Past and Future of Silicon Integrated Circuit Technology

@School of Physics, Xinjiang University
September 7, 2007

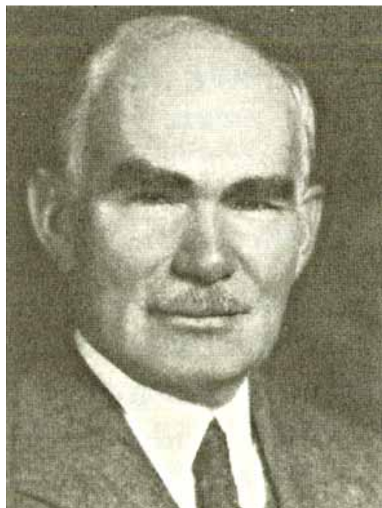
Tokyo Institute of Technology, Japan
Hiroshi IWAI

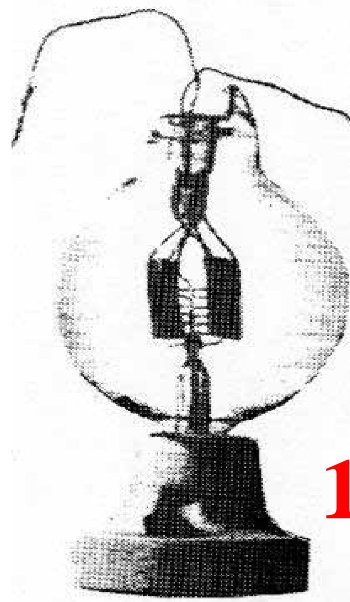
- Electronics is the

  Most important invention in the 20th century
- Electronics: Electronic Circuits or IC
- Electronic Circuits in 100 years

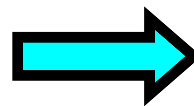
  Vacuum tube → ULSI

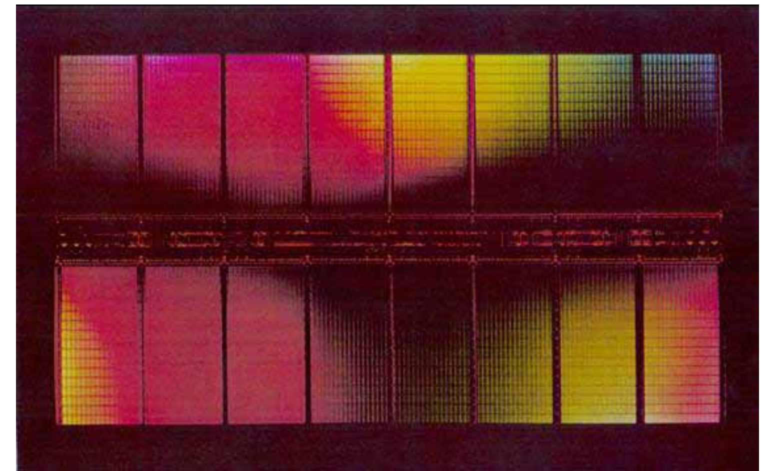
Last year was 100 year anniversary





1906





集積回路

IC Integrated Circuit

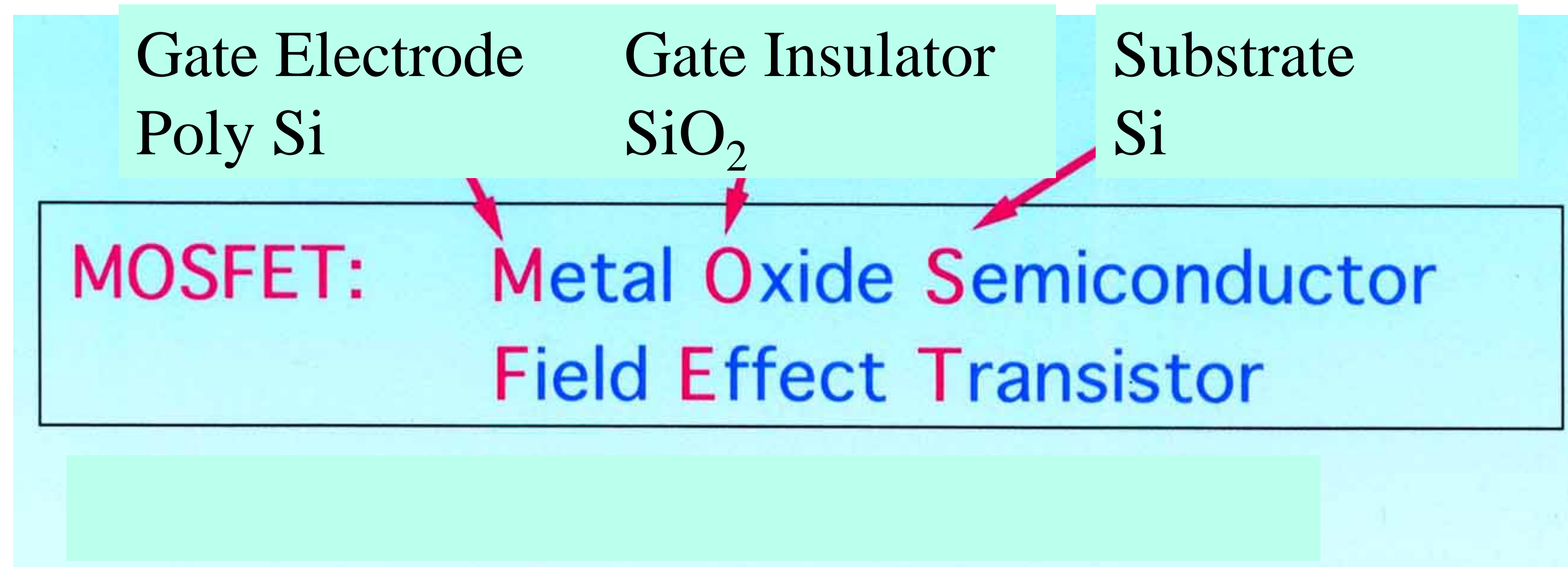
大規模集積回路

LSI Large Scale Integrated Circuit

VLSI Very Large Scale Integrated Circuit

ULSI Ultra Large Scale Integrated Circuit

Year	Device	Number of Transistors
1947年	Transistor	1
1960年代始	IC	Several
1970年代始	LSI	1,000
1980年代始	VLSI	10,000
1990年代始	ULSI	100,000
2000年代始	?LSI	10,000,000



Gate Electrode
Poly Si

Gate Insulator
SiO_2

Source

n-Si

n-Si

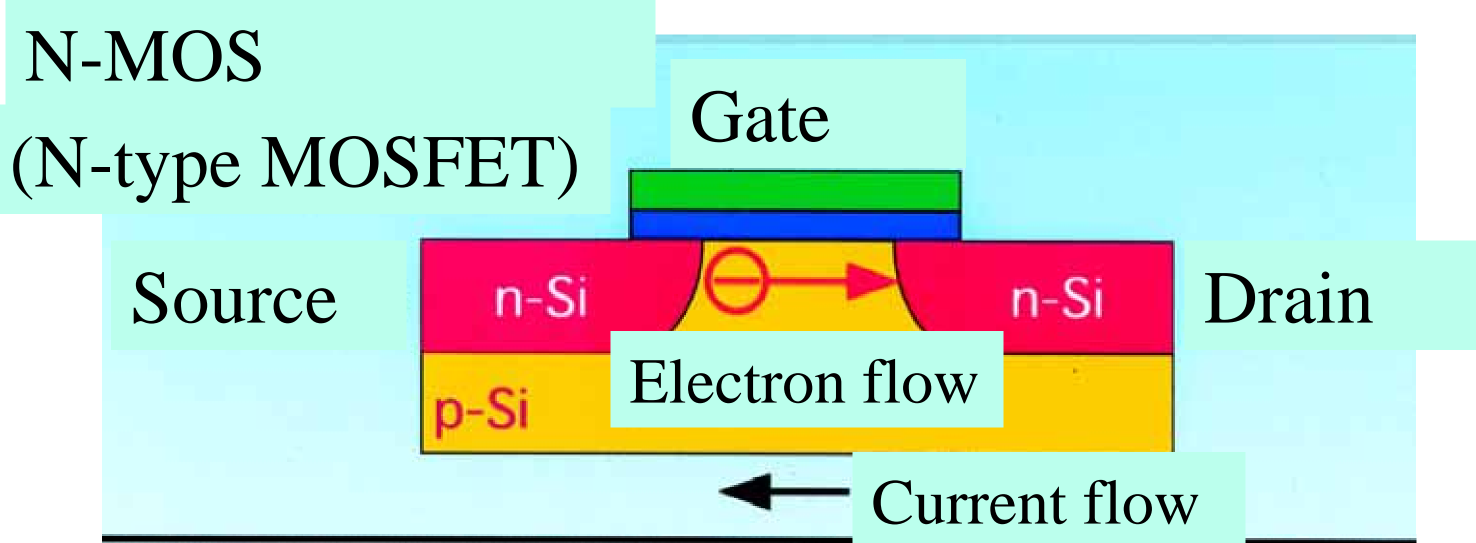
Drain

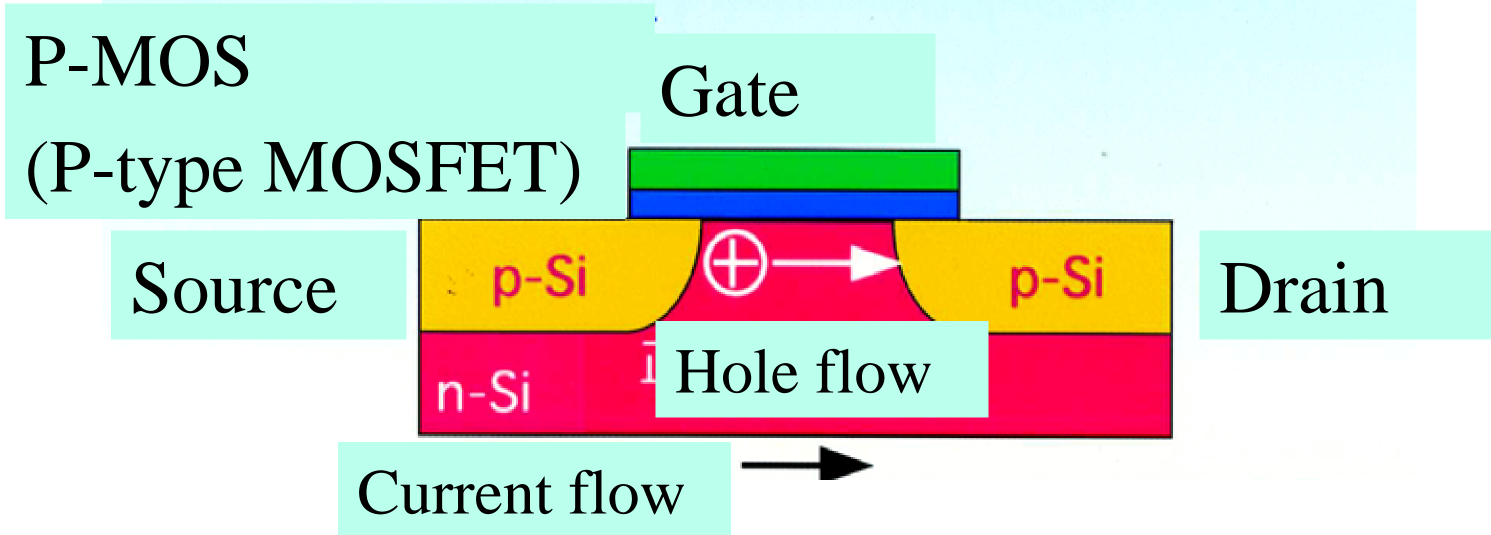
p-Si

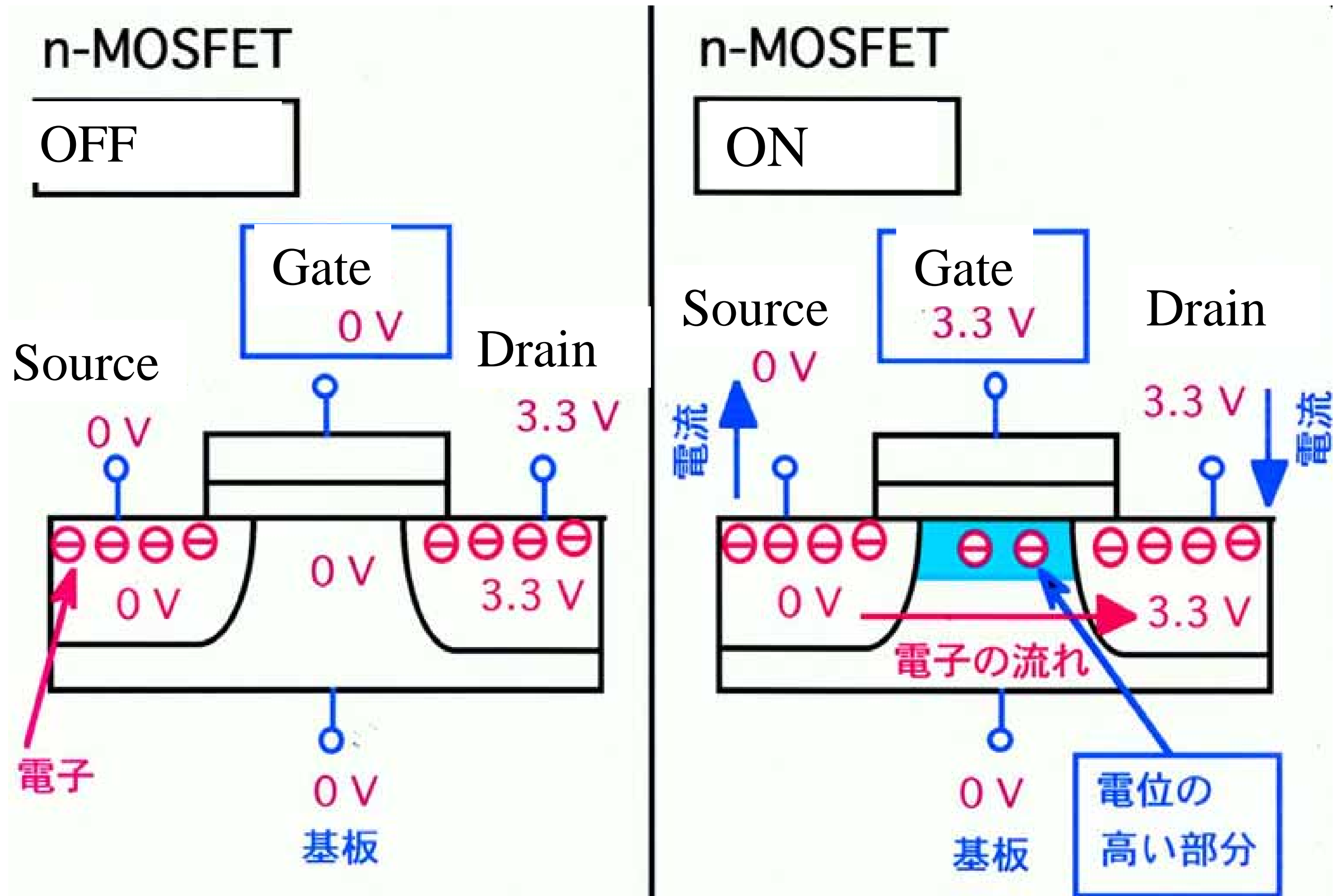
Si
Substrate

Channel

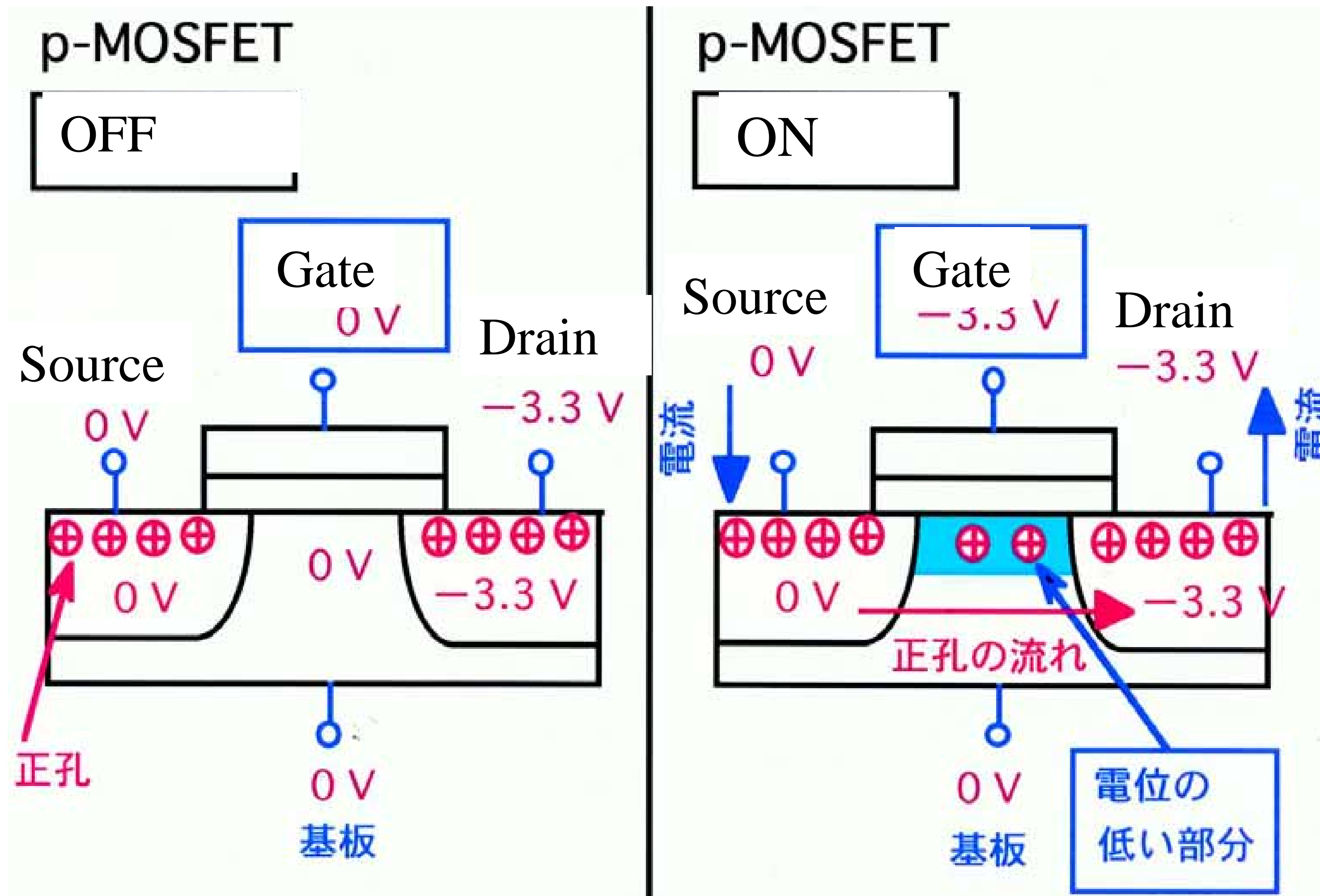
N-MOS (N-type MOSFET)


