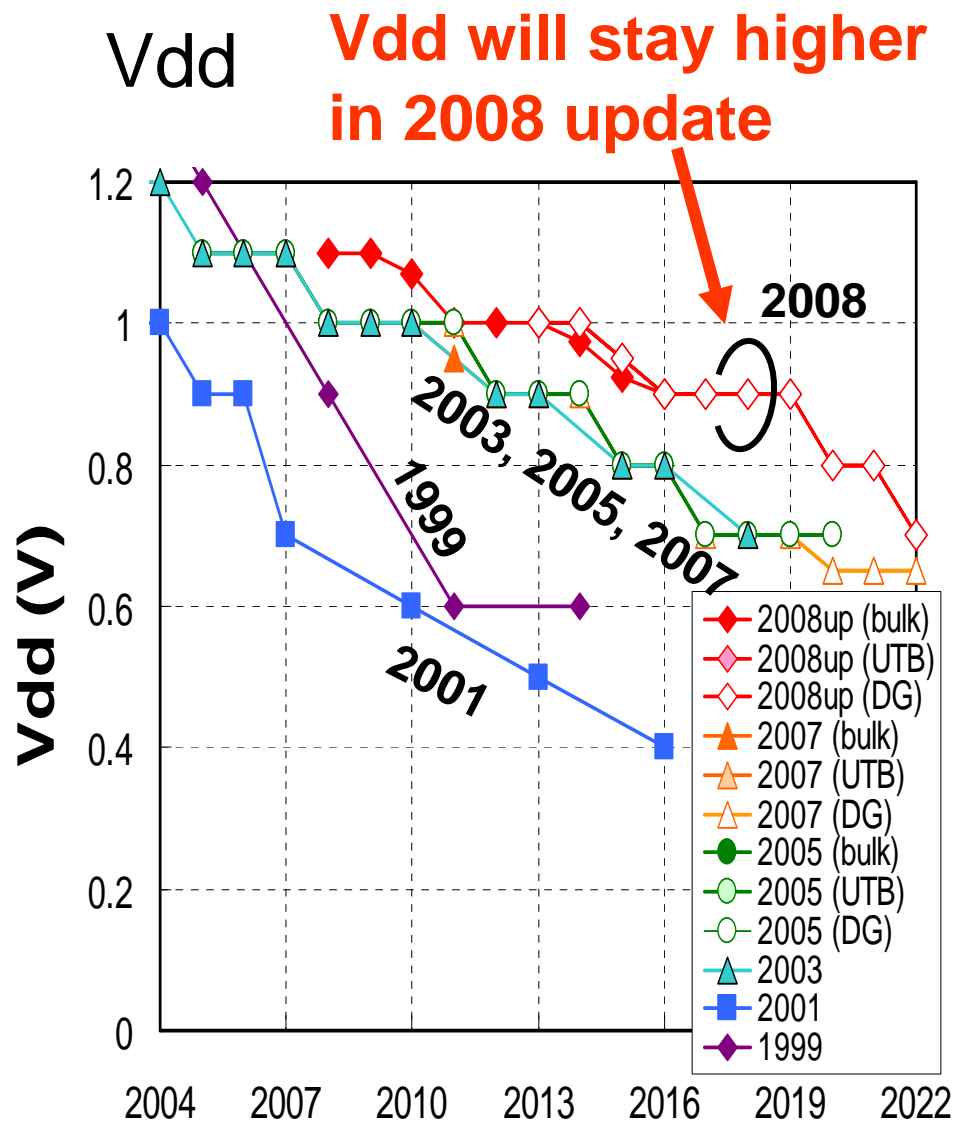
2008 Values are from ITRS Public Conf. and still under discussion

Source: ITRS and 2008 ITRS Summer Public Conf.

36
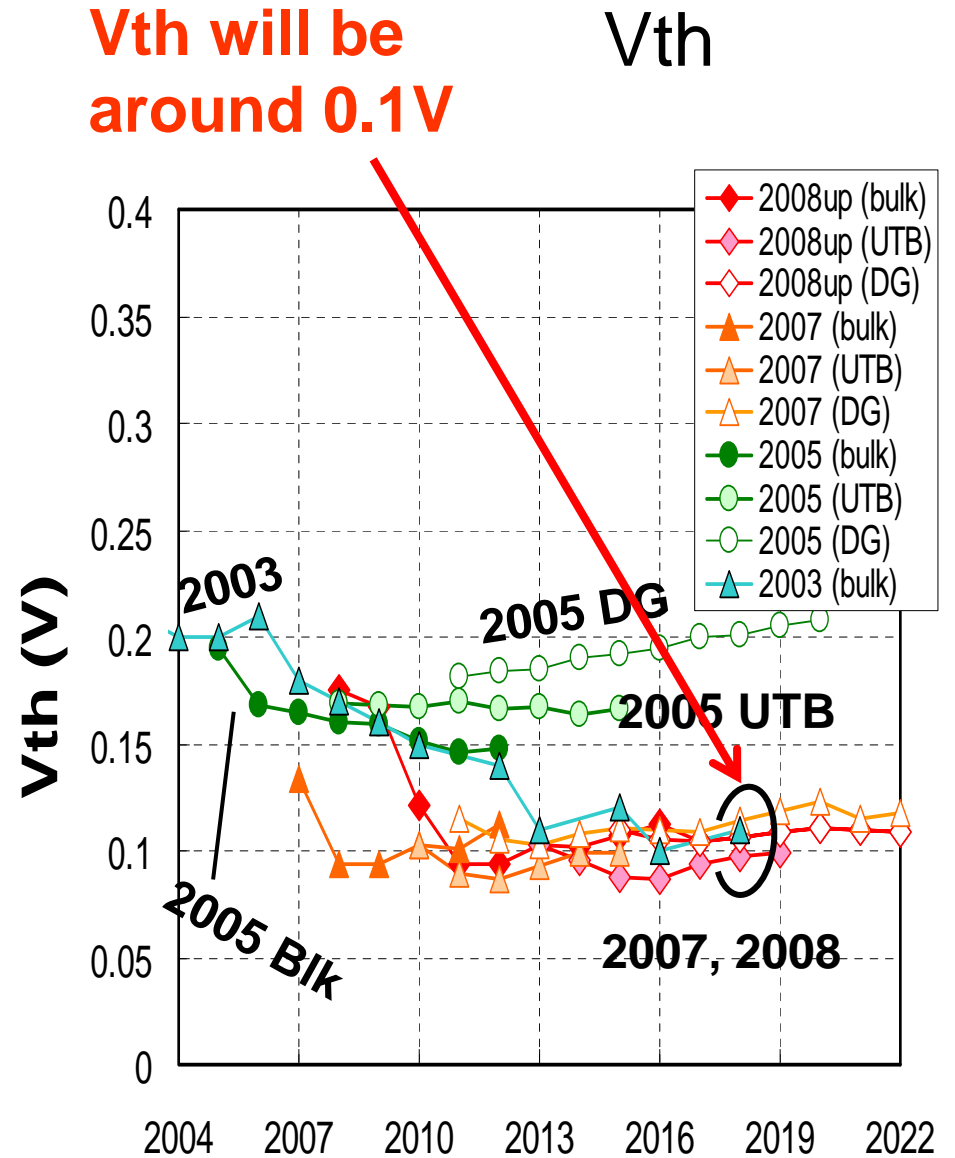
Improper down-scaling

Could we squeeze technologies for ultimate CMOS scaling?

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

Metal gate

Interfacial C (Quantum eff)

High-k oxd

Inversion C (Quantum eff)

Si

Interfacial C @Metal gate and Gate oxd. (EOT=0.2~0.3nm?)