N-MOS
(N-type MOSFET)
Gate
Source
n-Si
n-Si
Drain
p-Si
Electron flow
Current flow


P-MOS
(P-type MOSFET)
Gate
Source
p-Si
p-Si
Drain
n-Si
Hole flow
Current flow

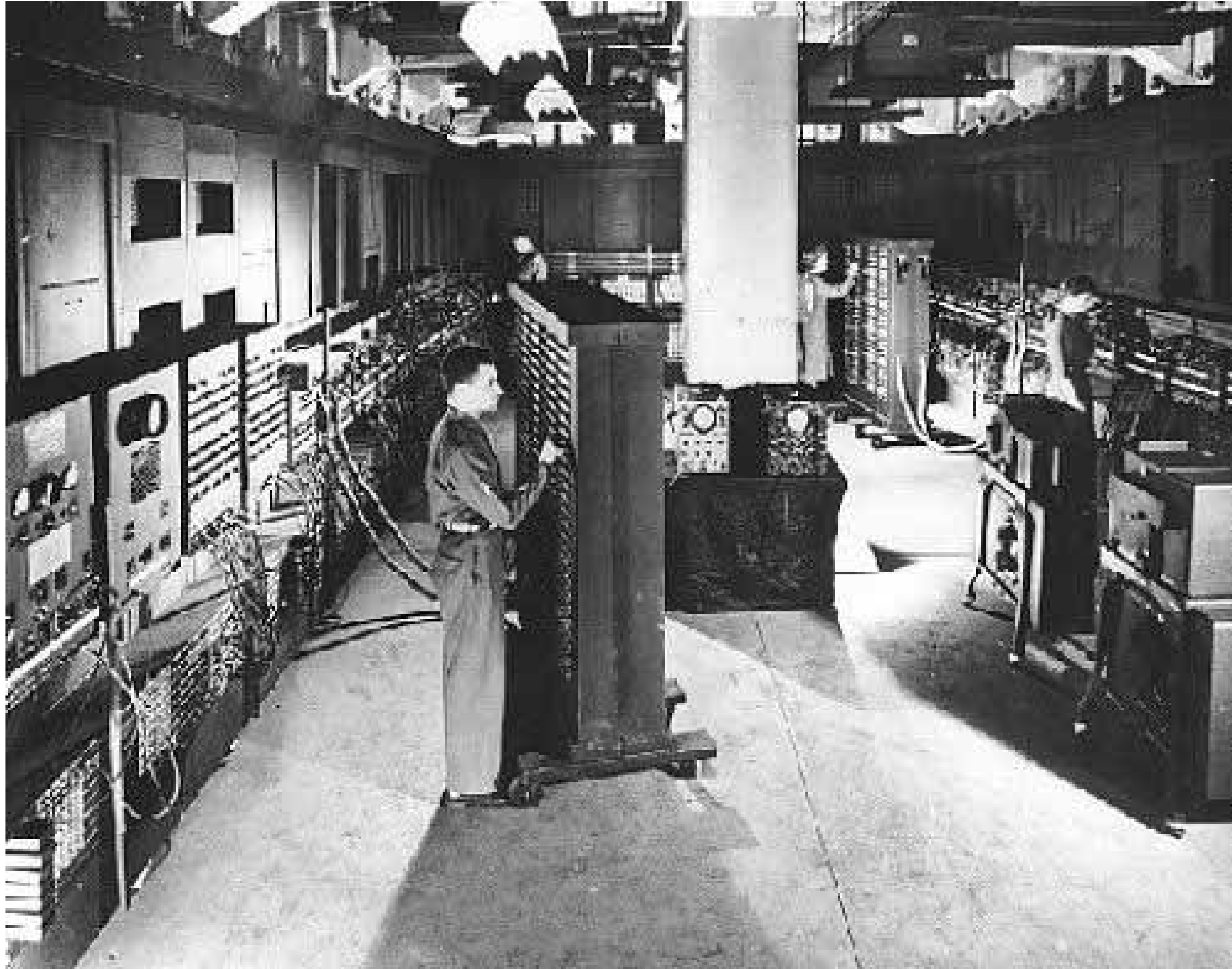


ON

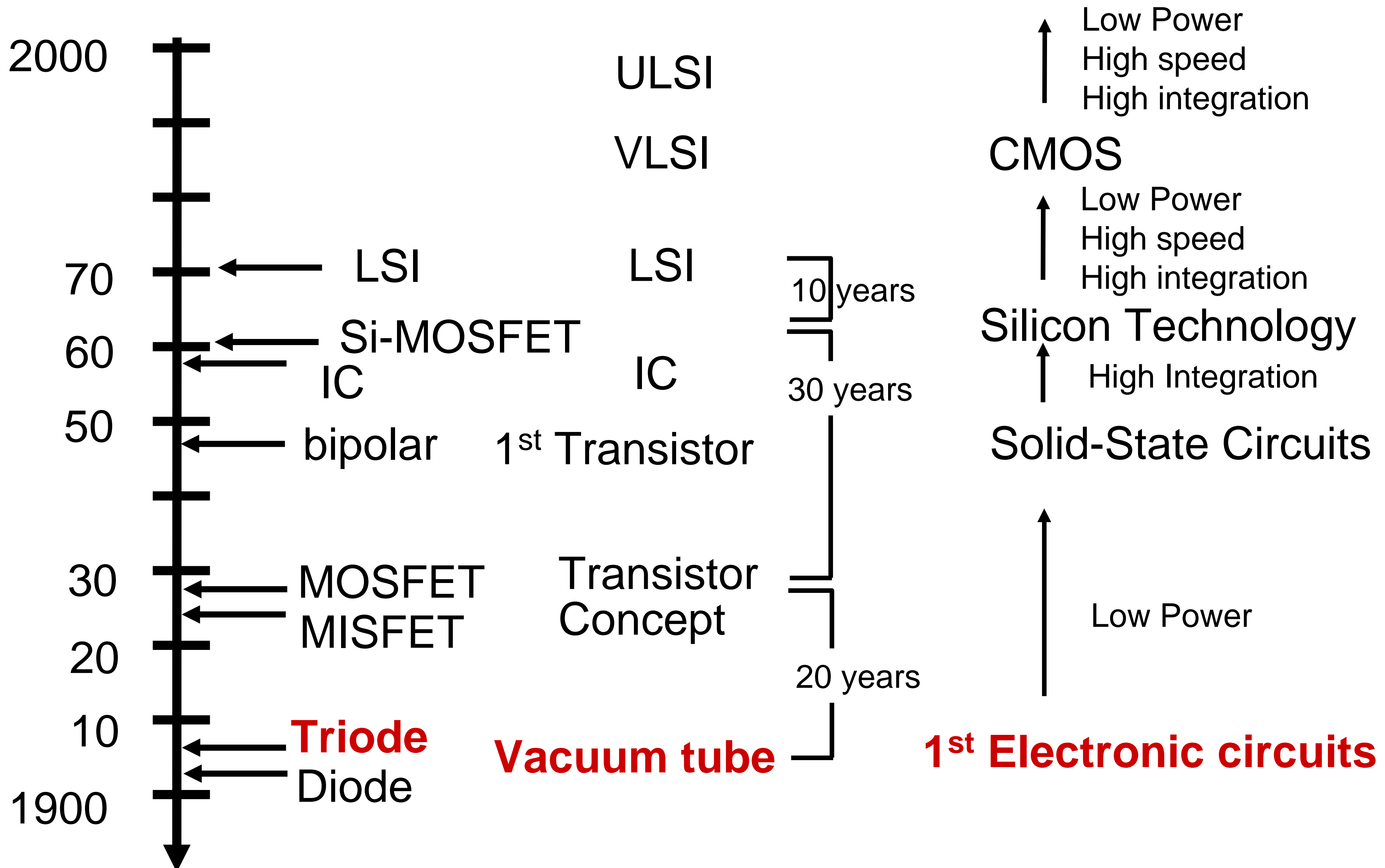

p-MOSFET
OFF
Gate
0 V
Source
0 V
Drain
−3.3 V
0 V
0 V
−3.3 V
正孔
0 V
基板
p-MOSFET
ON
Source
0 V
Gate
−3.3 V
Drain
−3.3 V
電流
電流
0 V
−3.3 V
正孔の流れ
0 V
基板
電位の
低い部分

First Computer Eniac: made of huge number of vacuum tubes 1946

Big size, huge power, short life time filament



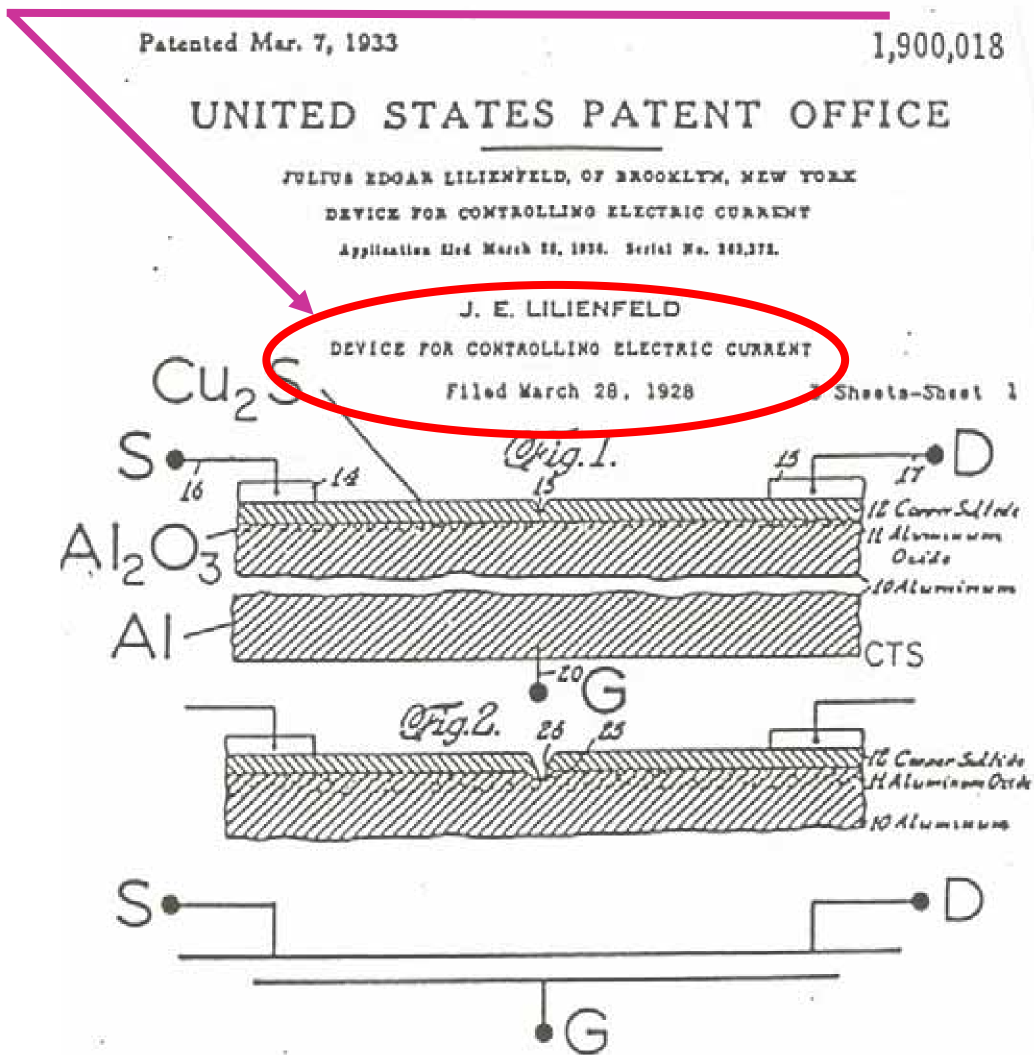
History of Electronic Devices



J. E. LILIENFELD

DEVICES FOR CONTROLLED ELECTRIC CURRENT

Filed March 28, 1928



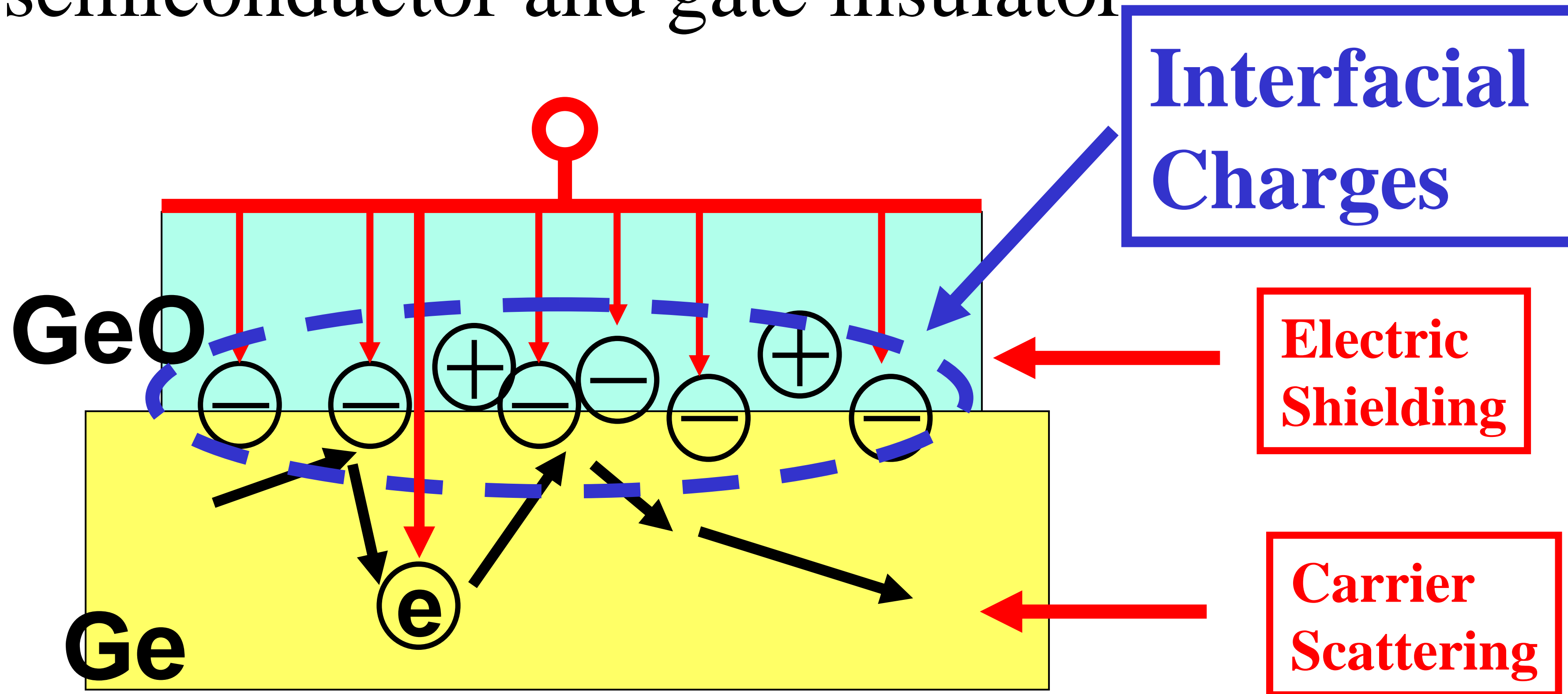
J.E.LILIENFELD

However, no one could realize MOSFET operation for more than 30 years.

Because of very bad interface property between the semiconductor and gate insulator

Even Shockley!

Very bad interface property between the semiconductor and gate insulator



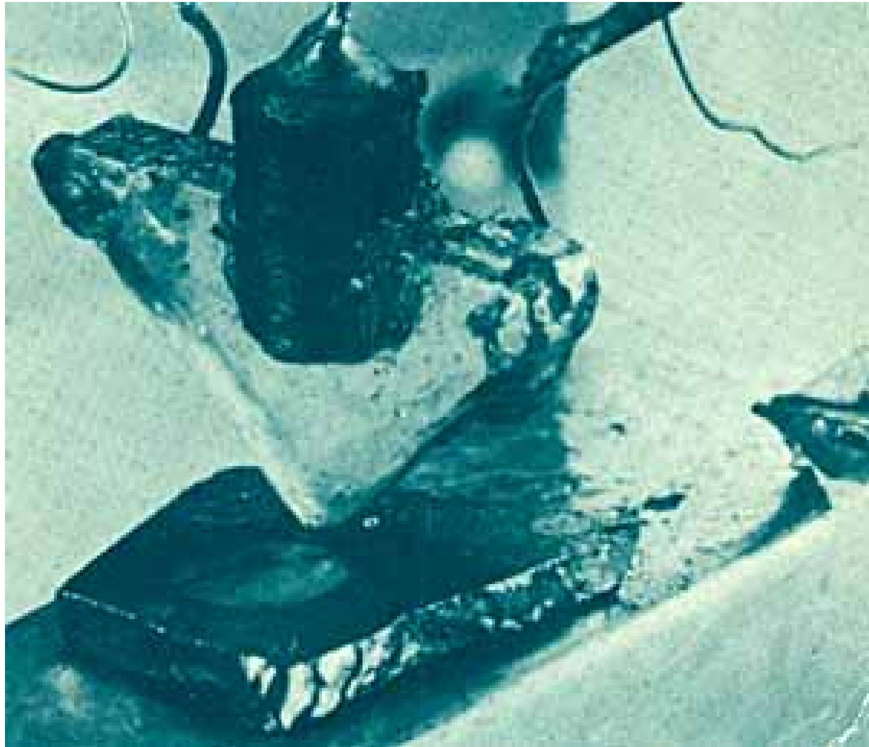
Drain Current was several orders of magnitude smaller than expected

Even Shockley!

1947: 1st Translator

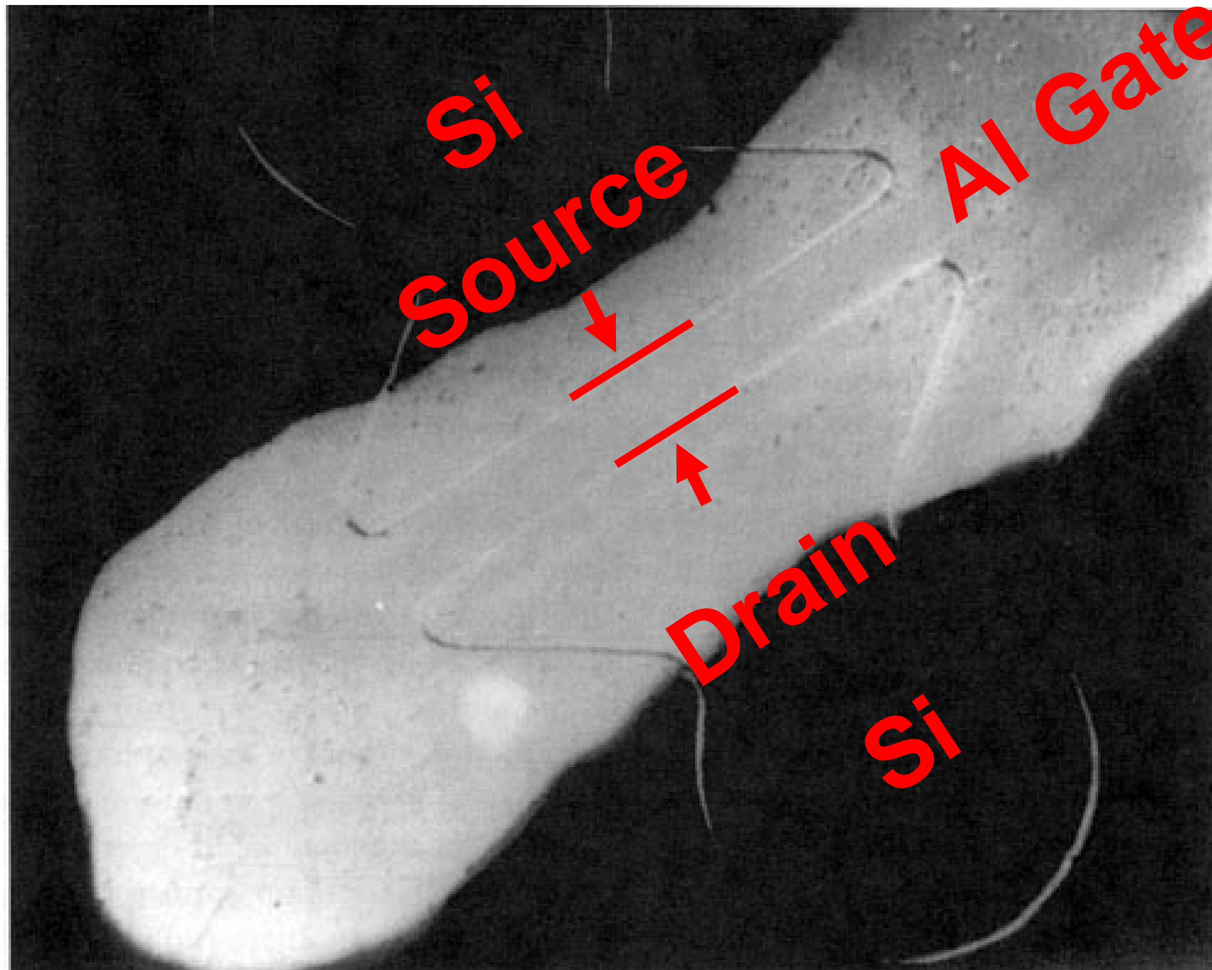
J. Bardeen, W. Bratten,

W. Shockley

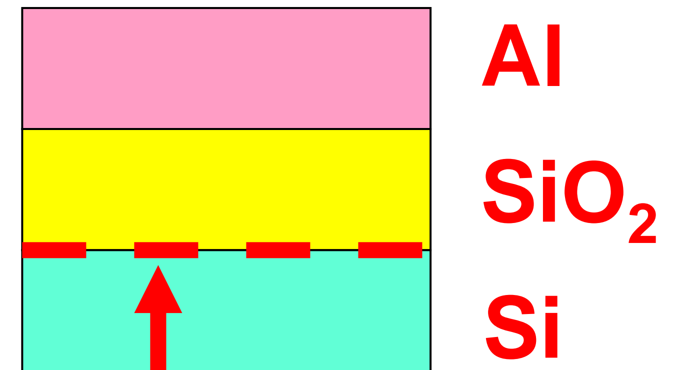


1960: First MOSFET by D. Kahng and M. Atalla

Top View



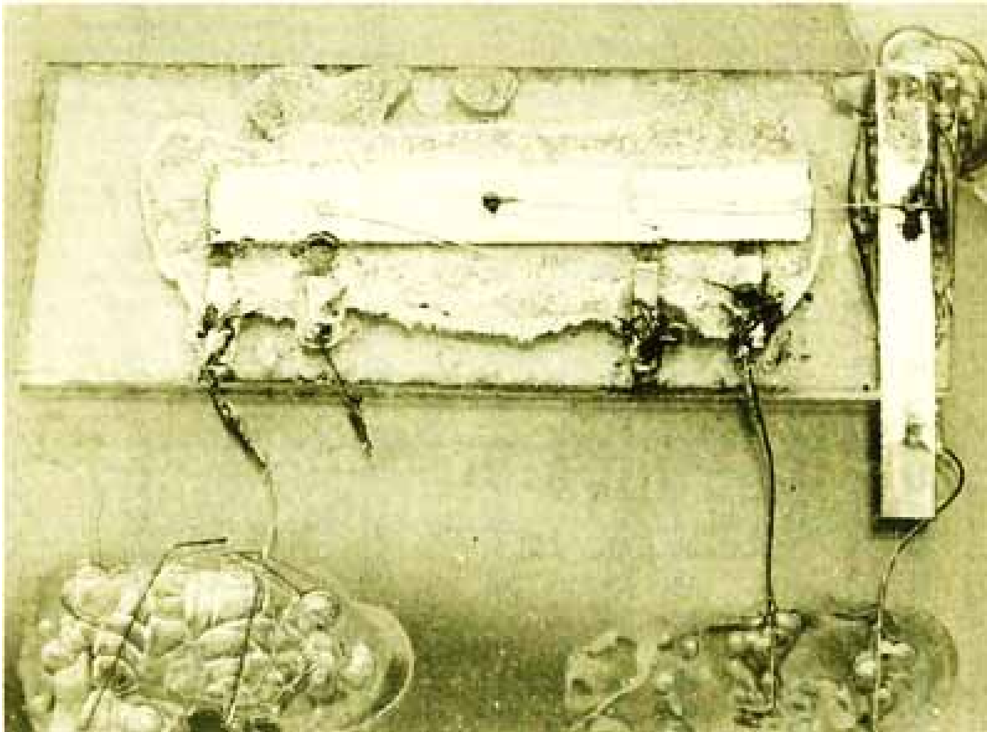
Cross-section



Exceptionally good interface!

Very small number of interface charges

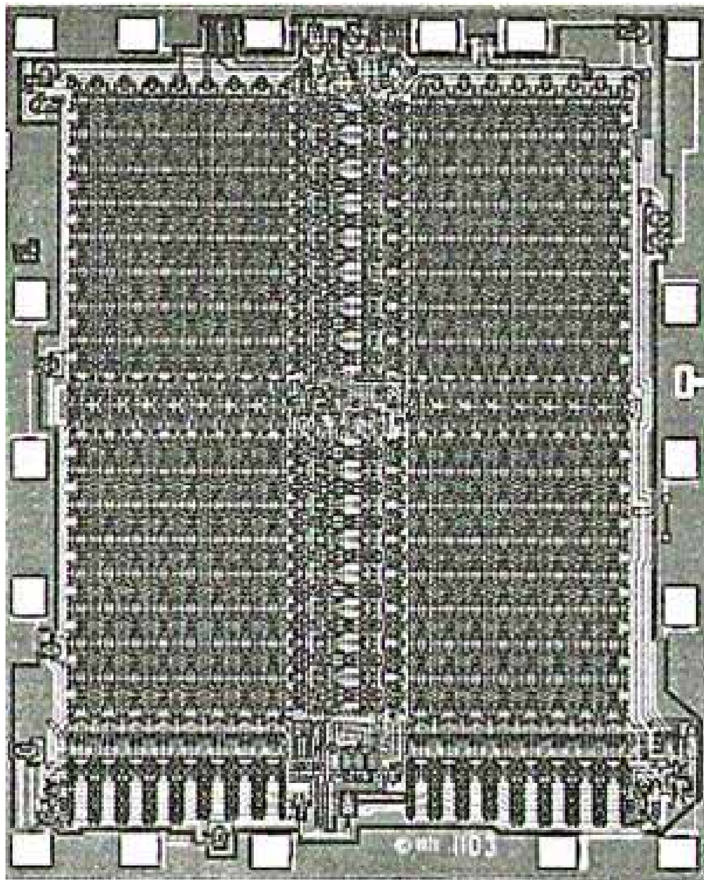
<u>1958: 1st Integrated Circuit</u>



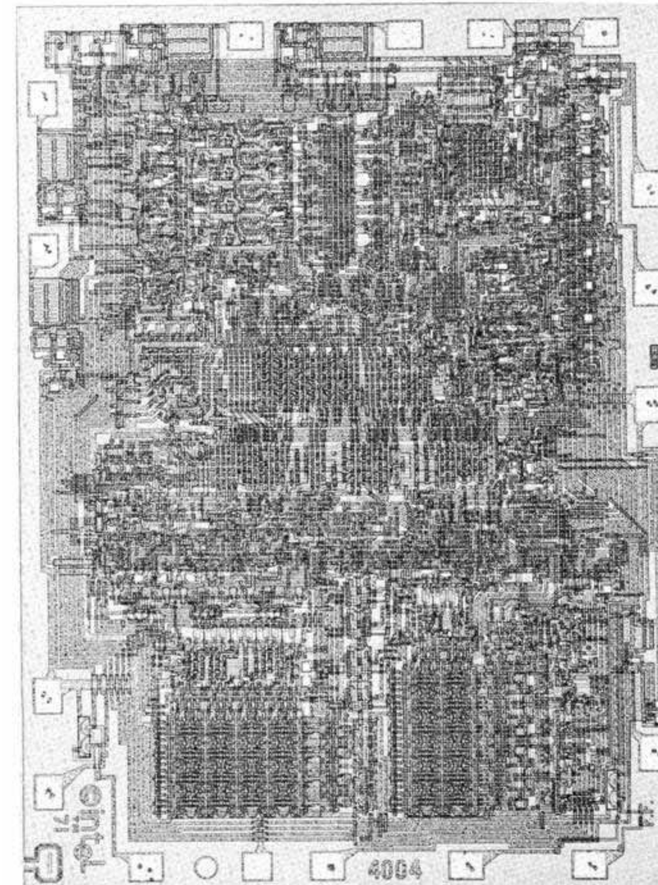
Jack S. Kilby

1970,71: 1st generation of LSIs

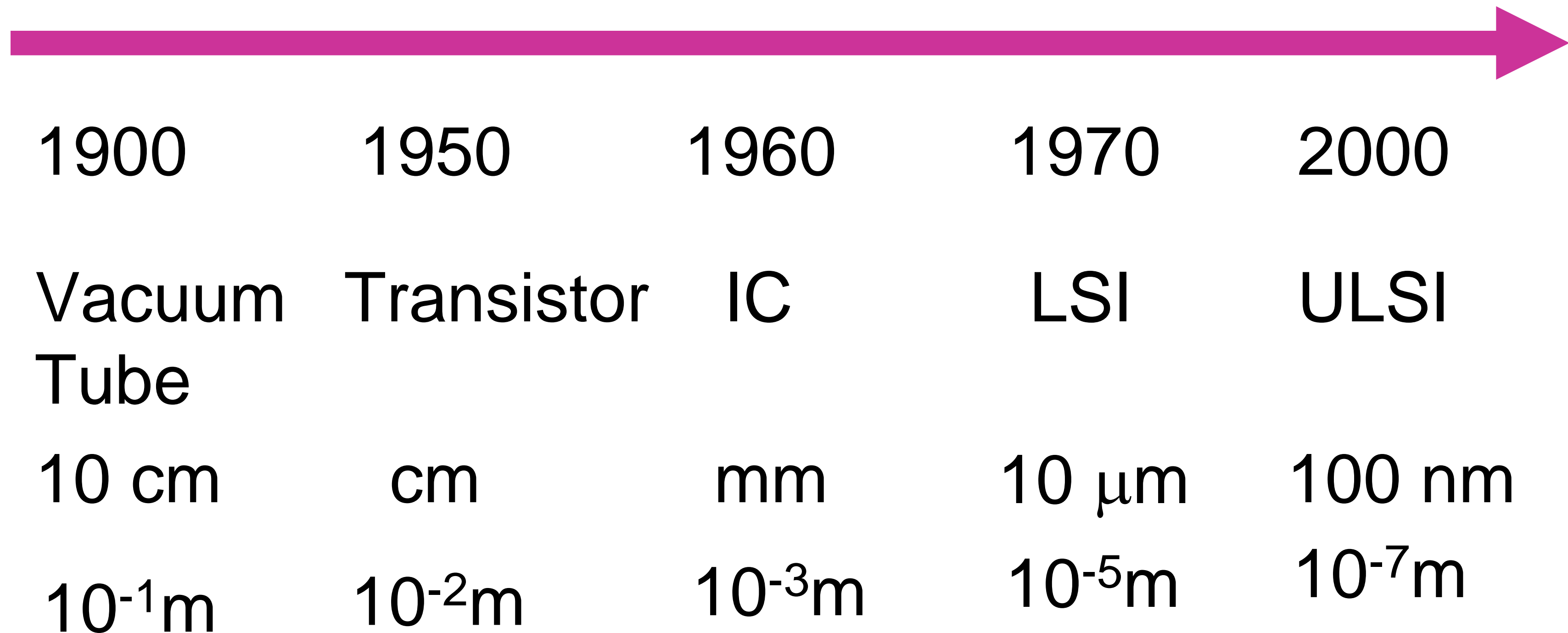
DRAM Intel 1103



MPU Intel 4004



Downsizing of the components has been the driving force for circuit evolution



In 100 years, the feature size reduced by one million times.
We have never experienced such a tremendous reduction in human history.