C_1

C_2

C_3

Gate Oxd C

Inversion C (EOT=0.3~0.5nm?)
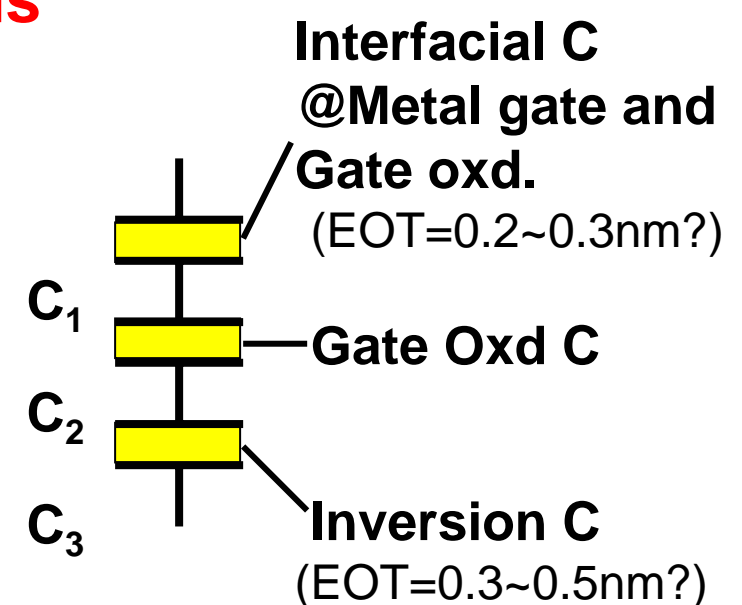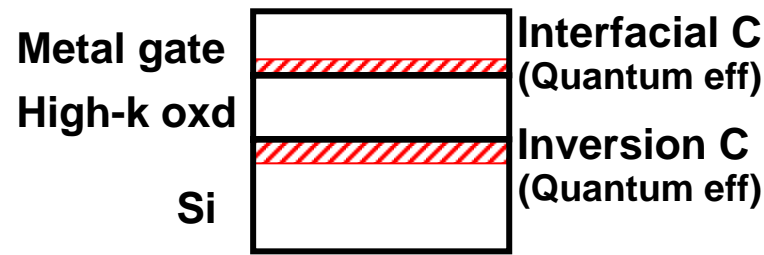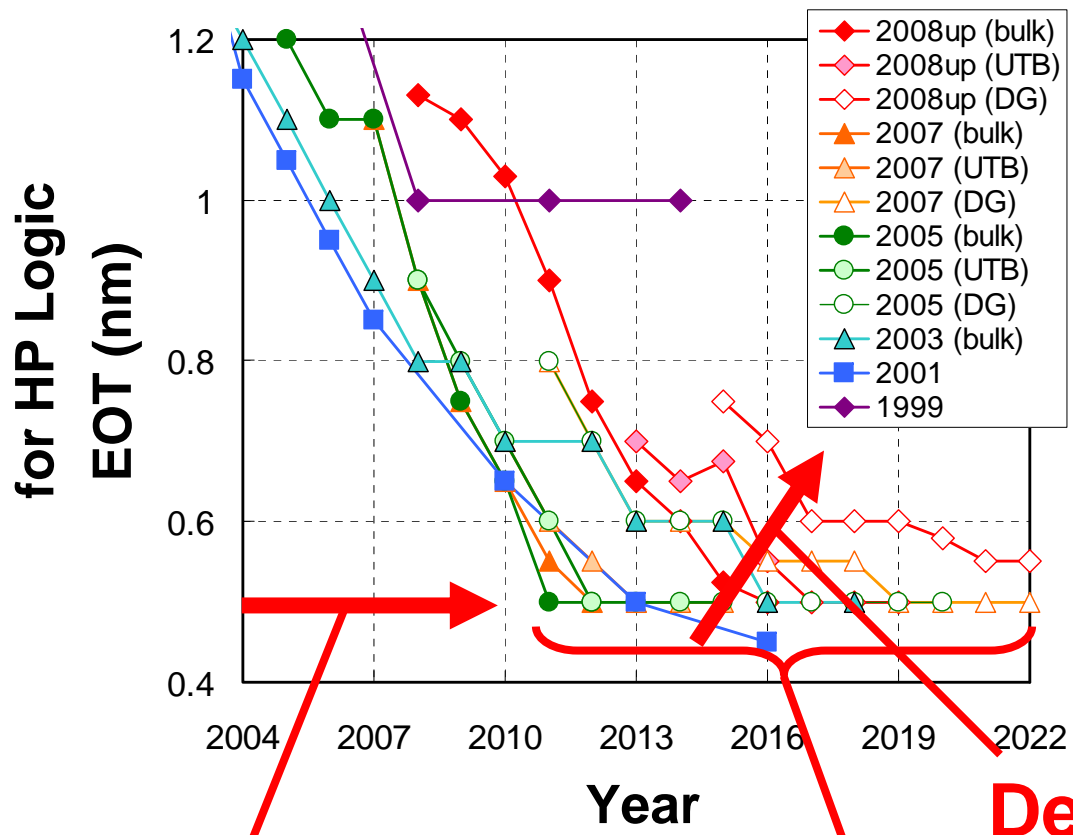
$EOT(C_1) + EOT(C_3) > 0.5nm$

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

Graph:
- y-axis: for HP Logic EOT (nm), ranging 0.4 to 1.2
- x-axis: Year, 2004 to 2022

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

Delay

Is 0.5nm real limit?

Saturation

37

EOT<0.5nm with Gain in Drive Current is Possible

La$_2$O$_3$ gate insulator

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

3.5

W/L=2.5/50μm
PMA 300ºC (30min)

insufficient compensation region [*]

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

4%up

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

14%up

Drain current (mA)

Drain voltage (V)
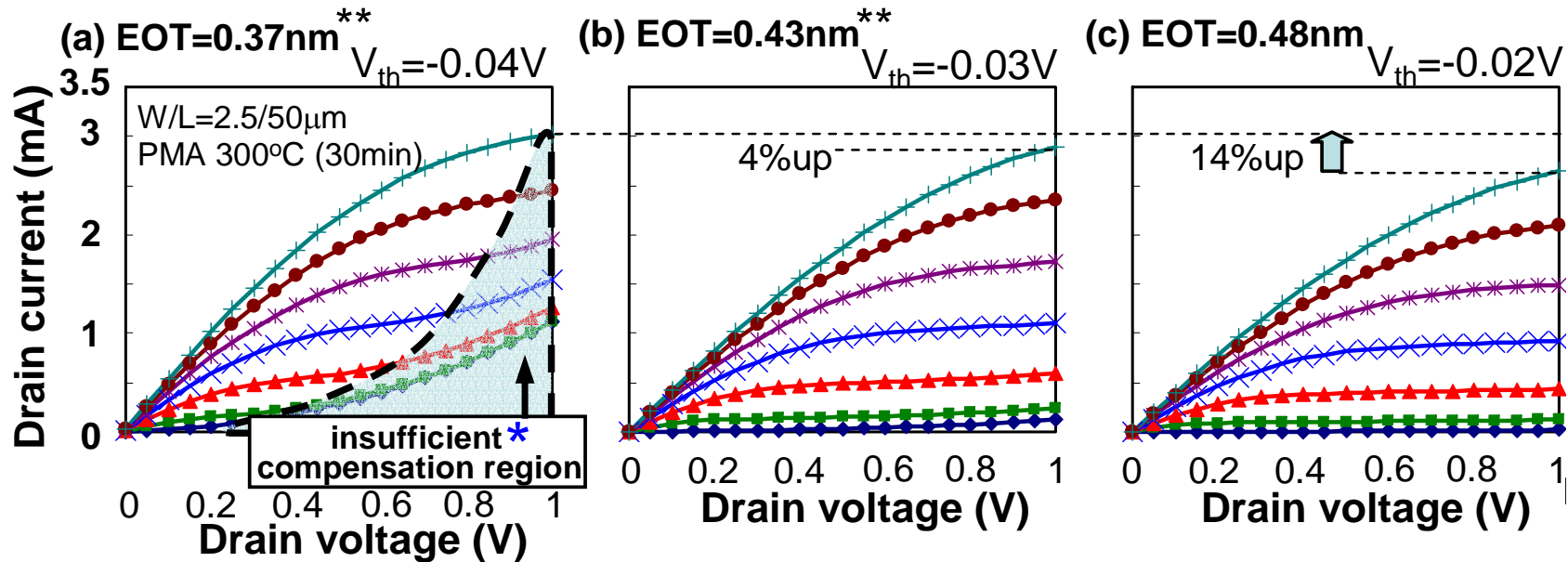
EOT scaling below 0.5nm

⬇

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.

Random Variability Reduction Scenario in ITRS 2007
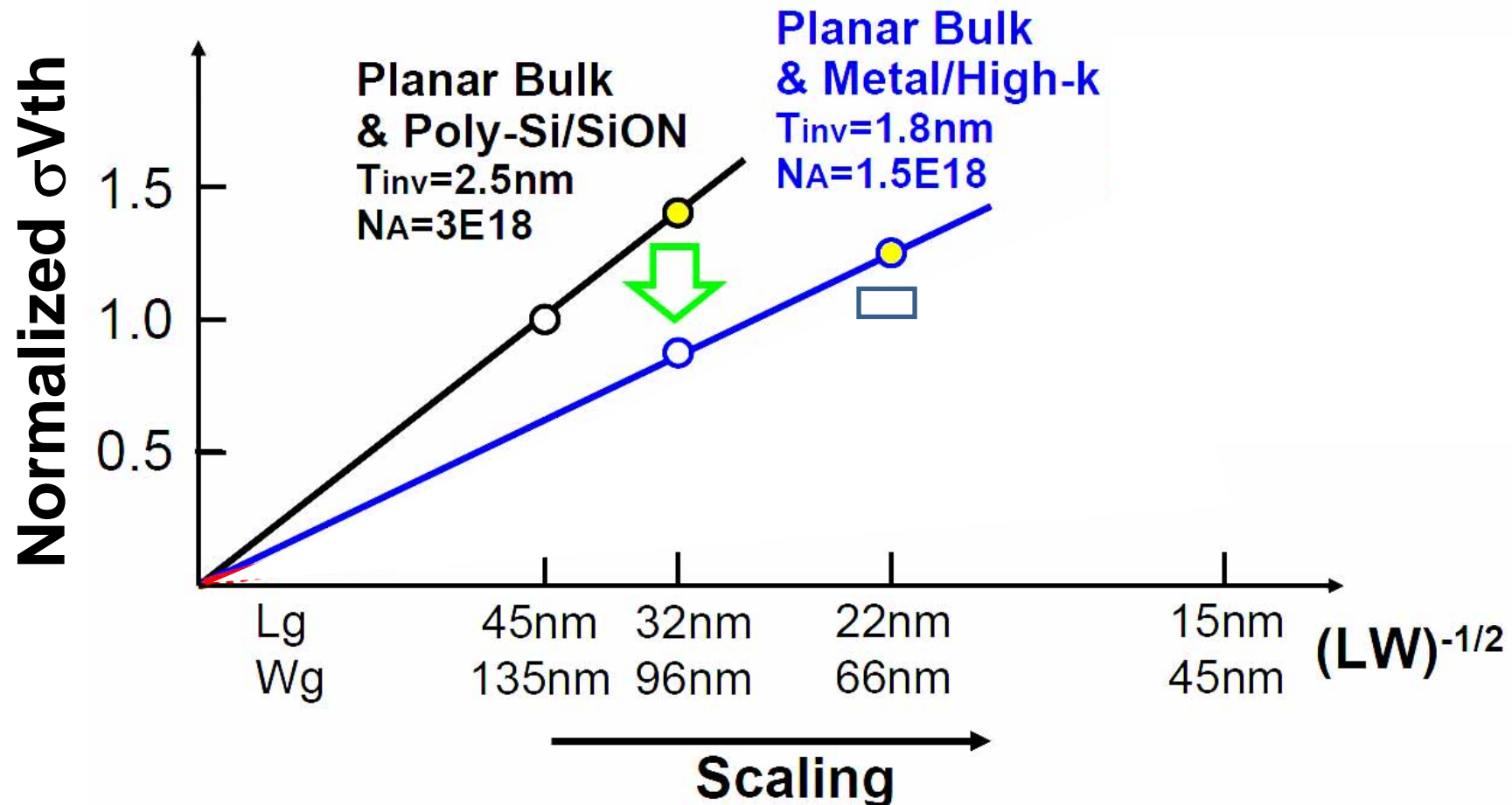
Assumption: Random dopant fluctuation is Main source of Random Variability: Line width roughness of Lg and Wg is not considered in this

Source: 2007 ITRS Winter Public Conf.

4. SRAM cell scaling

Intel® Xeon® 7400 Series (Dunnington)

45 nm high-k 6 cores

16MB shared L3 cache
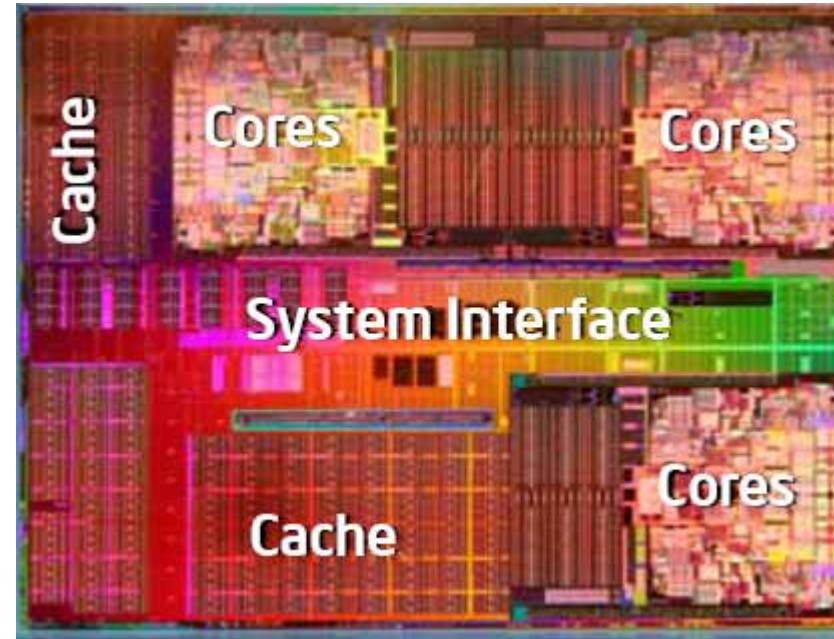
Cache occupies huge area

→ Cell size of SRAM should be minimized

→ However, 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

Intel's **SRAM** test chip trend

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

0.5 X every 2 years

Cell area (μm^2) vs Year

180nm, 130nm, 90nm, 65nm, 45nm, 32nm
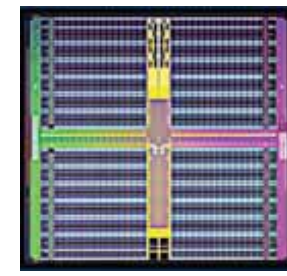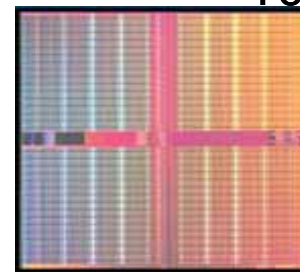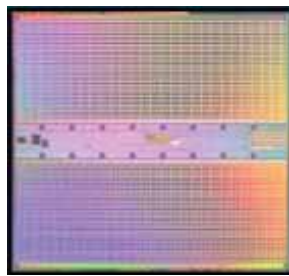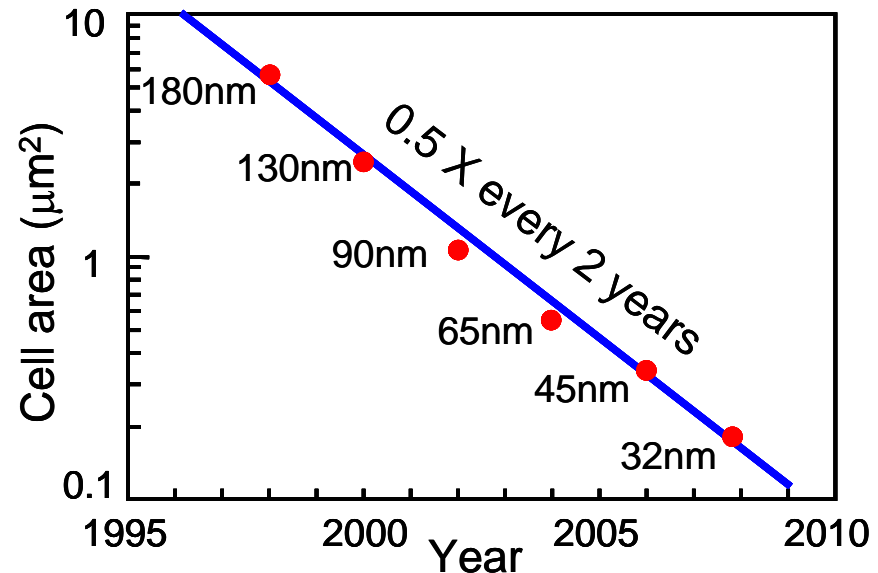
Technology	90 nm Process	65 nm Process	45 nm Process	32 nm Process
Cell size	1.0 μm^2 cell	0.57 μm^2 cell	0.346 μm^2 cell	0.182 μm^2 cell
Capacity	50 Mbit	70 Mbit	153 Mbit	291 Mbit
Chip area	109 mm^2	110 mm^2	119 mm^2	118 mm^2
Functional Si	February '02	April '04	January '06	September '07

43

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

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

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**

0.554 μm

0.180μm

Thin Spacer for S/D

25nm Lg

90nm pitch

44
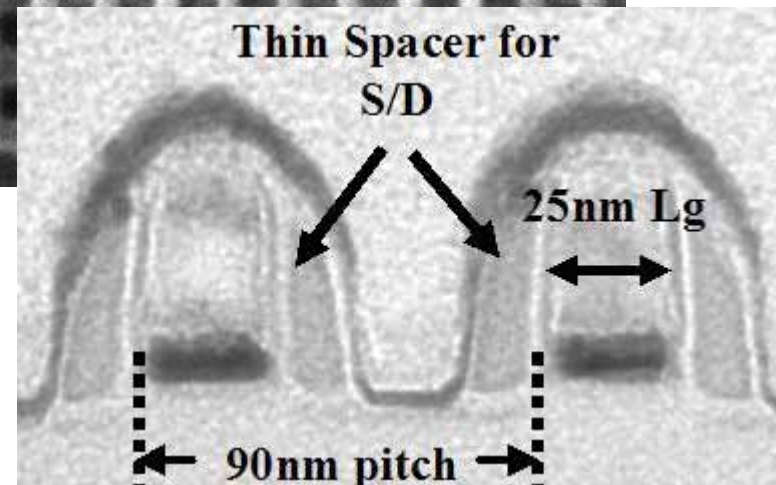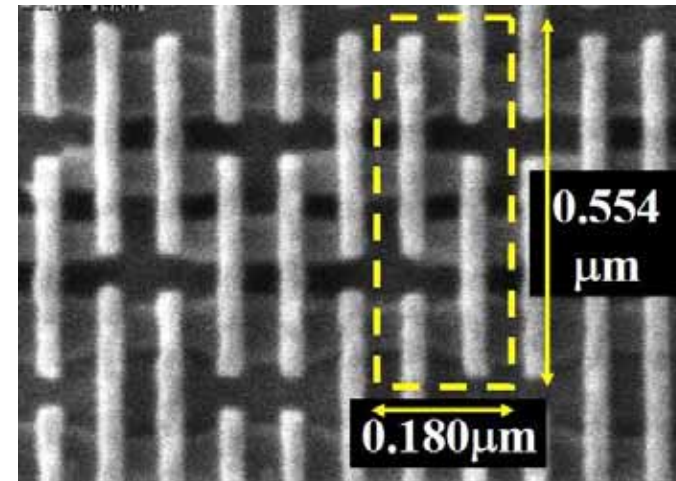
Cell size reduction trends

1/2 or 2/3 per cycle?