Downsizing

1. **Reduce Capacitance**

→ **Reduce switching time of MOSFETs**

Reduce power consumption

2. **Increase number of Transistors**

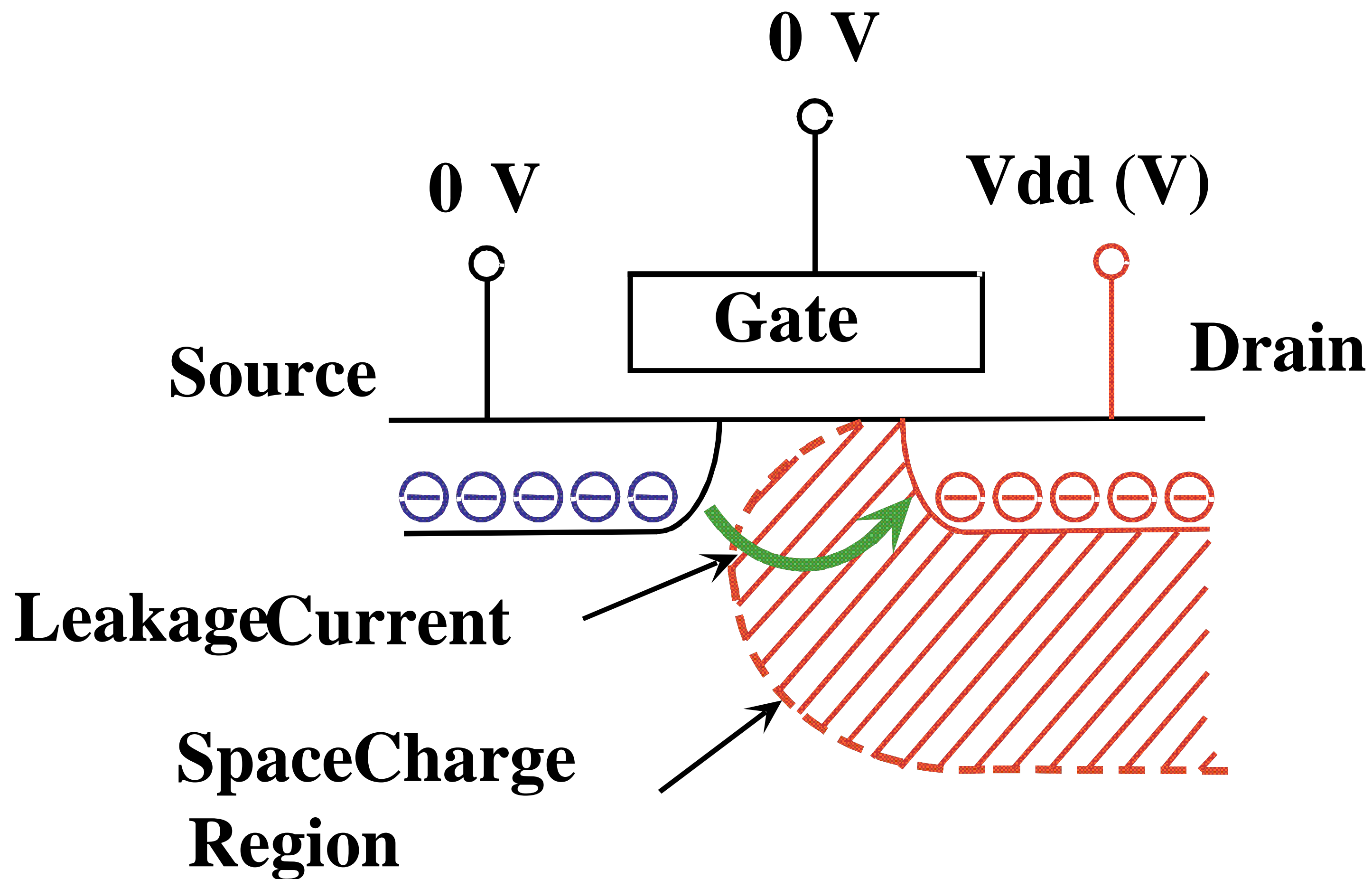
Increase functionality

→ Parallel processing

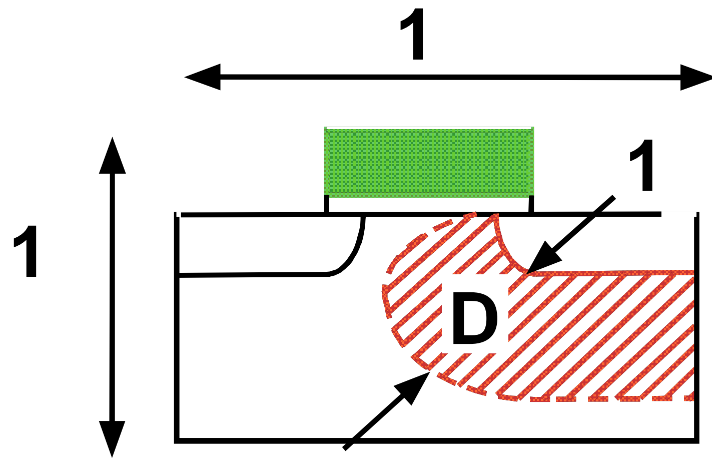
→ **Increase circuit operation speed**

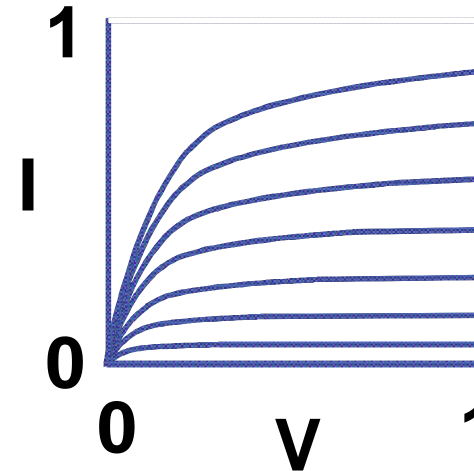
Thus, downsizing of Si devices is the most important and critical issue.

Short-channel effect at downsizing



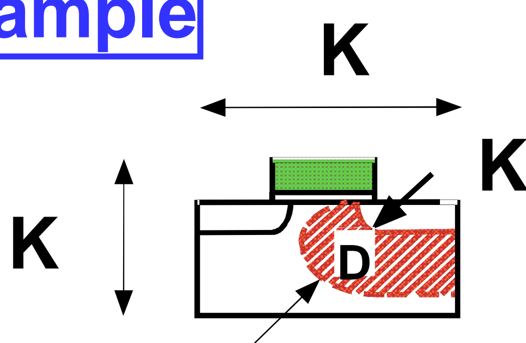
Scaling Method: by R. Dennard



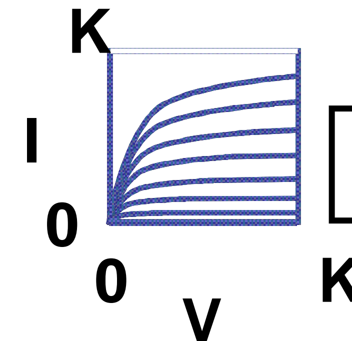


K=0.7 for example

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



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



I : K

R. Dennard et al. , IEEE J. Solid State Circuits, vol.9. P.256, 1974

Robert Dennad

Downscaling merit

Geometry & Supply voltage	L_g, W_g T_{ox}, V_d	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	$\tau = C_g V_d / I_d$ $\rightarrow$ KK/K= K
Clock frequency	f	1/K	$f = 1/\tau = 1/K$
Chip area	A_{chip}	α	Scaling α
Integration (# of Tr)	N	α/K^2	N $\rightarrow$ α/K^2
Power per chip	P	α	$fNCV^2/2$ $\rightarrow$ $K^{-1}(\alpha K^{-2})K(K^1)^2 = \alpha$

What will be real Downscaling ?

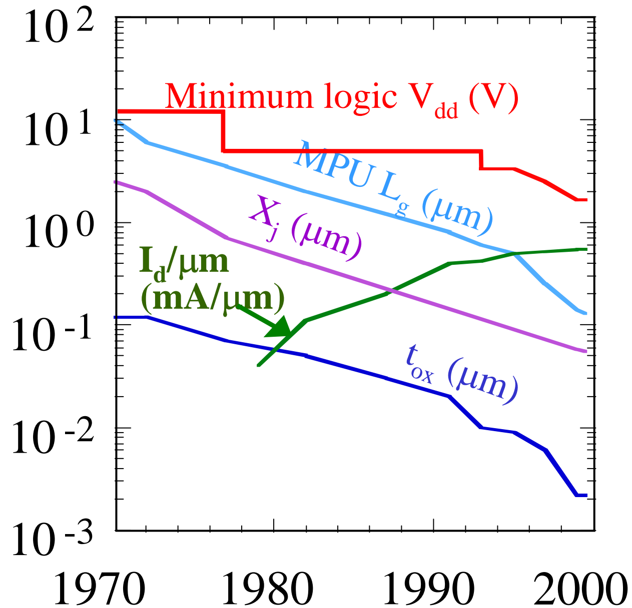
Is K the same for all the parameters?

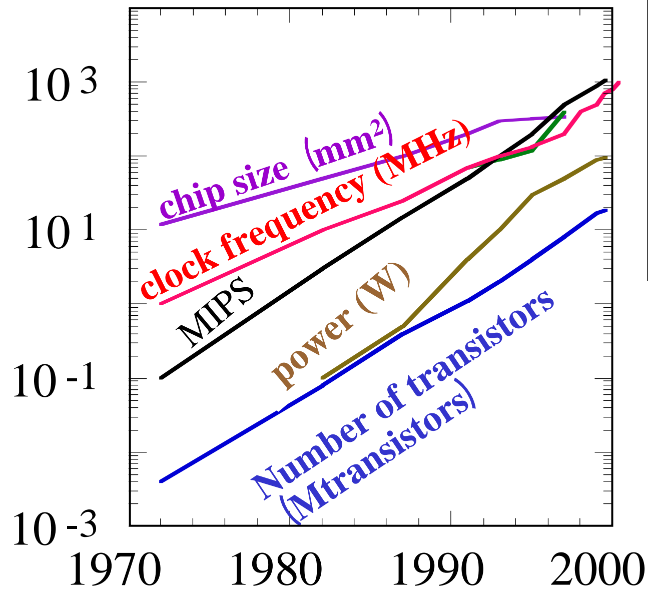
L_g, W_g, t_{ox}, V_d $\longrightarrow$ K? $A_{chip} \longrightarrow \alpha$?

$I_d \longrightarrow$ K? $I_d/\mu m \longrightarrow$ 1? f $\longrightarrow$ 1/K?

$C_g \longrightarrow$ K? $\tau \longrightarrow$ K? N $\longrightarrow \alpha/K^2$? P $\longrightarrow \alpha$?

Past downscaling trend





Past 30 years scaling

Merit: N, f increase

Demerit: P increase

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

Nano-CMOS:

What would be the
downsizing limit?

Many people wanted to say about the limit.
Past predictions were not correct!!

Period	Expected limit(size)	Cause
Late 1970's	1μm:	SCE
Early 1980's	0.5μm:	S/D resistance
Early 1980's	0.25μm:	Direct-tunneling of gate SiO_2
Late 1980's	0.1μm:	'0.1μm brick wall'(various)
2000	50nm:	'Red brick wall (various)
2000	10nm:	Fundamental?

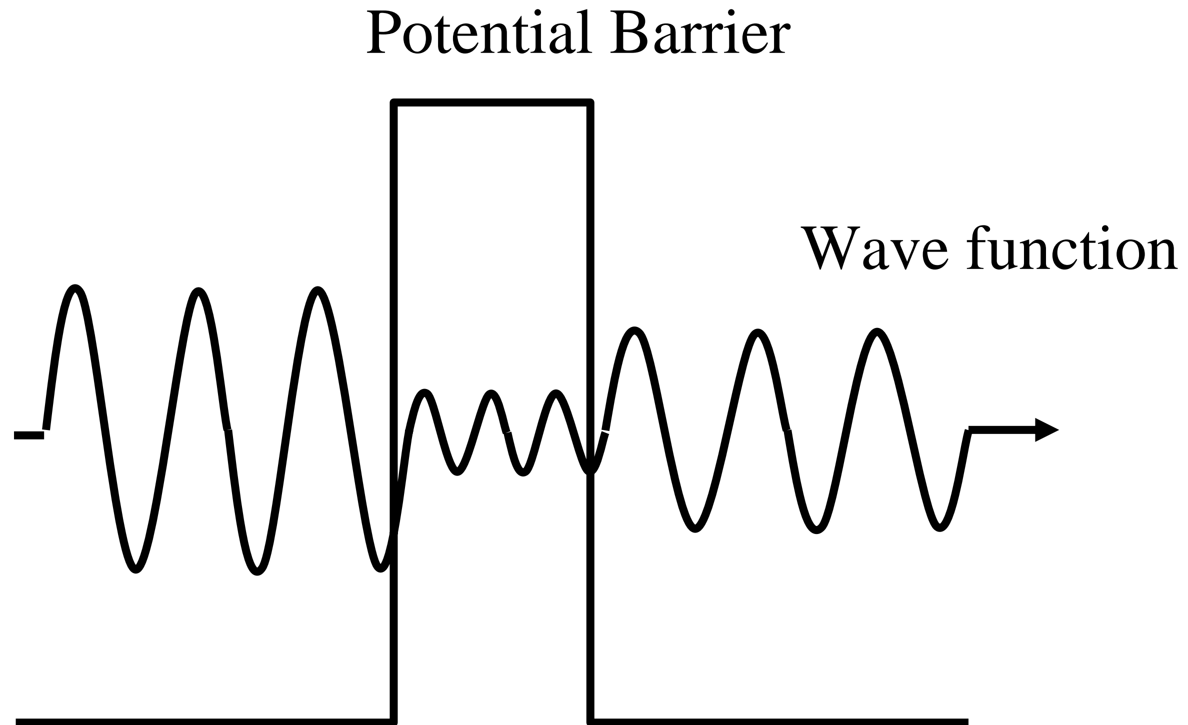
Historically, many predictions of the limit of downsizing.

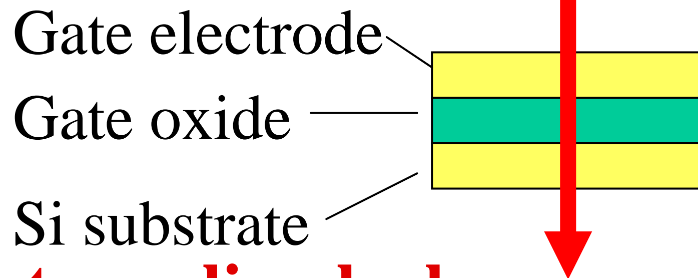
VLSI text book written 1979 predict that 0.25 micro-meter would be the limit because of direct-tunneling current through the very thin-gate oxide.

VLSI textbook

Finally, there appears to be a fundamental limit [10] of approximately quarter micron channel length, where certain physical effects such as the tunneling through the gate oxide and fluctuations in the positions of impurities in the depletion layers begin to make the devices of smaller dimension unworkable.

Direct-tunneling effect



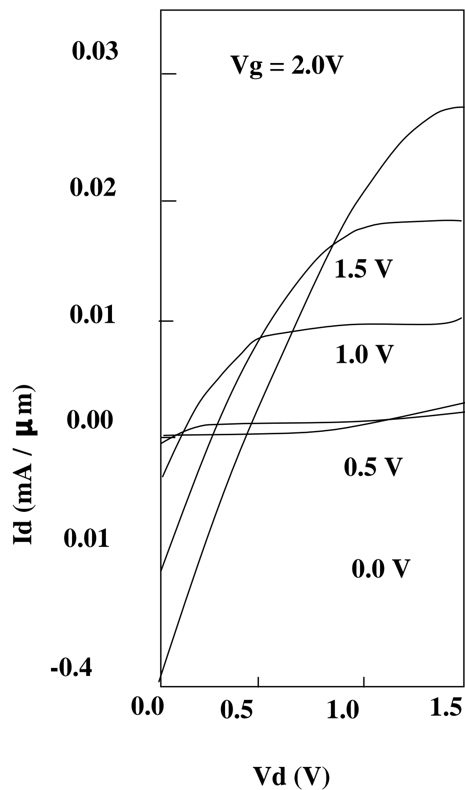


Direct tunneling leakage current start to flow when the thickness is 3 nm.

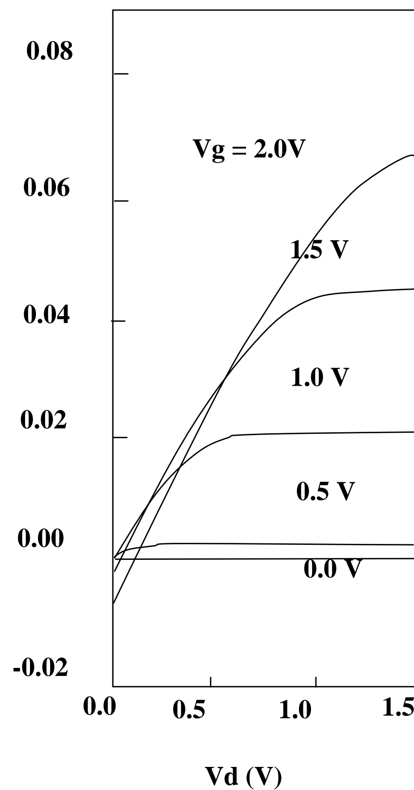
Direct tunneling leakage was found to be OK! In 1994!