Cell area (μm^2)

1

0.57μm²

0.5

Intel

0.35μm²

0.24μm²

0.18μm²

0.2

1/2 per cycle

2/3 per cycle

TSMC

IBM Alliance

0.15μm²

0.1

0.1μm²

65nm   45nm   32nm   22nm
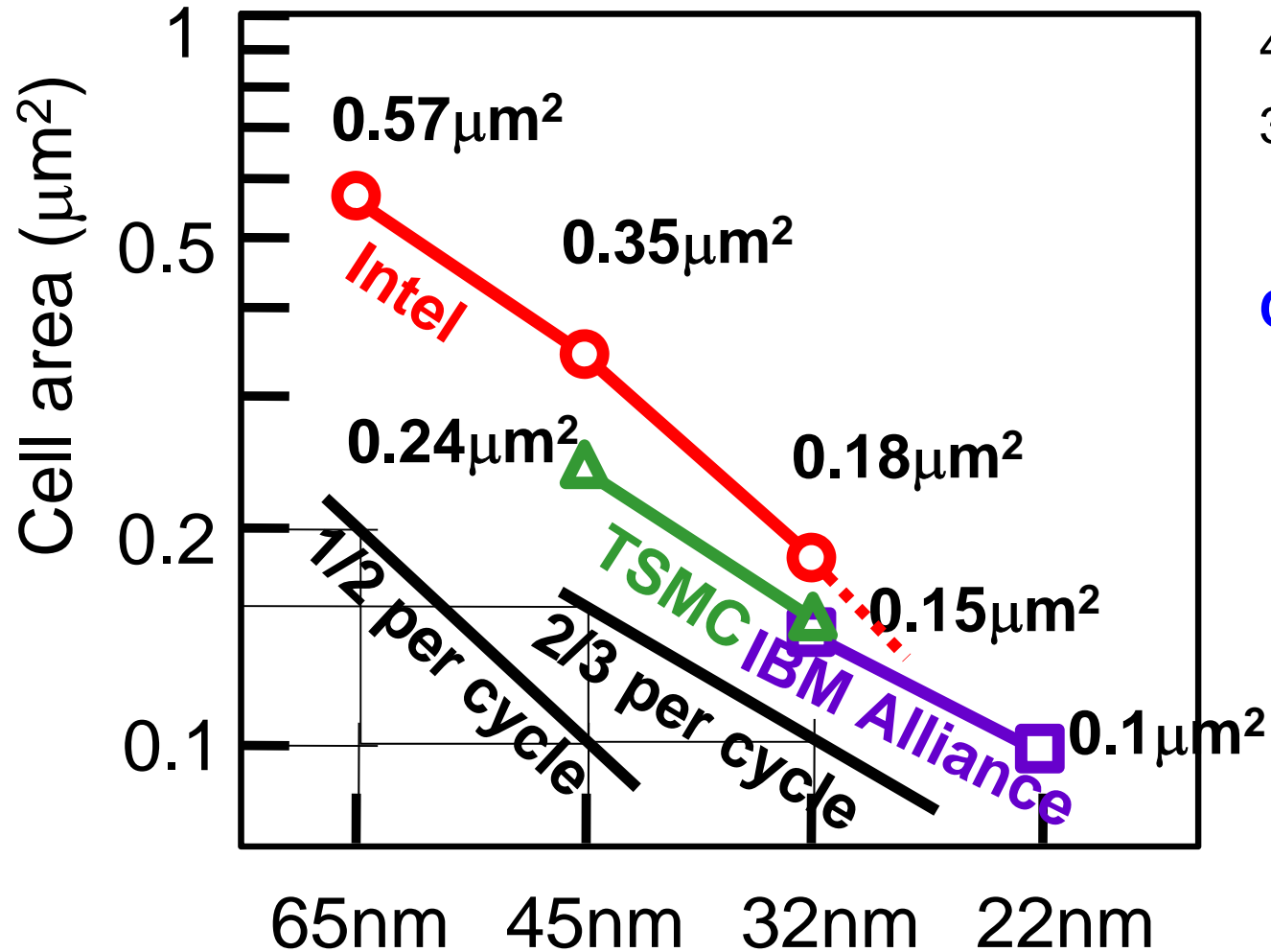
○ Intel

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

$$\sigma V_{Tran} =$$

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

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

Minimal oxide scale

HiK+MG

NMOS C_2

Normalized to 180nm

180nm   130nm   90nm   65nm   45nm

σV_{Tran}

NMOS σVTN (one device)

Minimum device

Nominal device

σV$_T$ (mV)

130nm   90nm   65nm   45nm
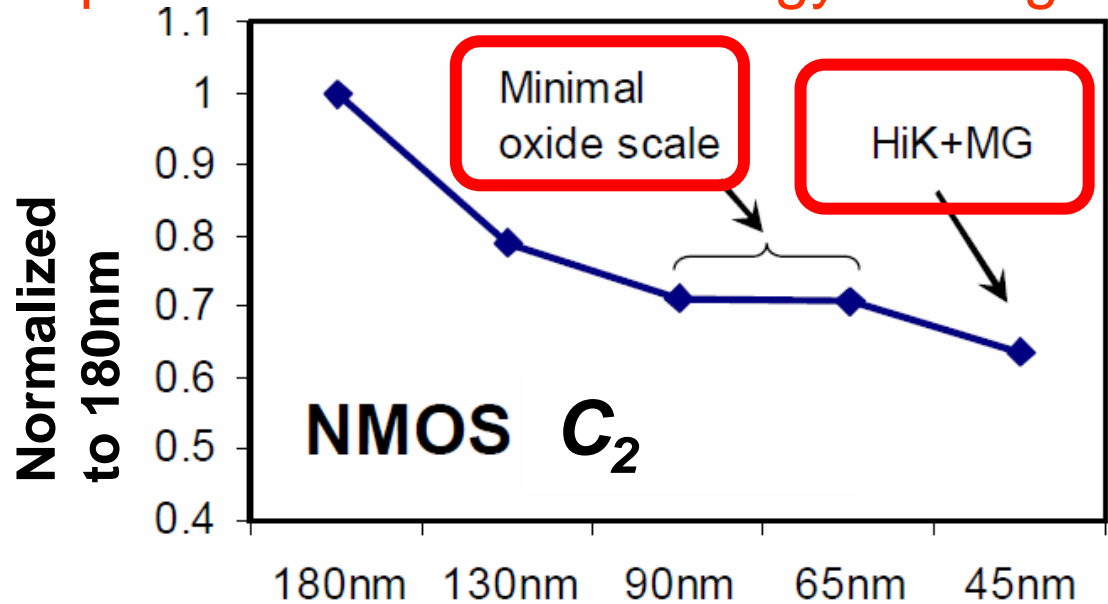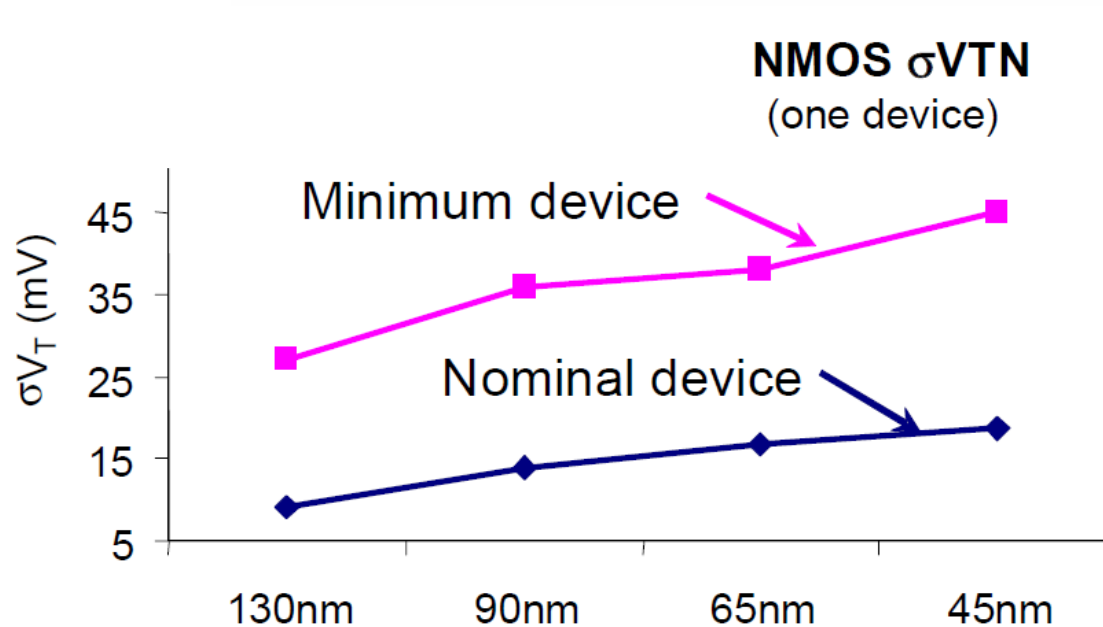
Source: K.J.Kuhn
IEDM 2007

46

Mismatch improvement by layout (Intel)

"tall" design

90nm :1.0 µm²
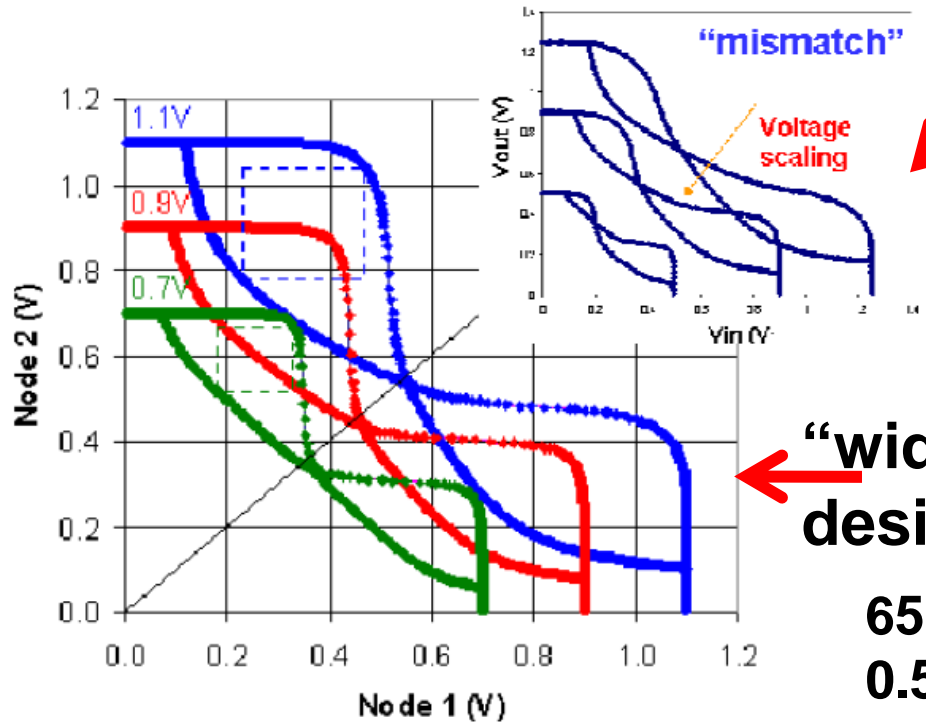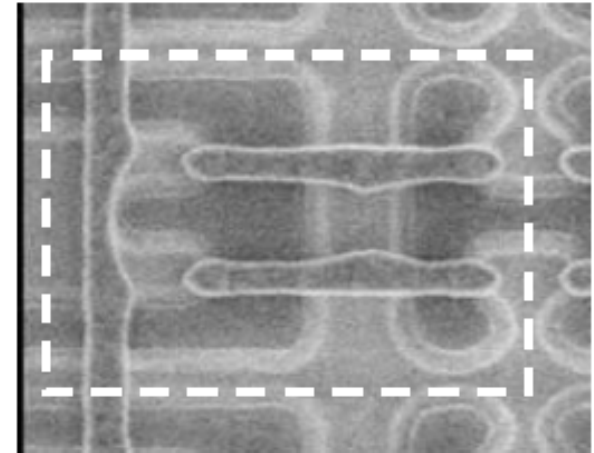
"wide" design

65nm : 0.57 µm²
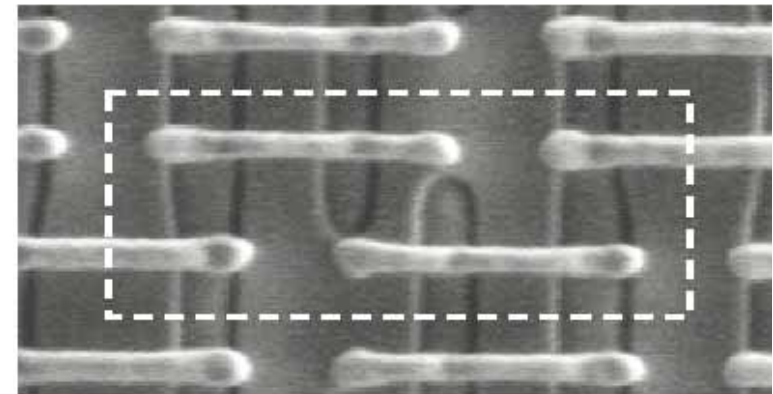
"wide" design (Square endcaps)

45nm 0.346 µm²
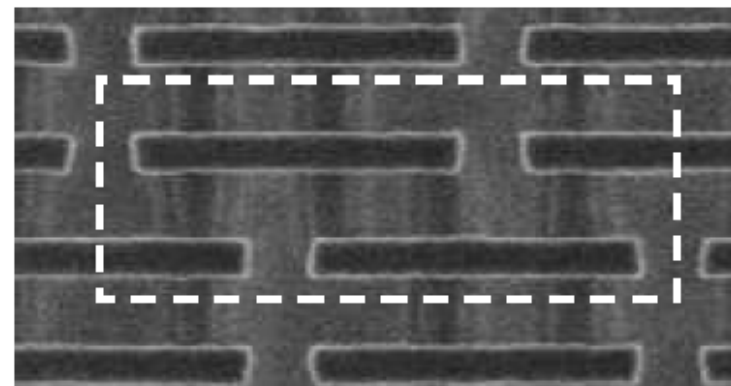
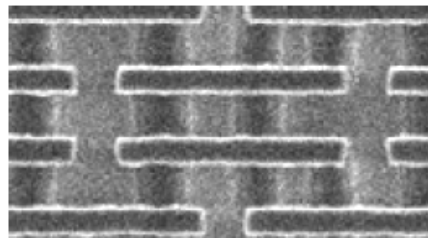**Double patterning for square endcap**

a) Pattern gate lines/spaces
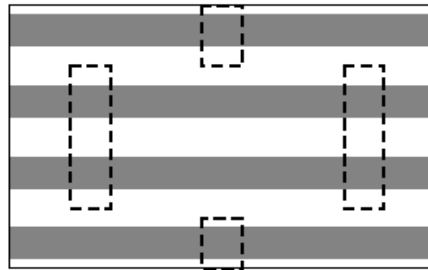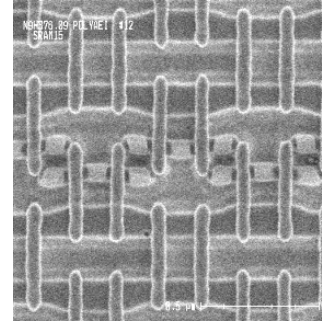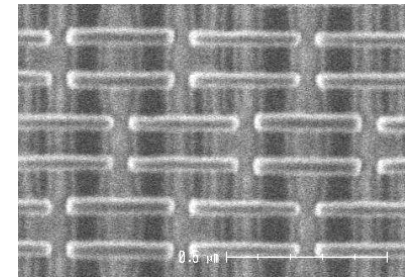
b) Pattern cut mask
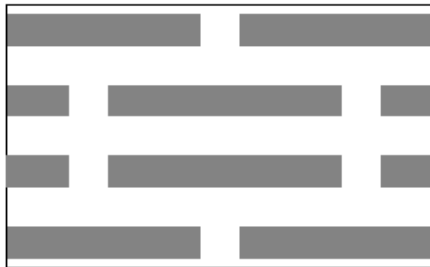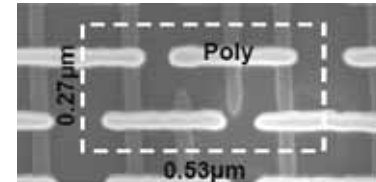
c) Final gate pattern

d) Intel 45nm SRAM cell

Source: M. Bohr, ICSICT2008

Cell evolution is similar

TSMC 45nm
IEDM 2007

TSMC 32nm
IEDM 2007

Poly

0.27 μm

0.53μm
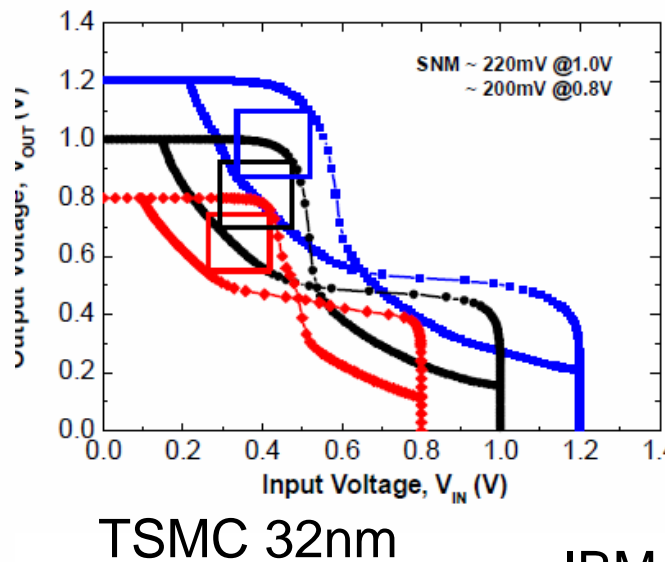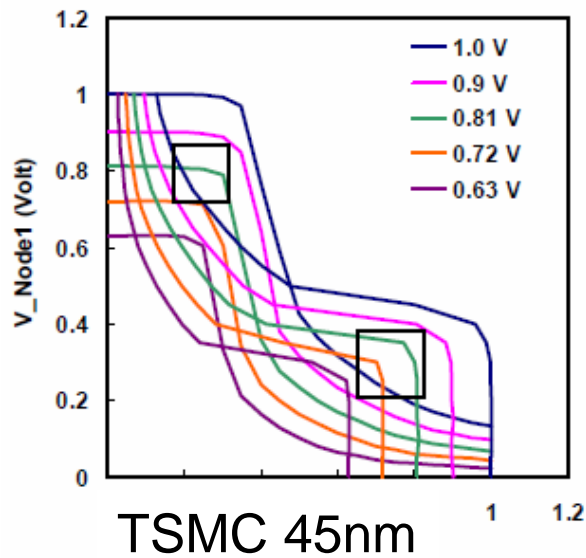
IBM Alliance 32nm
IEDM 2004

0.554 μm

0.180μm
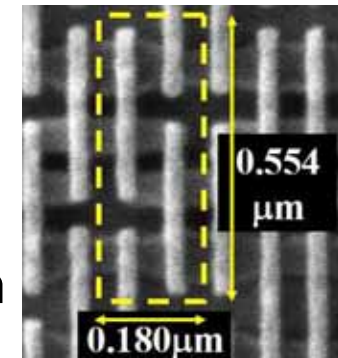
IBM Alliance 22nm
IEDM 2008

TSMC 45nm

SNM ~ 220mV @1.0V
~ 200mV @0.8V

TSMC 32nm

V_{DD} =0.8, 1.0, 1.2V

154 mV

148 mV

SNM [V]

V_{DD} [V]