MOSFETs with 1.5 nm gate oxide

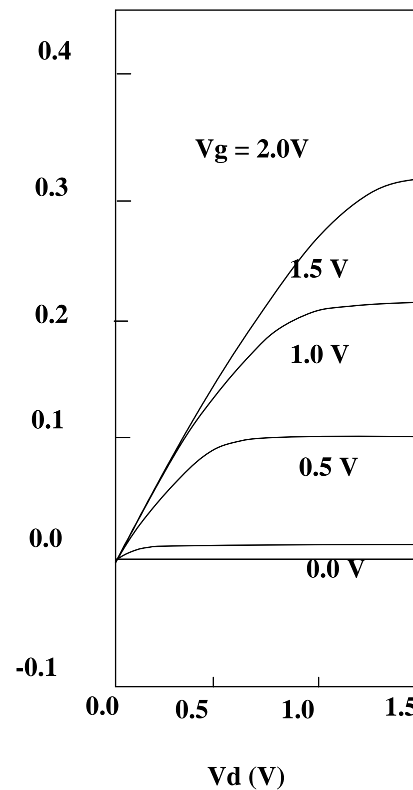
Lg = 10 μm



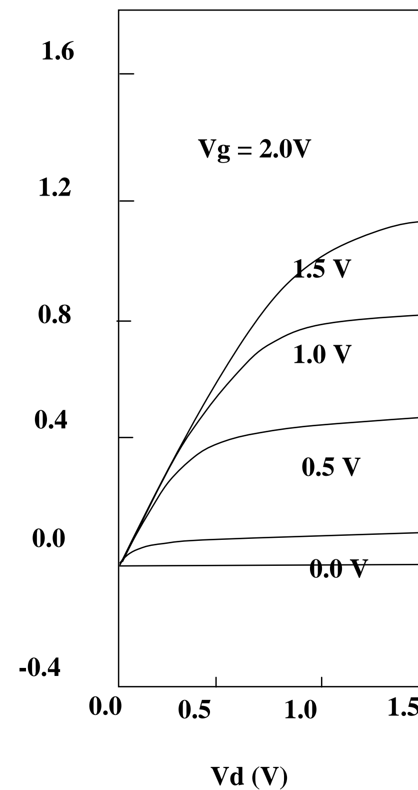
Lg = 5 μm



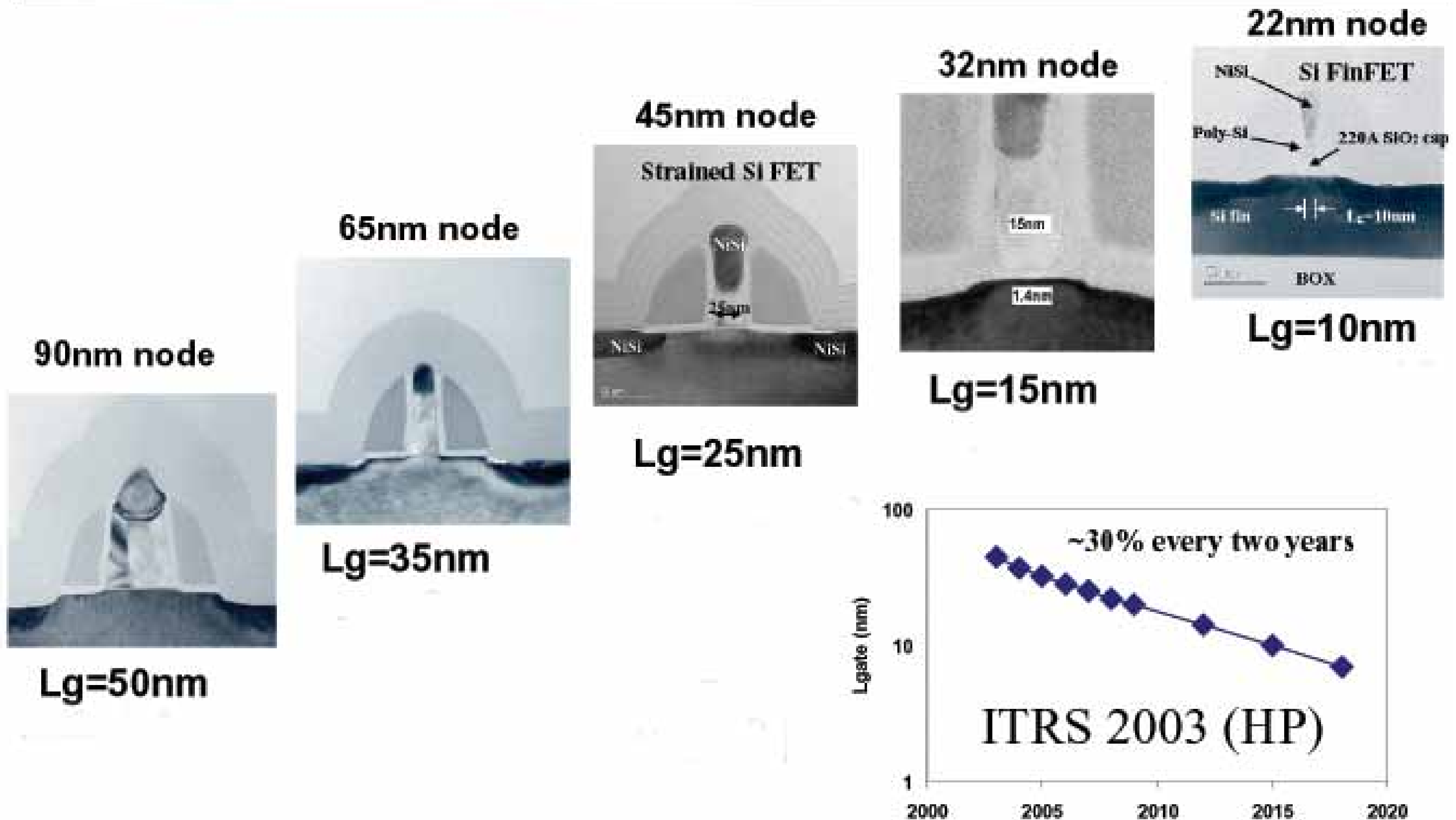
Lg = 1.0 μm



Lg = 0.1μm



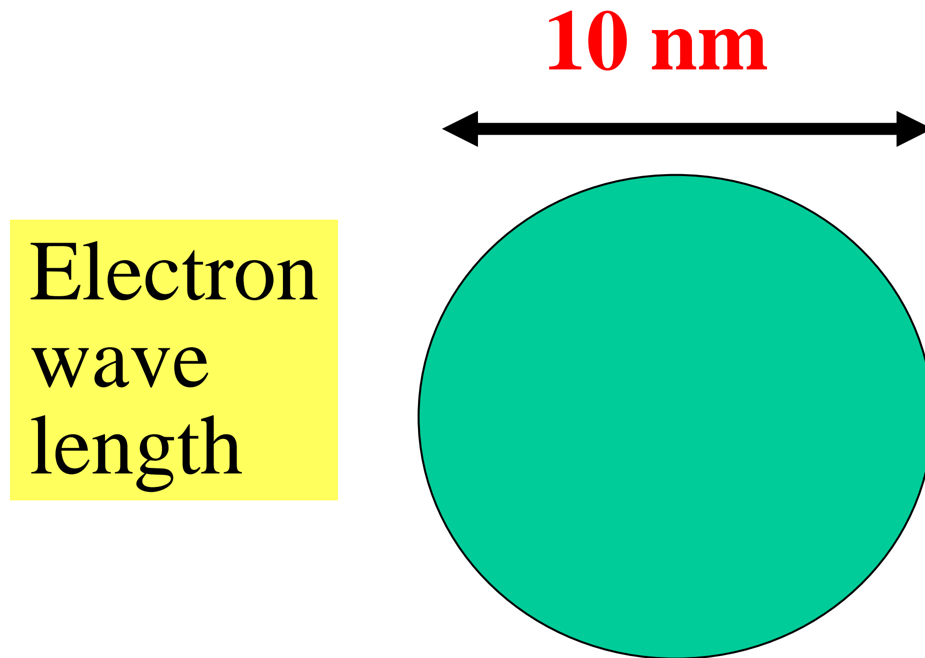
Transistor Scaling Continues

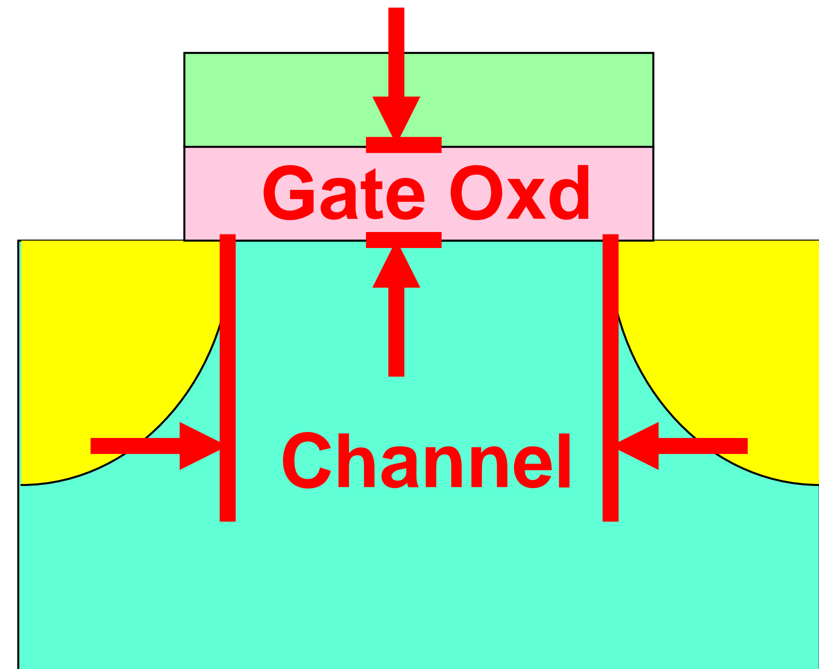


Qi Xinag, ECS 2004, AMD

Downsizing limit?

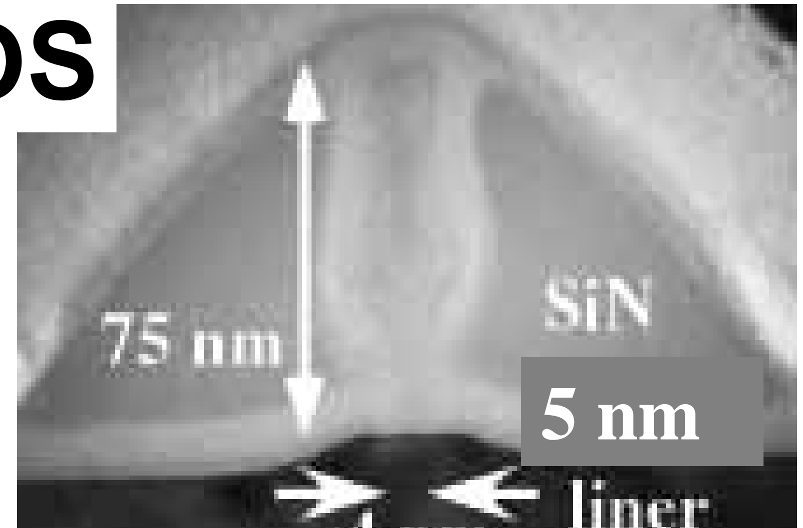
Channel length?



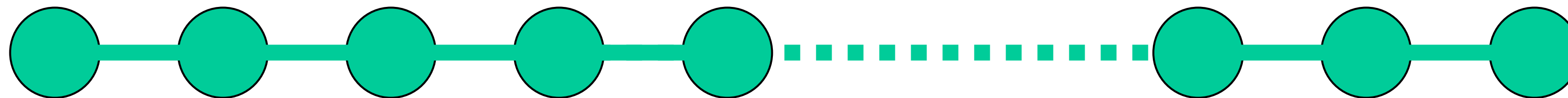


5 nm gate length CMOS

Is a Real Nano Device!!

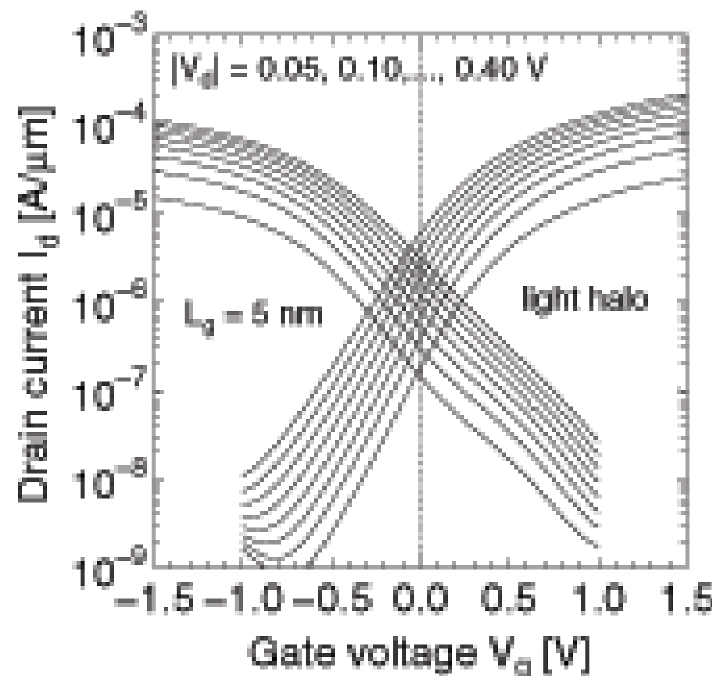


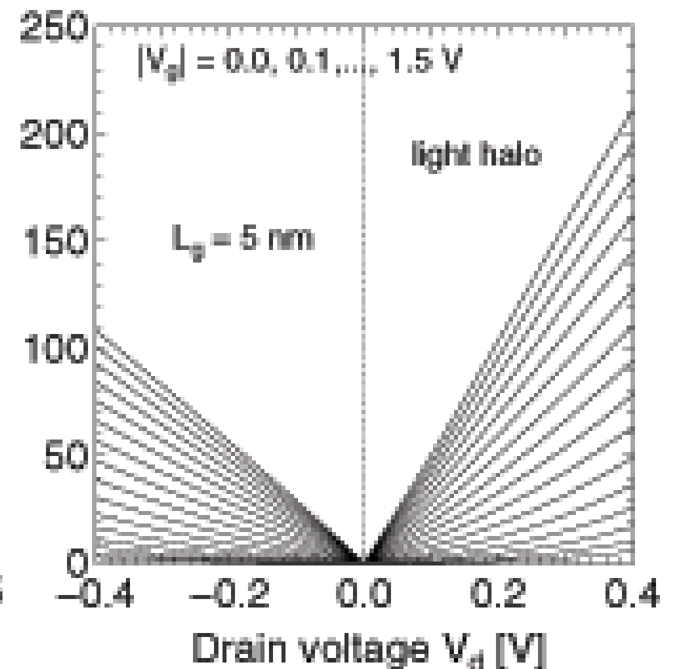
Length of 18 Si atoms



H. Wakabayashi et.al, NEC

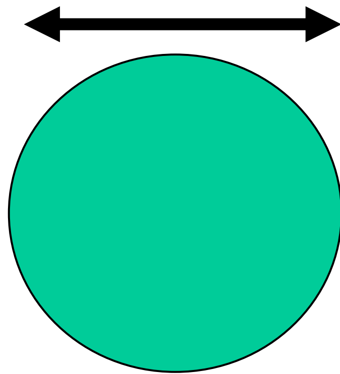
IEDM, 2003





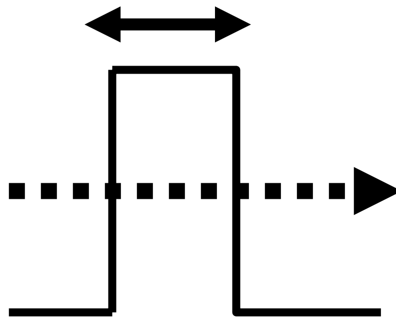
Electron wave length

10 nm



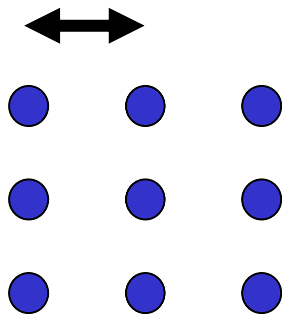
Tunneling distance

3 nm



Atom distance

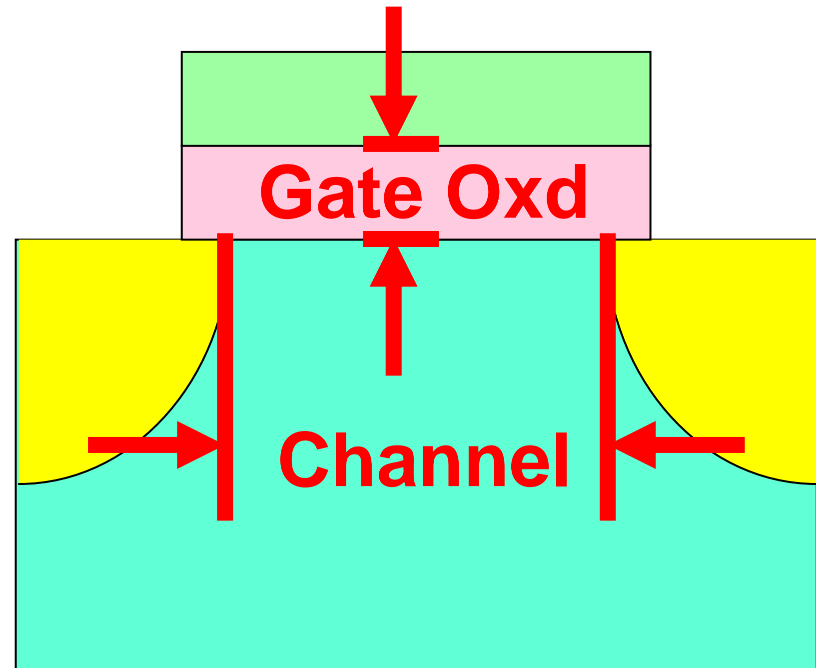
0.3 nm



Downsizing limit!

Channel length

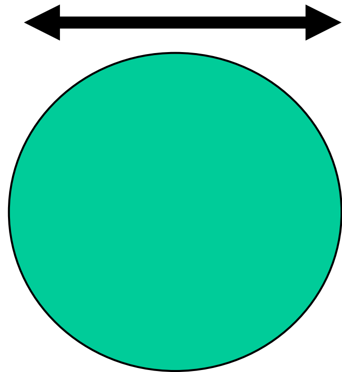
Gate oxide thickness



Prediction now!

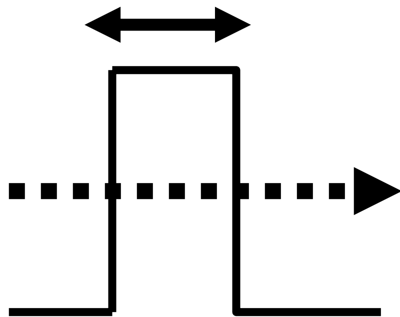
Electron wave length

10 nm



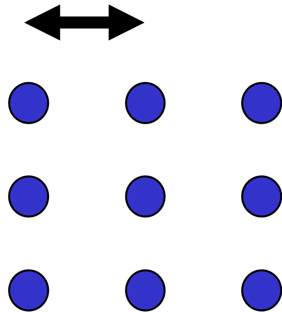
Tunneling distance

3 nm



Atom distance

0.3 nm



MOSFET operation

Lg = 2 ~ 1.5 nm?

But, no one knows future!

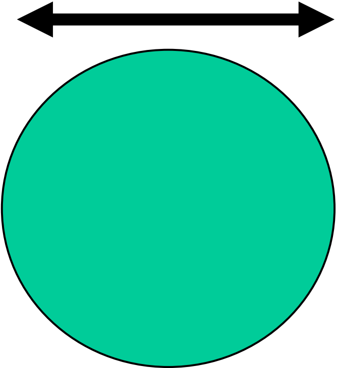
Prediction now!

Gate length

Prediction at present

Electron wave length

10 nm

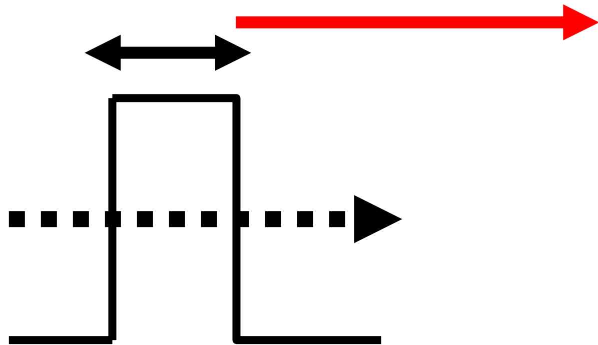


Practical limit because of off-leakage between S and D?

Lg = 5 nm?

Tunneling distance

3 nm

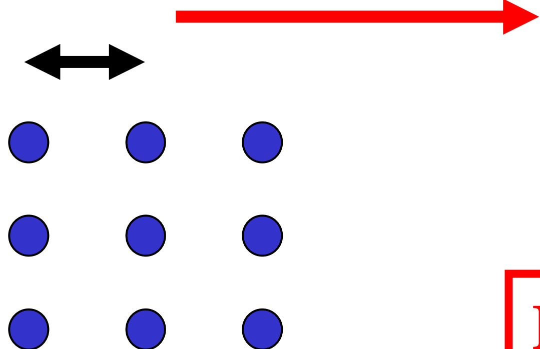


MOSFET operation

Lg = 2 ~ 1.5 nm?

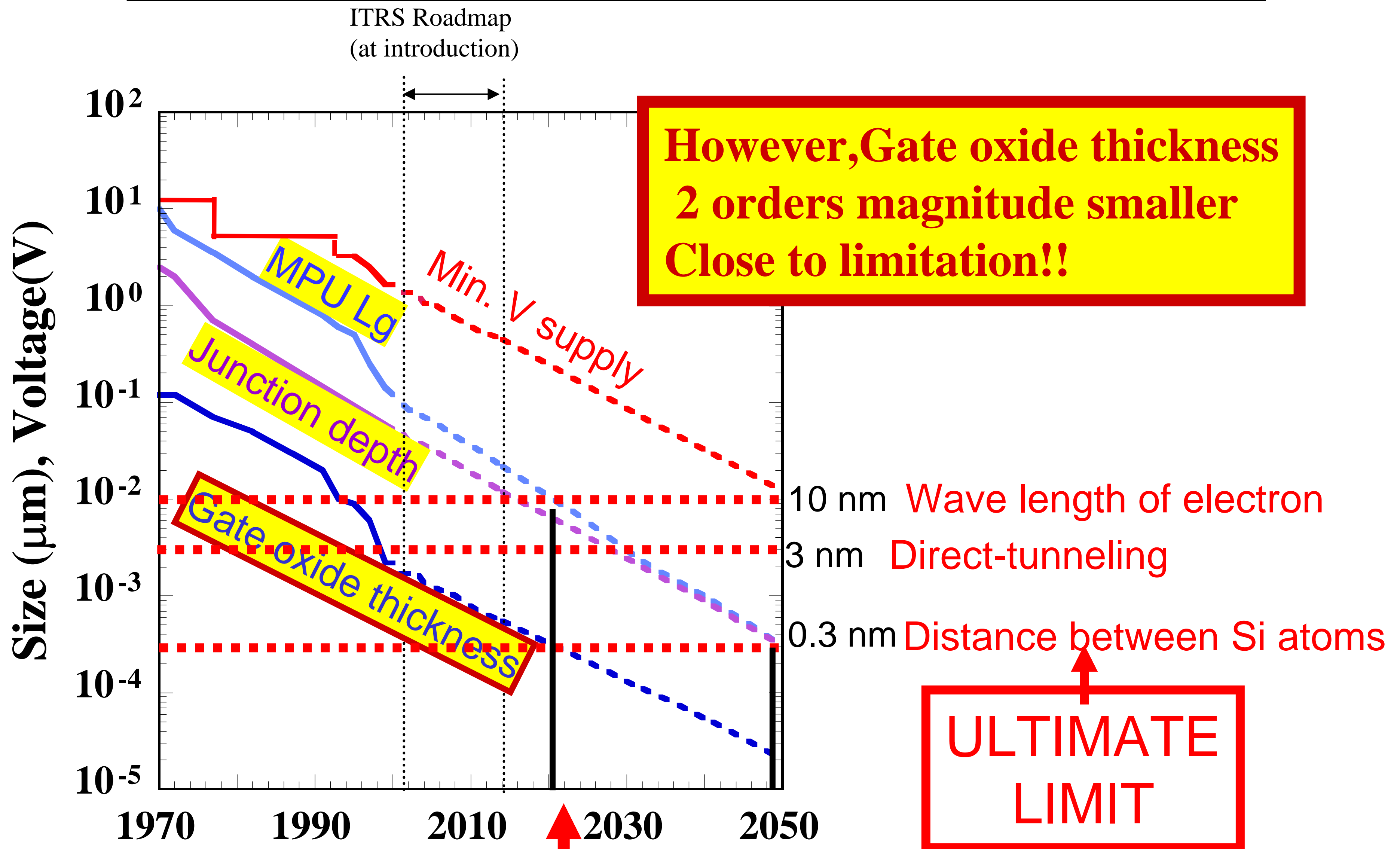
Atom distance

0.3 nm



But, no one knows future!

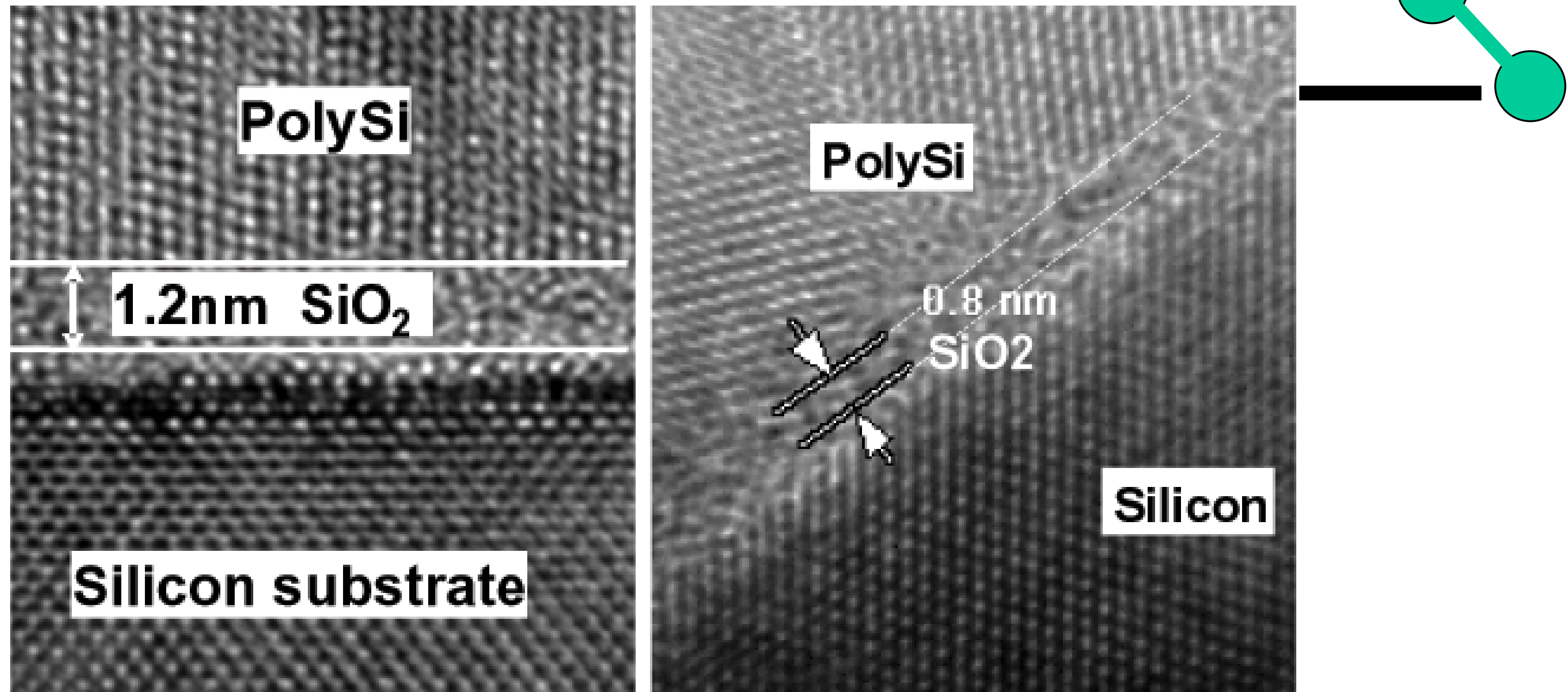
Ultimate limitation



Lg: Gate length downsizing will continue to another 10-15 years

0.8 nm Gate Oxide Thickness MOSFETs operates!!

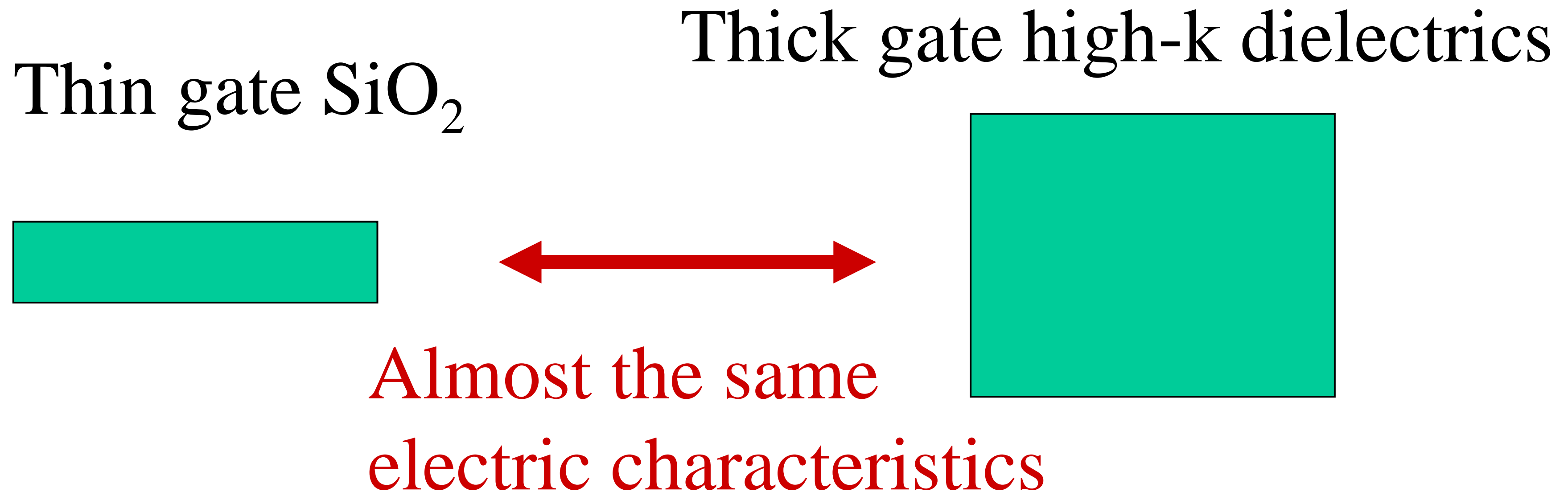
0.8 nm: Distance of 3 Si atoms!!



- 1.2nm physical SiO2 in production (90nm logic node)
- 0.8nm physical SiO2 in research transistors

By Robert Chau, IWGI 2003

There is a solution!
To use high-k dielectrics



However, very difficult and big challenge!

Remember MOSFET had not been realized without Si/SiO_2!

Choice of High-k

Candidates ■

● Gas or liquid at 1000 K

○ Radio active

Unstable at Si interface

① Si + MO_X M + SiO_2

② Si + MO_X MSi_X + SiO_2

③ Si + MO_X M + MSi_XO_Y

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
● H																	● He
Li	Be											● B	● C	● N	● O	● F	● Ne
① Na	Mg											Al	Si	● P	● S	● Cl	● Ar
K	Ca	Sc	② Ti	① V	① Cr	① Mn	① Fc	① Co	① Ni	① Cu	① Zn	① Ga	① Ge	● As	● Se	● Br	● Kr
● Rh	Sr	Y	Zr	① Nb	① Mo	Tc	① Ru	① Rb	① Pd	● Ag	① Cd	① In	① Sn	① Sb	① Te	● I	● Xe
● Cs	③ Ba	★	Hf	① Ta	① W	① Re	① Os	① Ir	● Pt	● Au	● Hg	● Tl	① Pb	① Bi	○ Po	○ At	○ Rn
○ Fr	○ Ra	☆	○ Rf	○ Ha	○ Sg	○ Ns	○ Hs	○ Mt									

★	La	Ce	Pr	Nd	○ Pm	Sm	Eu	Gd	Tb	Dy	Ho	Er	Tm	Yb	Lu
☆	○ Ac	○ Th	○ Pa	○ U	○ Np	○ Pu	○ Am	○ Cm	○ Bk	○ Cf	○ Es	○ Fm	○ Md	○ No	○ Lr

HfO_2 based dielectrics are selected as the first generation materials, because of their merit in
1) band-offset,
2) dielectric constant
3) thermal stability

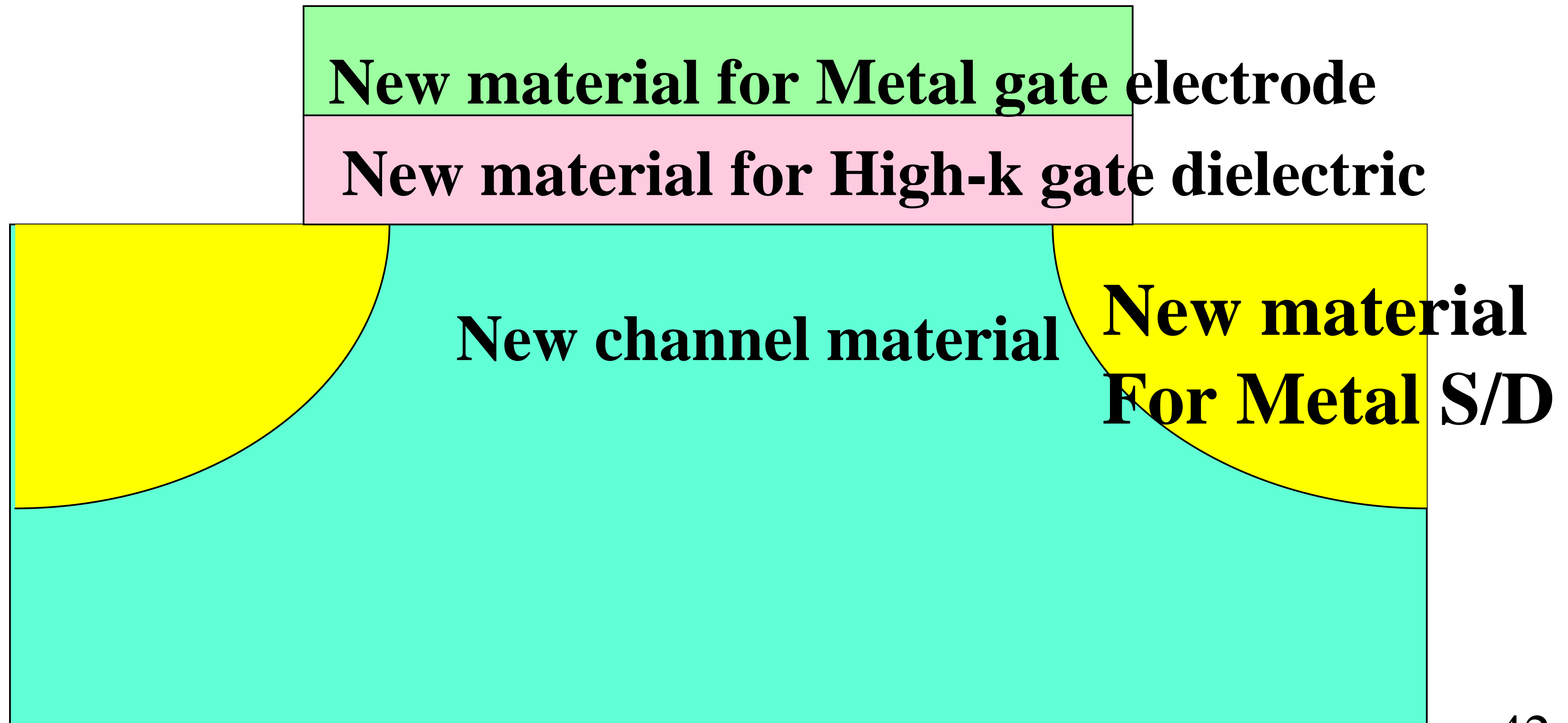
La_2O_3 based dielectrics are thought to be the next generation materials, which may not need a thicker interfacial layer

R. Hauser, IEDM Short Course, 1999

Hubbard and Schlom, J Mater Res 11 2757 (1996)

New material research will give us many future possibilities and the most important for Nano-CMOS!