V_{OUT}

V_{IN}

IBM Gr. 32nm

48
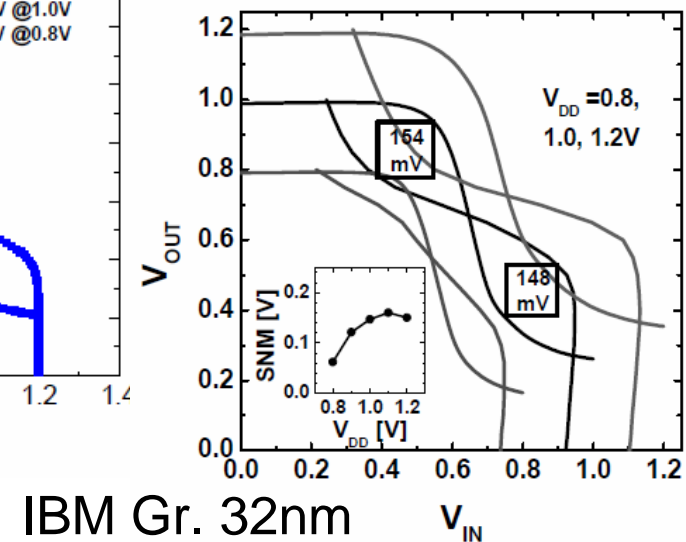
Most Difficult part of SRAM down-scaling is
Vdd down-scaling

Density of on-chip cache SRAM memory is high
and thus, V_{th} 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
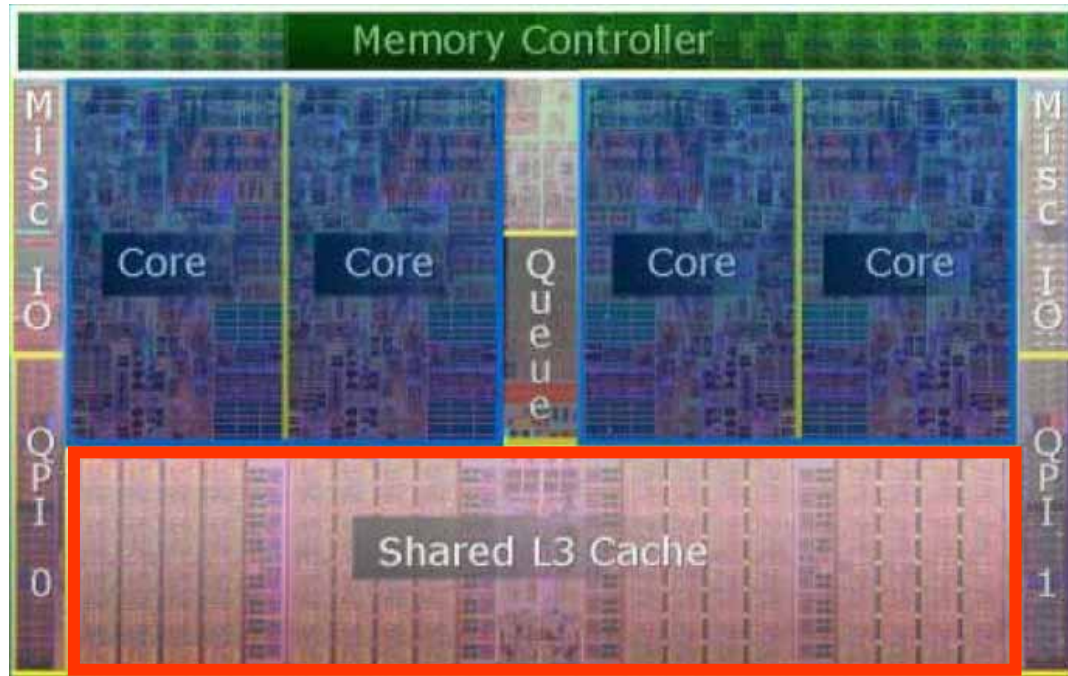
Nehalem(Intel) 2,4 or 8 Cores

Voltage/Frequency Partitioning
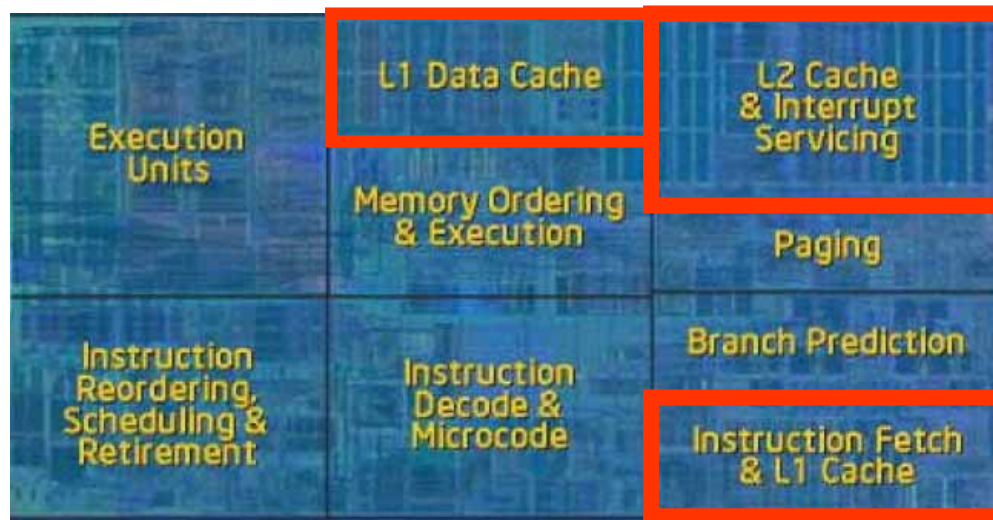
- DDR Vcc
- Core Vcc
- Uncore Vcc

Dynamic Power Management

Chip

8T SRAMCell

32kB L1 I -cache

32kB L1 D-cache
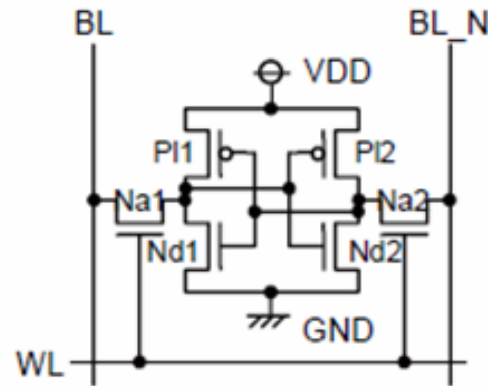
256kB L2 -cache
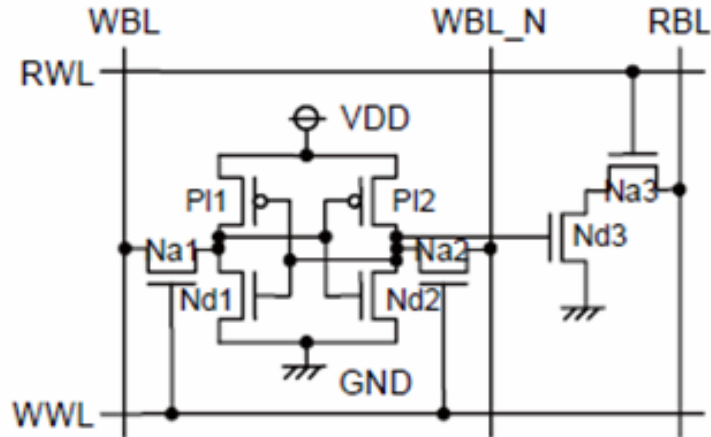
6T SRAMCell

8 MB L3 cache

Core

6T and 8T Cell

6T Cell
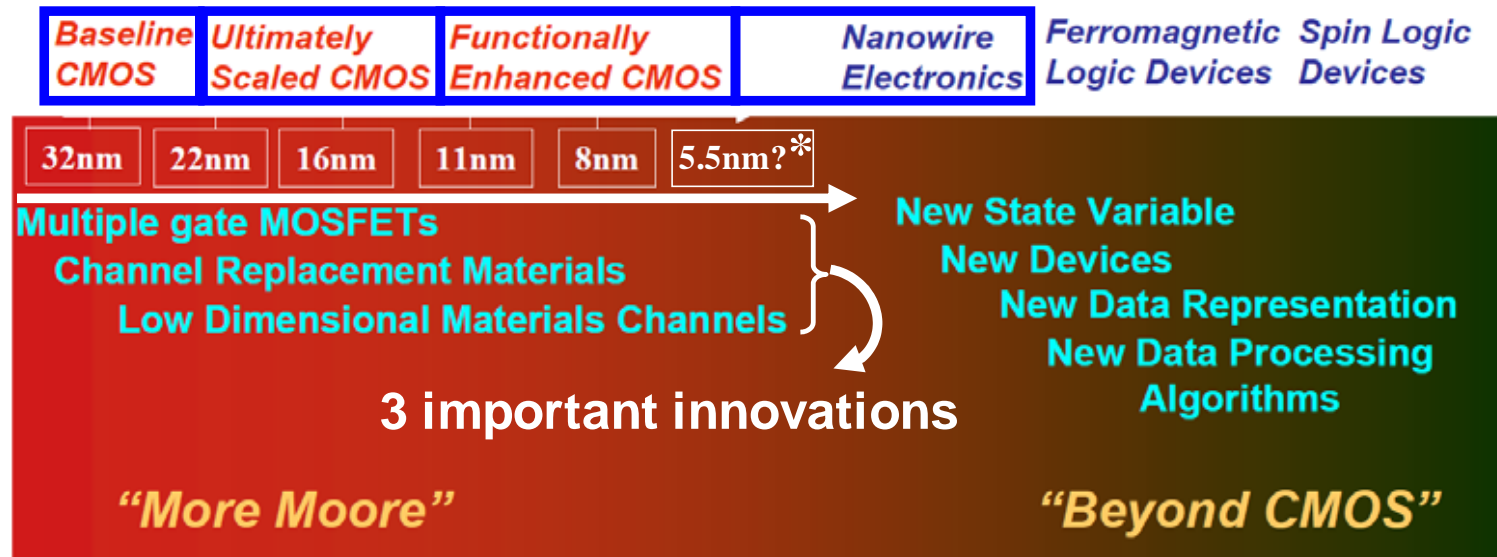

Cell size is small

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.

| Baseline CMOS | Ultimately Scaled CMOS | Functionally Enhanced CMOS | Nanowire Electronics | Ferromagnetic Logic Devices | Spin Logic Devices |

| 32nm | 22nm | 16nm | 11nm | 8nm | 5.5nm?* |

Multiple gate MOSFETs
Channel Replacement Materials
Low Dimensional Materials Channels

New State Variable
New Devices
New Data Representation
New Data Processing Algorithms

3 important innovations

"More Moore"   "Beyond CMOS"

ITRS figure edited by Iwai

Source: 2008 ITRS Summer Public Conf.

*5.5nm? was added by Iwai

53

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

Drain   Source

Trade off

Carrier scattering probability

Small   **Large**

# of quantum channel

Small   **Large**

Increase in Ion (Id-sat)

1 ⟷ **2** E-k band

E

QC
QC
QC
QC

3 dense nanowires
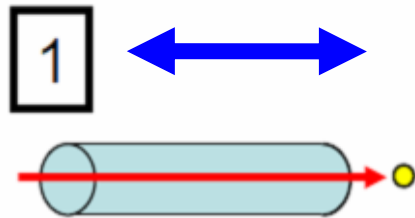
High Conduction (1D)

Go=77.8μS/wire

3D stacking
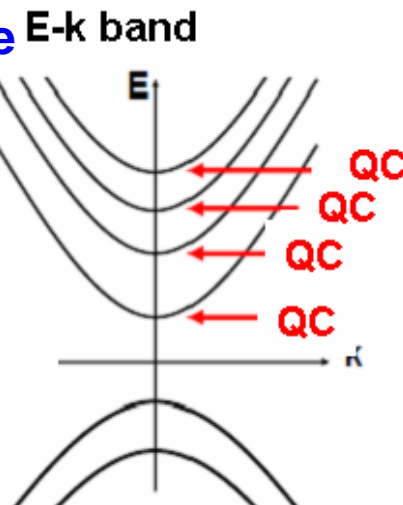
Multiple quantum channel (QC) used for conduction
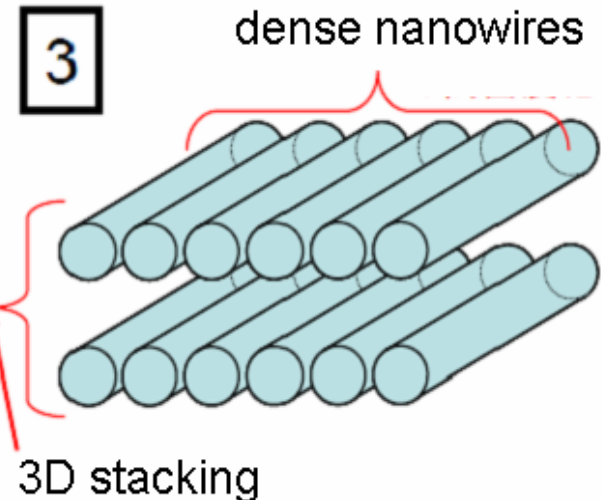
High-density lateral and vertical integration

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

Si nanowire FET as a strong candidate after CMOS limitation

1. Compatibility with current CMOS process

2. Good controllability of I_{OFF}

3. High drive current

cut-off

source   drain

Gate: OFF

1D ballistic conduction

Multi quantum Channel

High integration of wires

E

Quantum channel
Quantum channel
Quantum channel

Quantum channel

k

TEM image
10nm

gate electrode

wire

sidewall
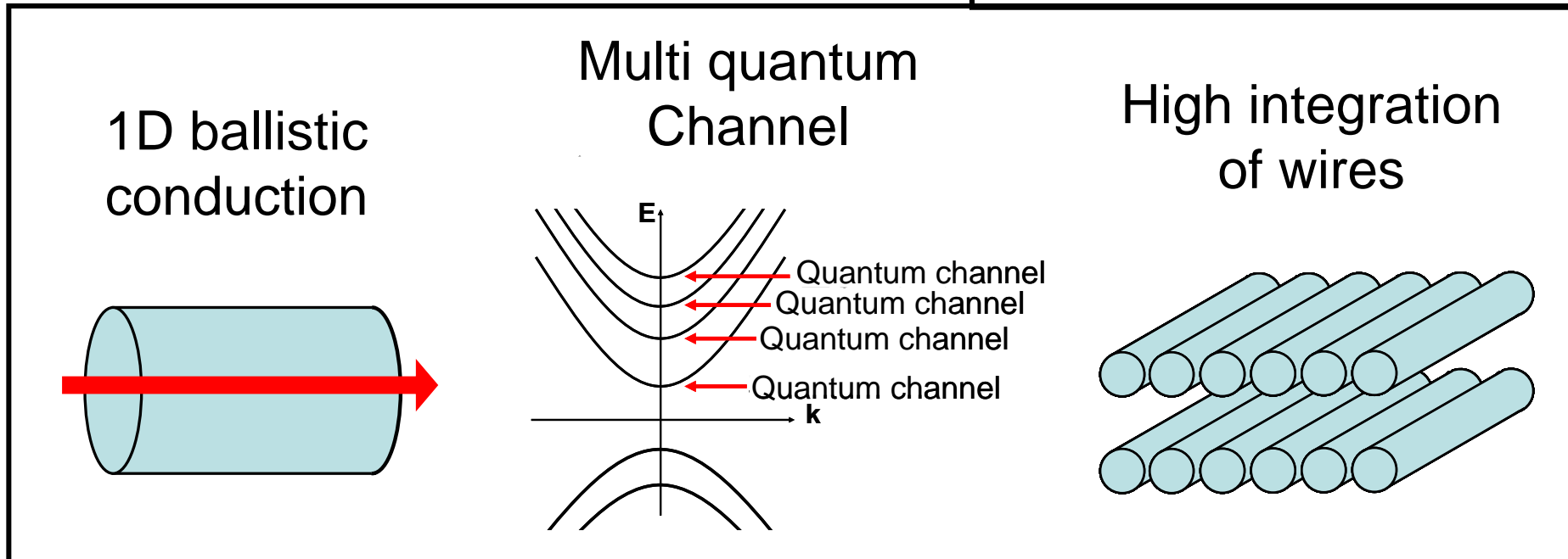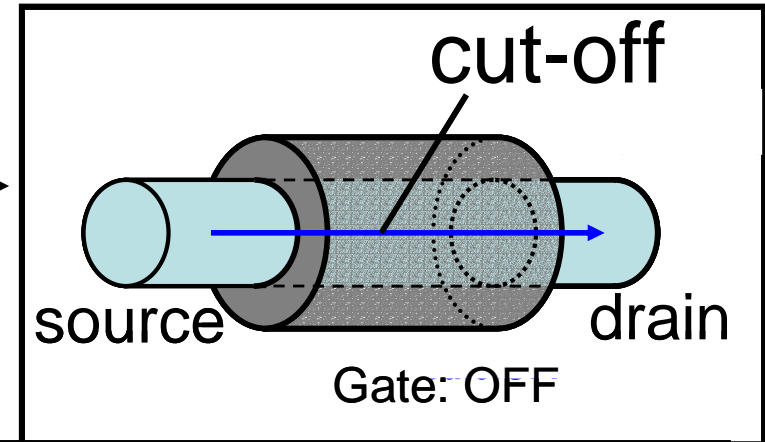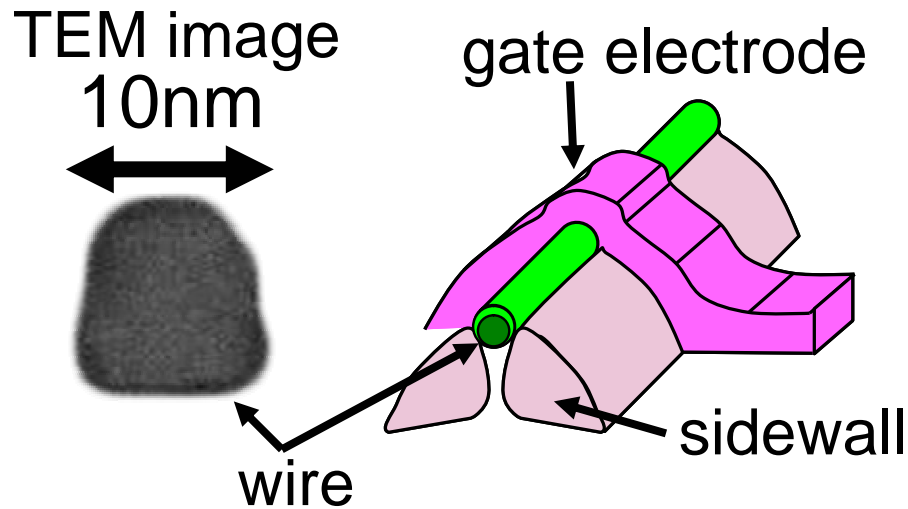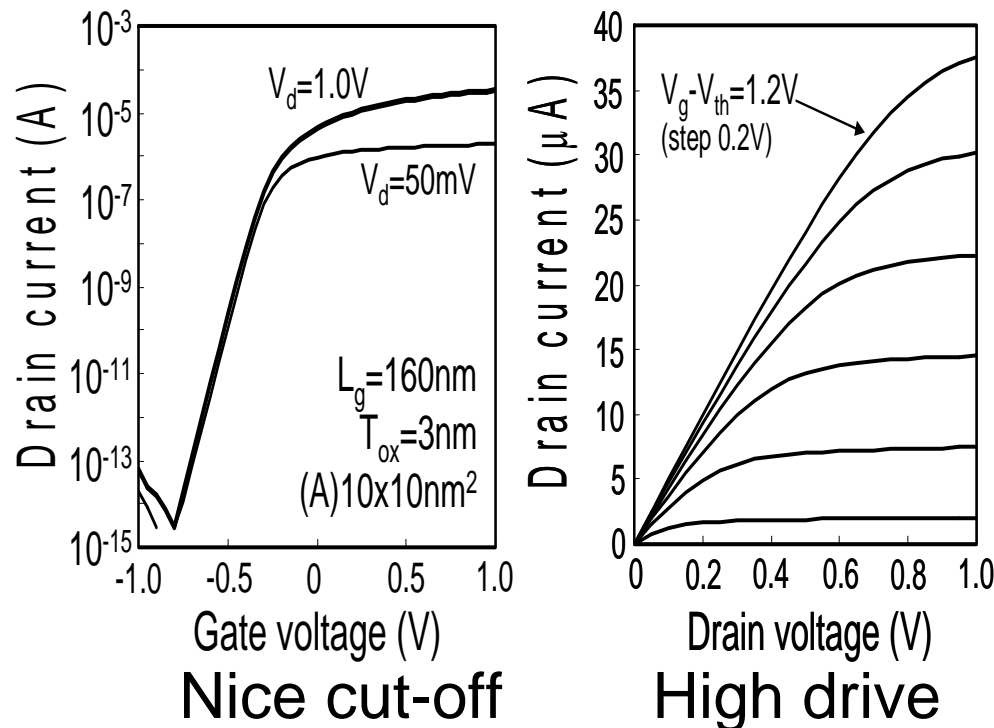
Advantage of Si nanowire