Not only for high-k!



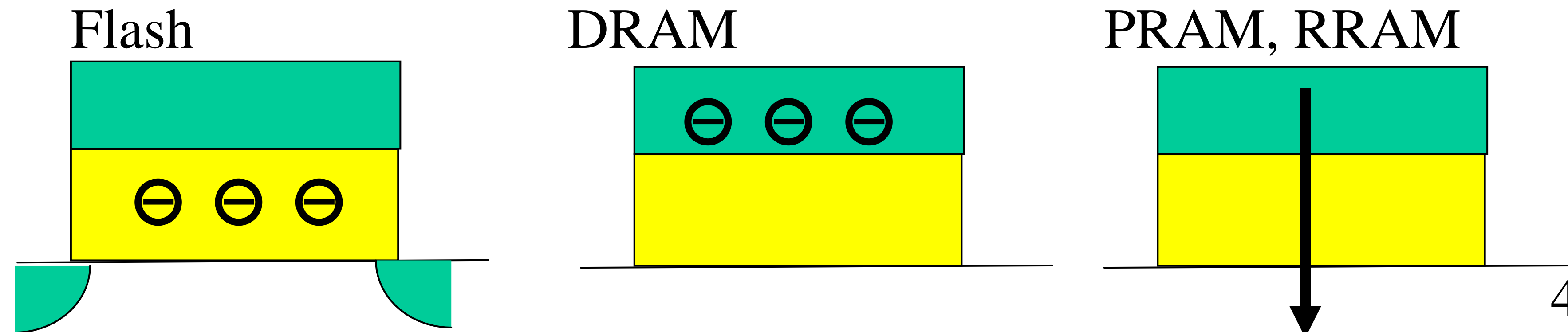
New materials are important for Not only nano-CMOS logic MOSFETs, But also for new memories!

Limit of high-density memories, such as flash, DRAM will be solved by new materials.

Flash: floating gate → gate insulator charge trap like SONOS, MNOS

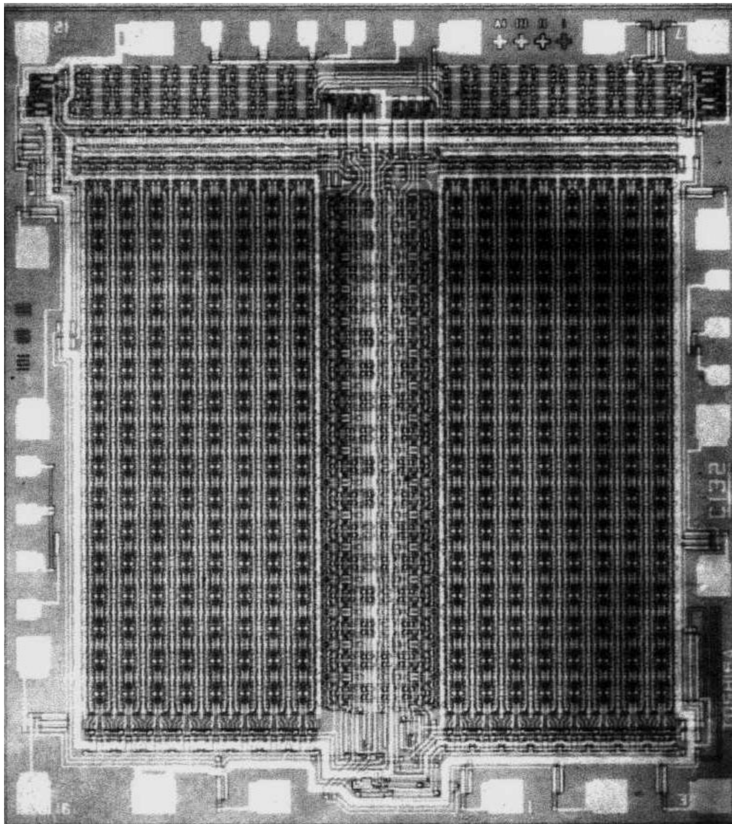
DRAM → New high-k insulator

New memory → PRAM, RRAM

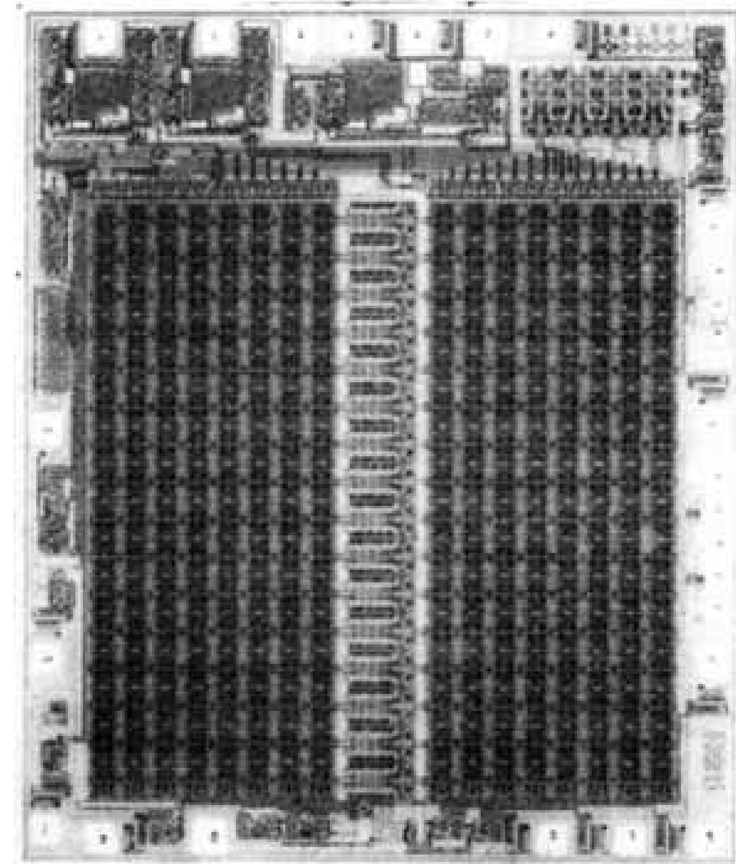


1970s: 10 years after single MOSFETs,

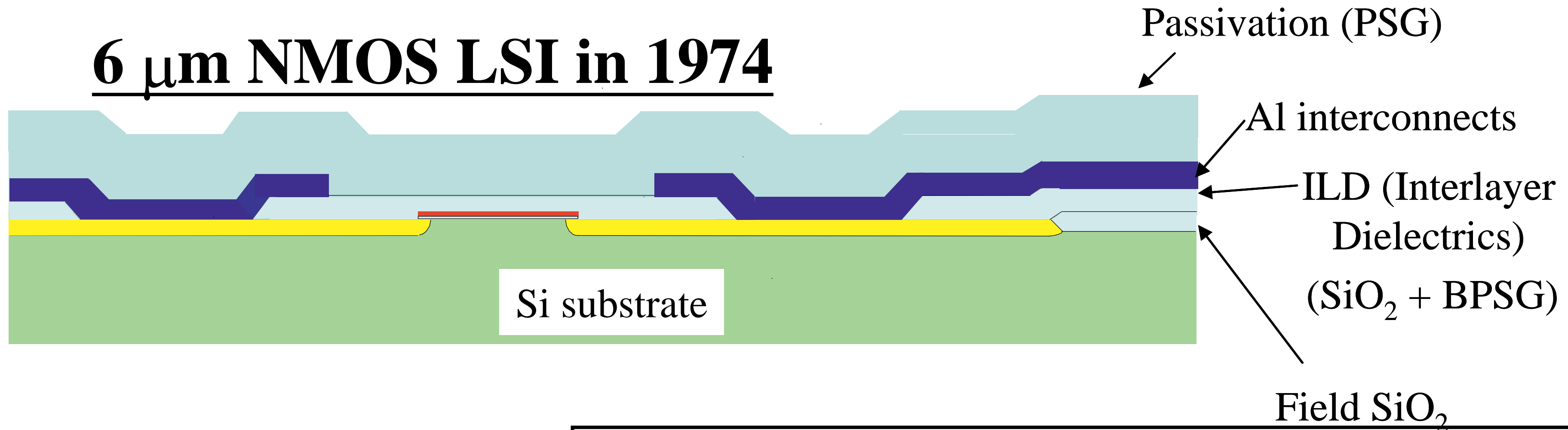
PMOS 1kbit DRAM Toshiba(1974)



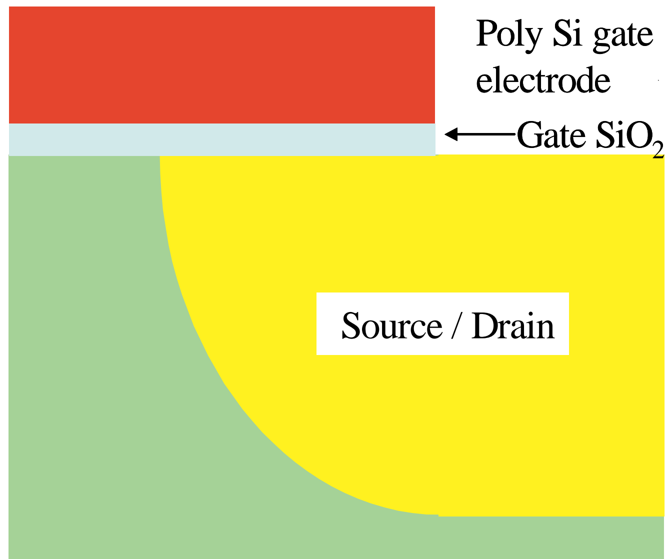
NMOS 1k bit SRAM Toshiba (1974)



6 μm NMOS LSI in 1974



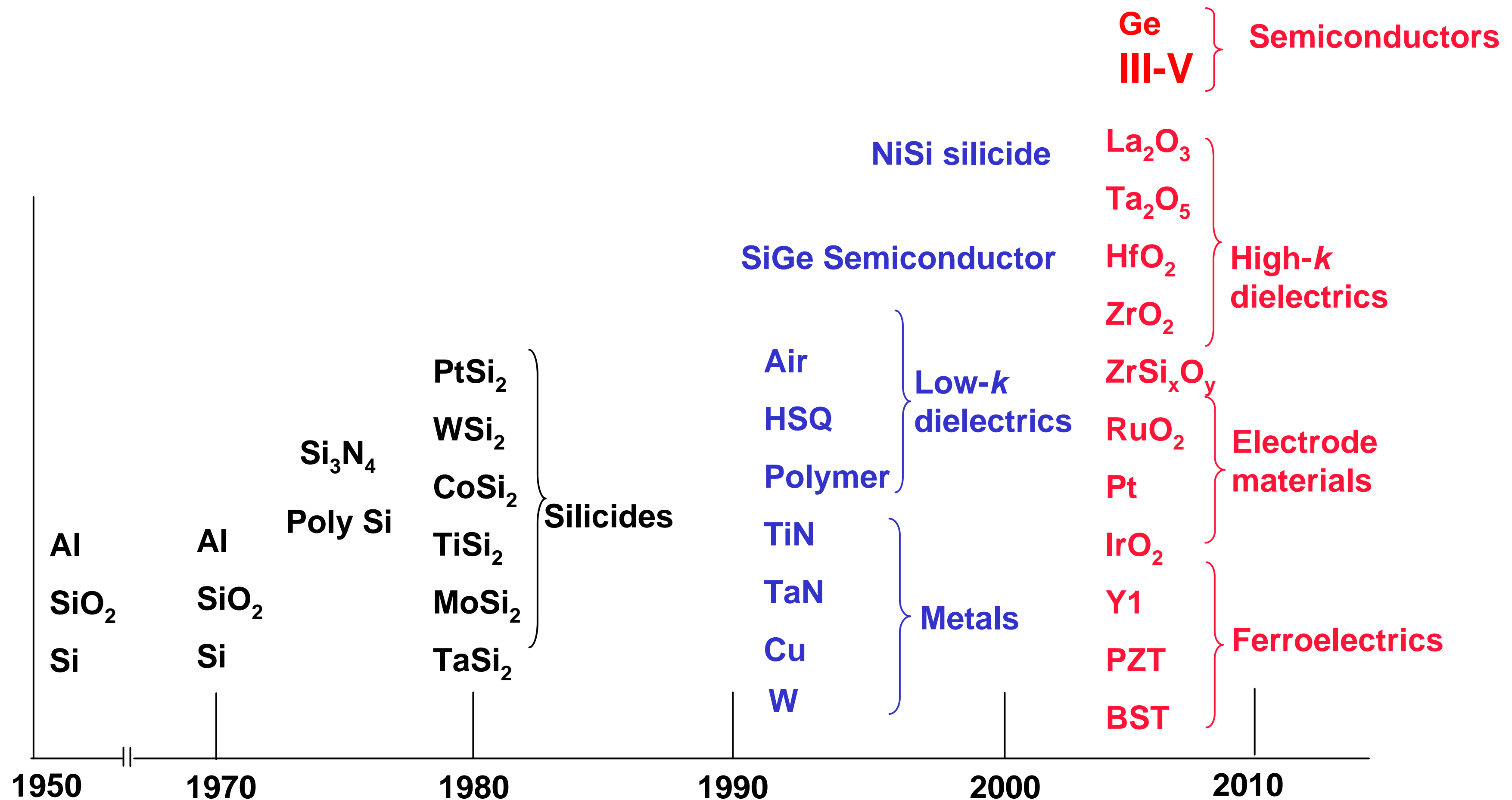
magnification



Layers	Materials	Atoms
1. Si substrate	1. Si	1. Si
2. Field oxide	2. SiO_2	2. O
3. Gate oxide	3. BPSG	3. P
4. Poly Si	4. Al	4. B
5. S/D	5. PSG	5. Al
6. Interlayer		(H, N, Cl)
7. Aluminum		
8. Passivation		

Just examples!
Many other candidates

New materials

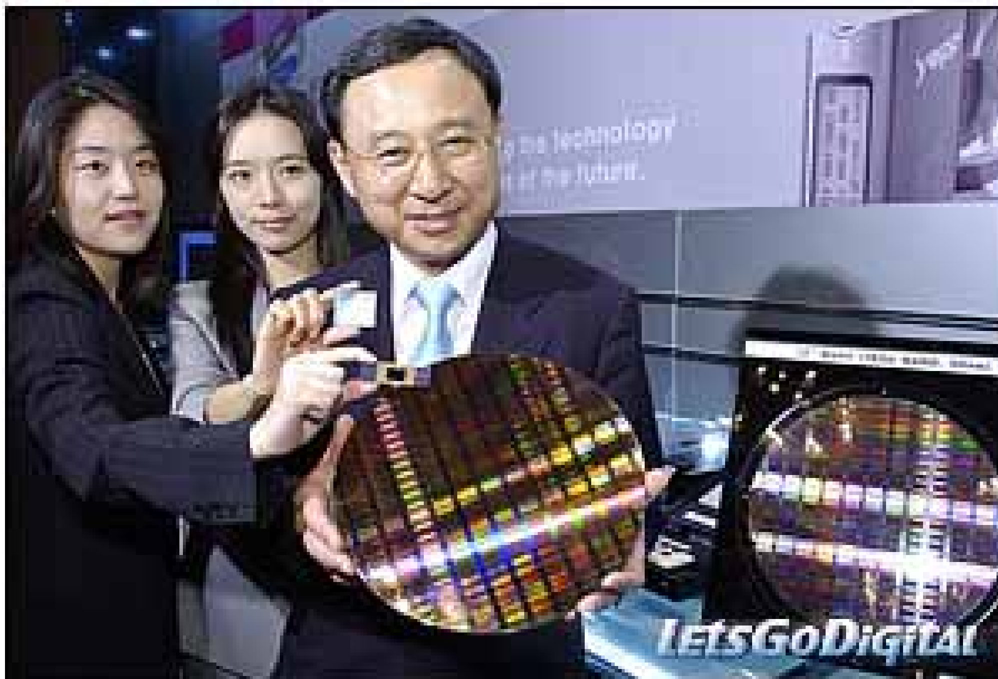


Y. Nishi, Si Nano Workshop, 2006,

(S. Sze, Based on invited talk at Stanford Univ., Aug. 1999)