Large drive current

Spec. in 2019 by ITRS
$I_{ON}=2.3mA/\mu m$
$L_g=11nm$, $T_{ox}=0.6nm$

Our nanowire FET
$I_{ON}=0.25mA/\mu m$ (with 2010
Litho. tech.)
$L_g=160nm$
$T_{ox}=3.0nm$

With 2019 litho. tech.

$I_{ON}=2.3mA/\mu m$ will be
obtained even with
$L_g=80nm$ and $T_{ox}=1.5nm$

$L_g=160nm$, $T_{ox}=3nm$

Drain current (A)

$V_d=1.0V$

$V_d=50mV$

$L_g=160nm$
$T_{ox}=3nm$
(A)10x10nm^2

Gate voltage (V)

Nice cut-off

Drain current (μA)

$V_g-V_{th}=1.2V$
(step 0.2V)

Drain voltage (V)

High drive

by the courtesy of Professor H.Iwai

56

Our roadmap for R &D

Source: H. Iwai, IWJT 2008

2007 2010 2015 2020 2025 2030 2035

Extended CMOS: More Moore + CMOS logic

- Research
- Developmen
- Production

PJT(2007~2012)

Beyond the horizon

Si Channel

Si Fin, Tri-gate

Natural direction of downsizing

Si Nano wire

Mechanical Stress, Roughness, Surface control

III-V and Ge Nano wire

Nanowire

Selection

Problem: High-k gate oxides, etching of III-V wire

Tube

CNT

Selection

Tube, Ribbon

Selection

Ribbon

Graphene

ITRS

ITRS Beyond CMOS

High conduction By 1D conduction

Cloud

Horizon

More Moore

Extended CMOS

More Moore ??

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

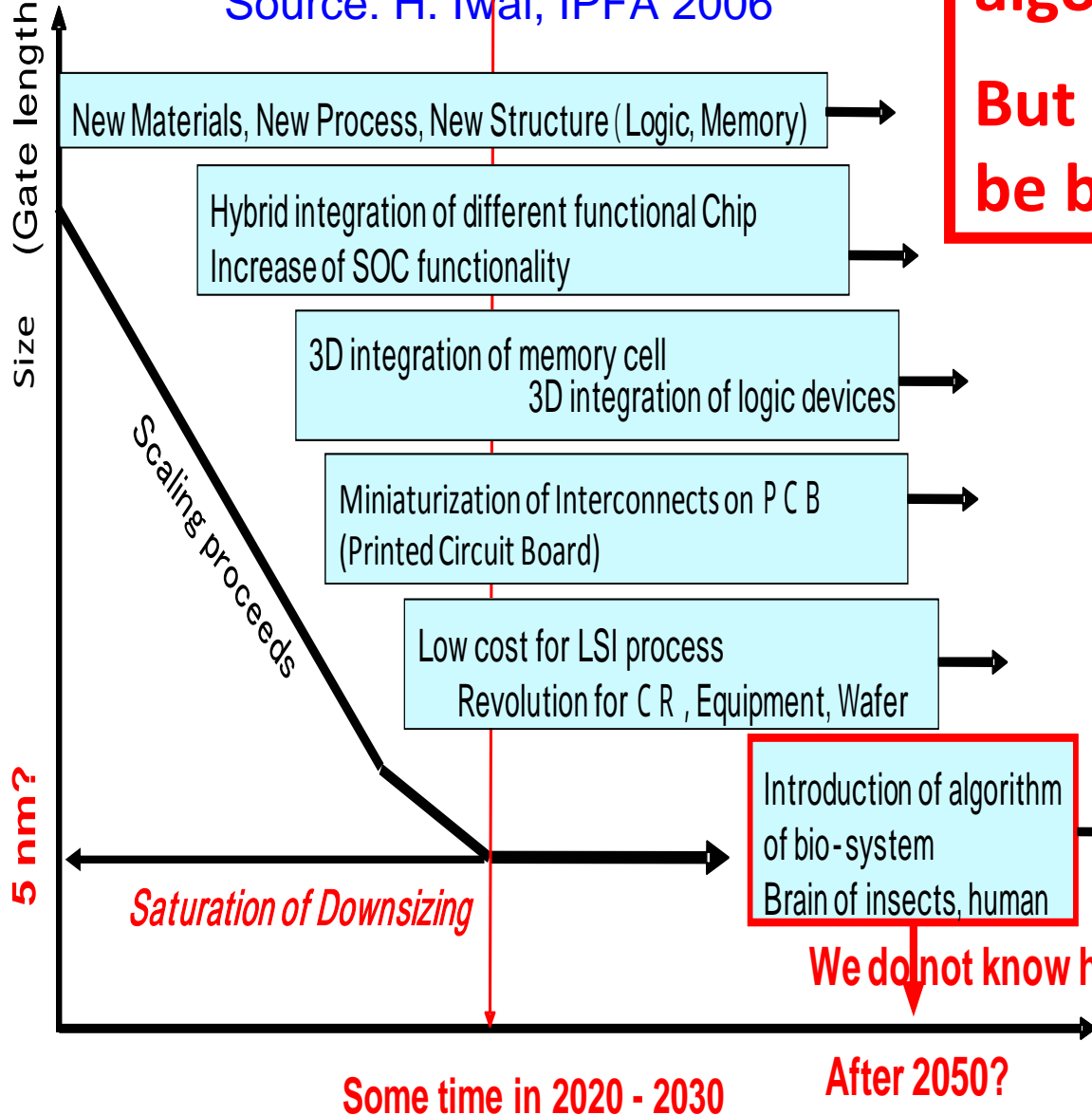Chirality determines conduction types: metal or semiconductor

Graphene:

Graphene formation technique

Suppression of off-current
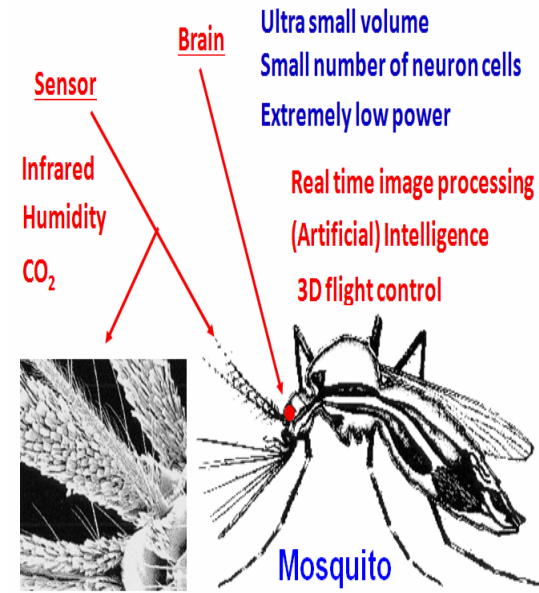
Very small bandgap or no bandgap (semi-metal)

Control of ribbon edge structure which affects bandgap
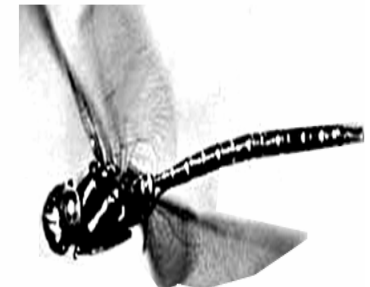
57

Long term roadmap for development

Source: H. Iwai, IPFA 2006

(Gate length etc)

Size

Scaling proceeds

5 nm?

Saturation of Downsizing

NewMaterials, NewProcess, NewStructure 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
(Printed Circuit Board)

Lowcost for LSI process
Revolution for Equipment, Wafer

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

Some time in 2020 - 2030

We do not know how?

After 2050?

We do know system and algorithms are important!

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

Sensor

Brain

Ultra small volume
Small number of neuron cells
Extremely low power

Infrared
Humidity
CO_2

Real time image processing
(Artificial) Intelligence
3D flight control

Mosquito

Dragonfly brain has even
further higher performance

Thank you
for your attention!