Now: After 45 Years from the 1st single MOSFETs

32 Gb and 16Gb NAND, SAMSUNG

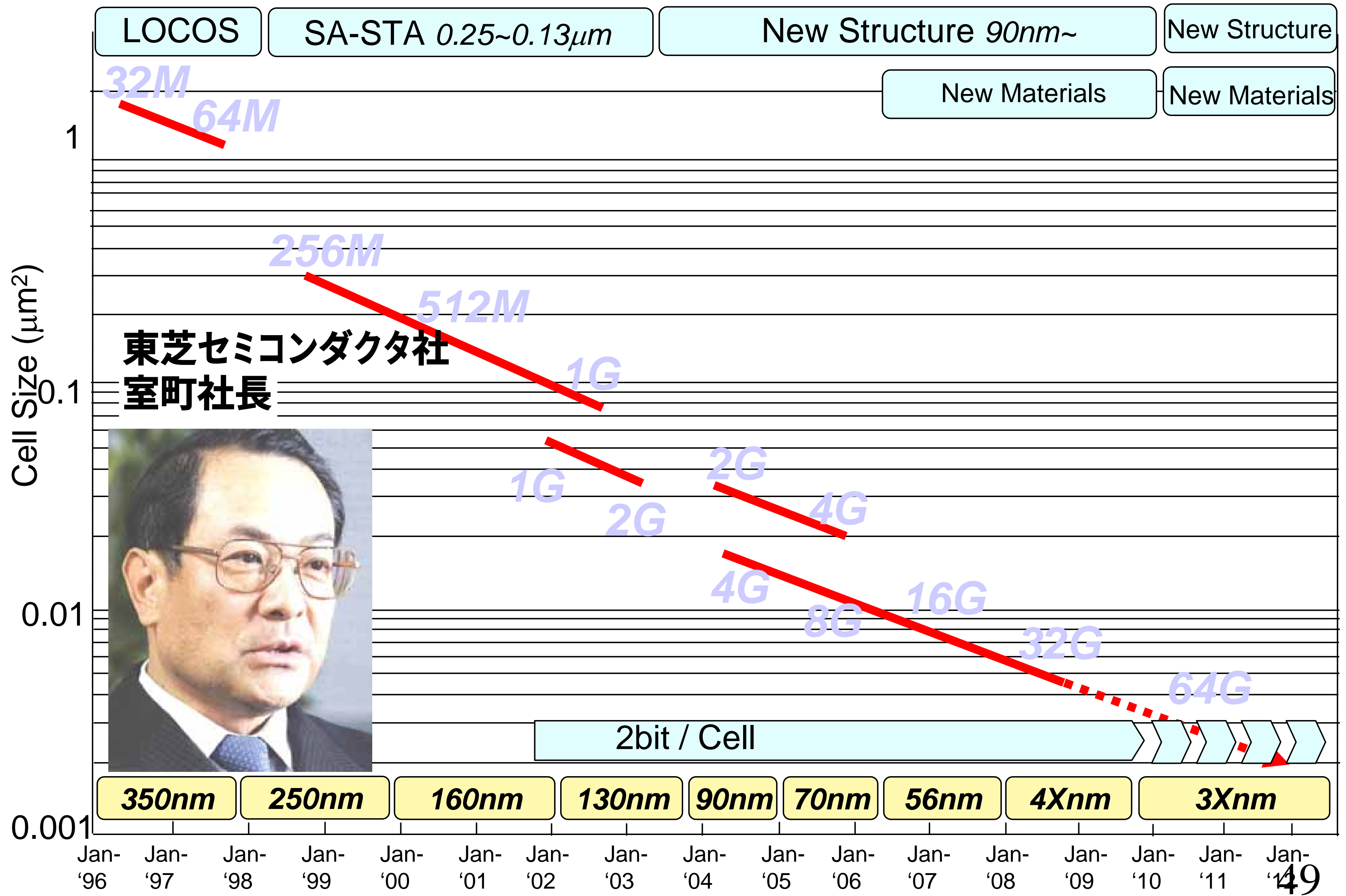


Samsung's NAND flash trend

Capacity	Node	Announcement	Product
512Mbit	120nm	2000	2001
1Gbit	100nm	2001	2002
2Gbit	90nm	2002	2003
4Gbit	70nm	2003	2004
8Gbit	60nm	2004	2005
16Gbit	50nm	2005	2006
32Gbit	40nm	2006	2007?
256Gbit	20nm	2010	2011?

Even Tbit is possible!

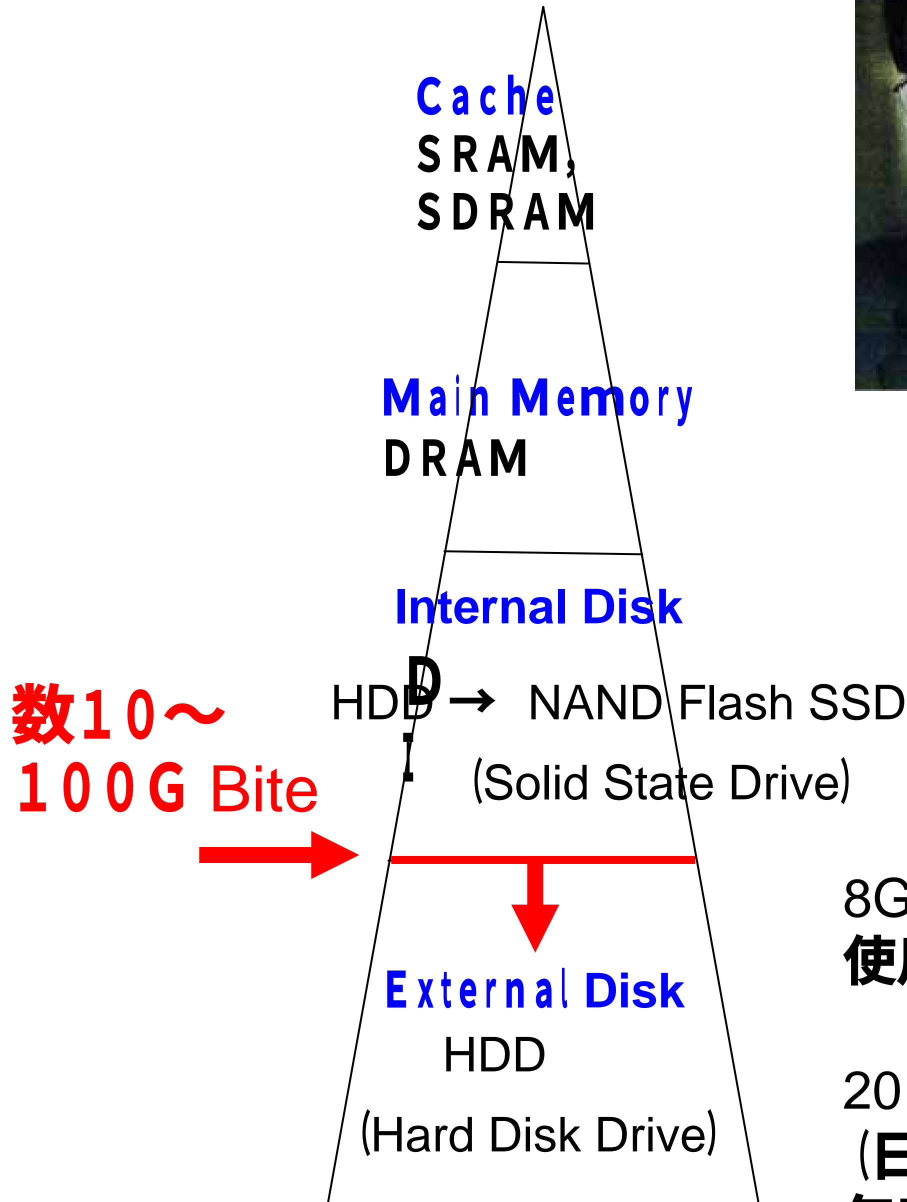
東芝のNAND Flashの製品ロードマップ (ISTF2006)



Already 32 Gbit:

larger than that of world population
comparable for the numbers of neurons
in human brain

Samsung announced 256 Gbit will be produced in 2010.
Only 4 years from now.
256Gbit: larger than those of # of stars in galaxies



NADN Flash発明者
東北大学　舛岡教授

ようやくNAND FlashのＳＳＤが
PCのDiskに２００７年から
使用され始めた

→　PCを軽く、起動時間を短く、
消費電力を少なく

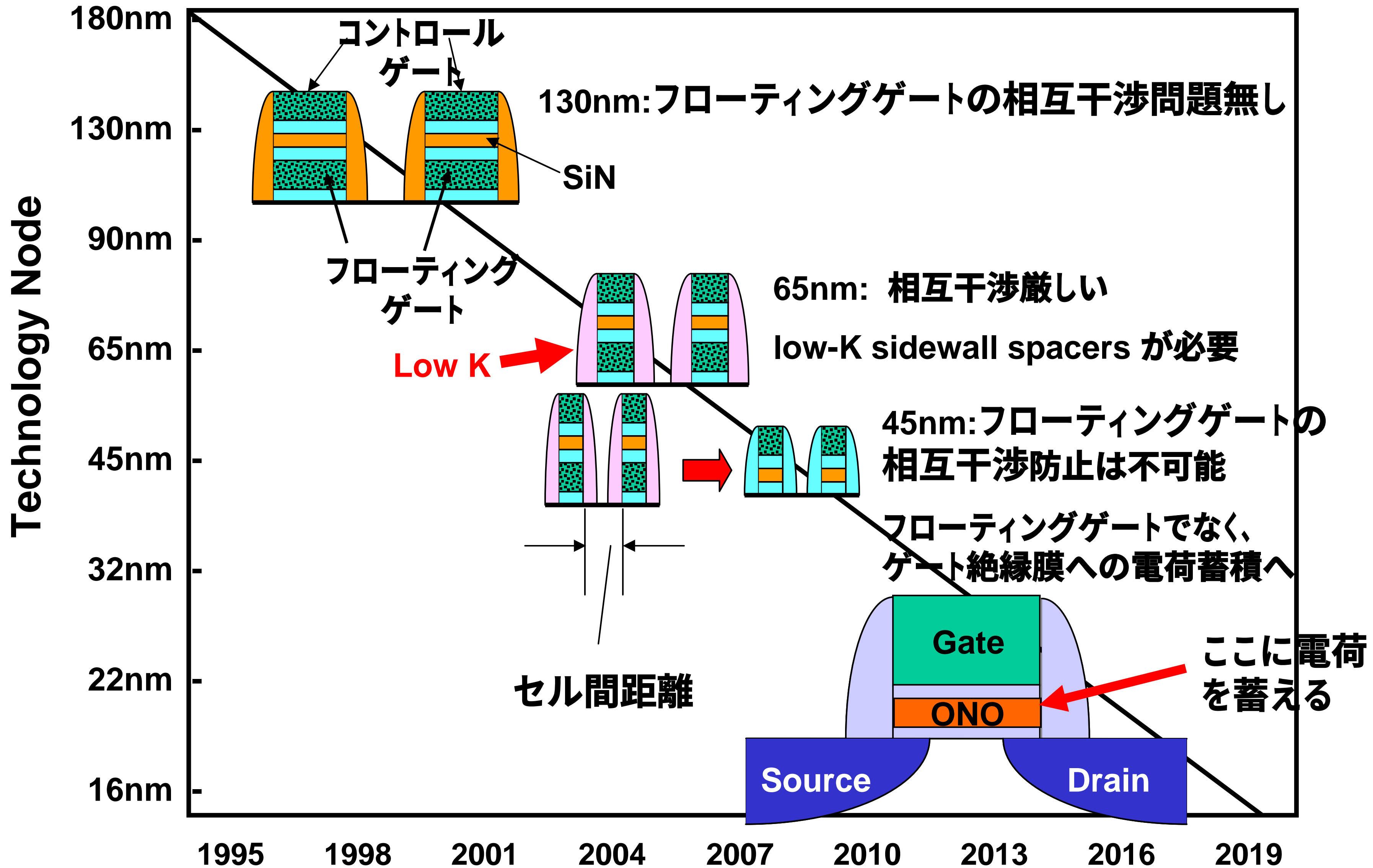
8G～１６Gbit NAND Flashメモリを
使用で32~64GBiteのＳＳＤ

2010年頃には１００ＧBite 1万円の売値
（日経エレクトロニクス）
年率５０%の製造コストの低下

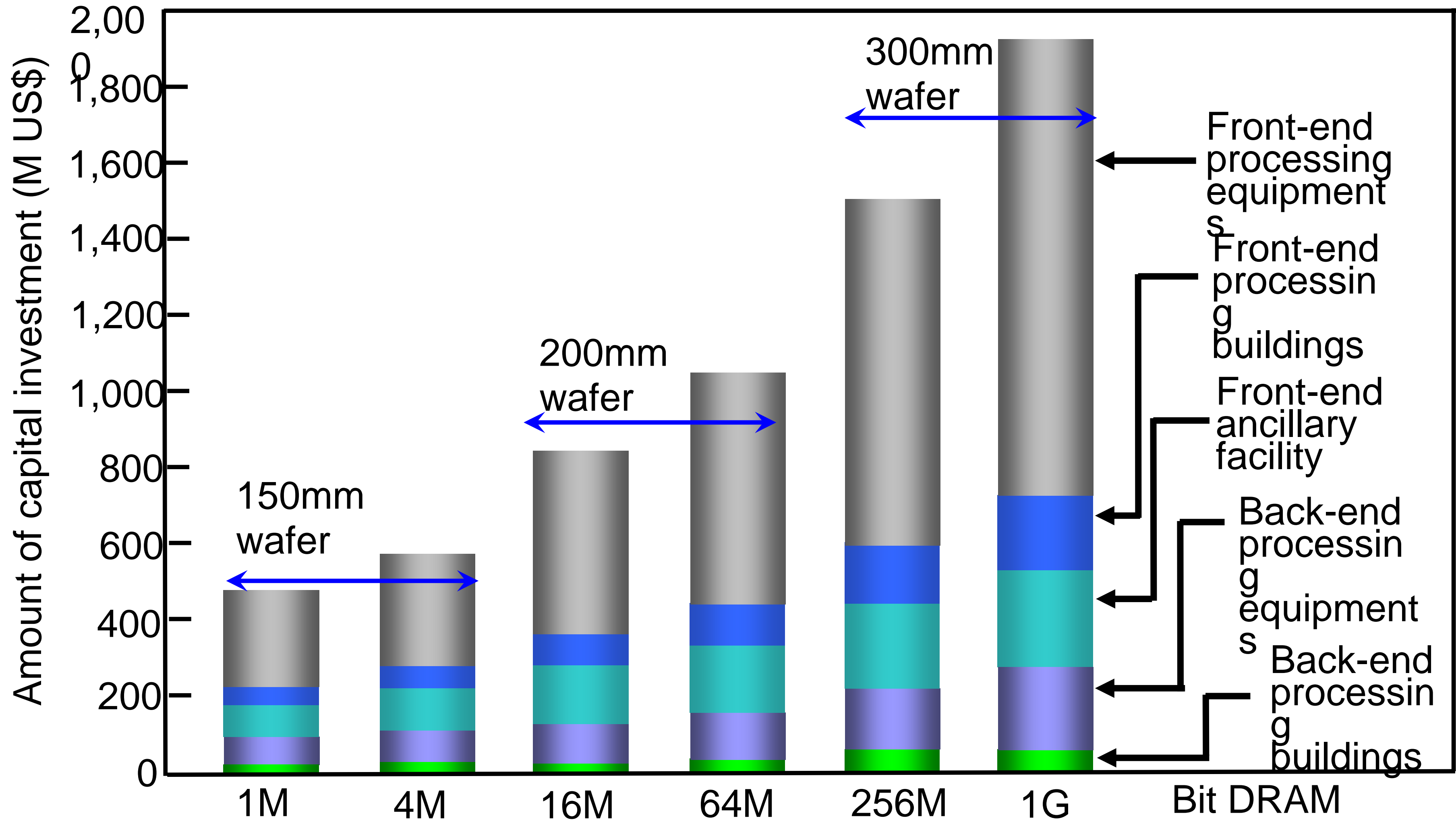
従来型メモリ高密度化の限界

フローティングゲート型NANDフラッシュメモリ

②セル間距離微細化の限界:フローティングゲート相互干渉



Amount of capital investment on Each of DRAM generation



Reference :Semiconductor Industry Research Institute Japan

Examples of foundry facility: UMC

Fab 12A (Tainan, Taiwan)
- US$3 billion investment
- Production since 2001
- 38K wafers/month by E/06
- 90, 65nm in production

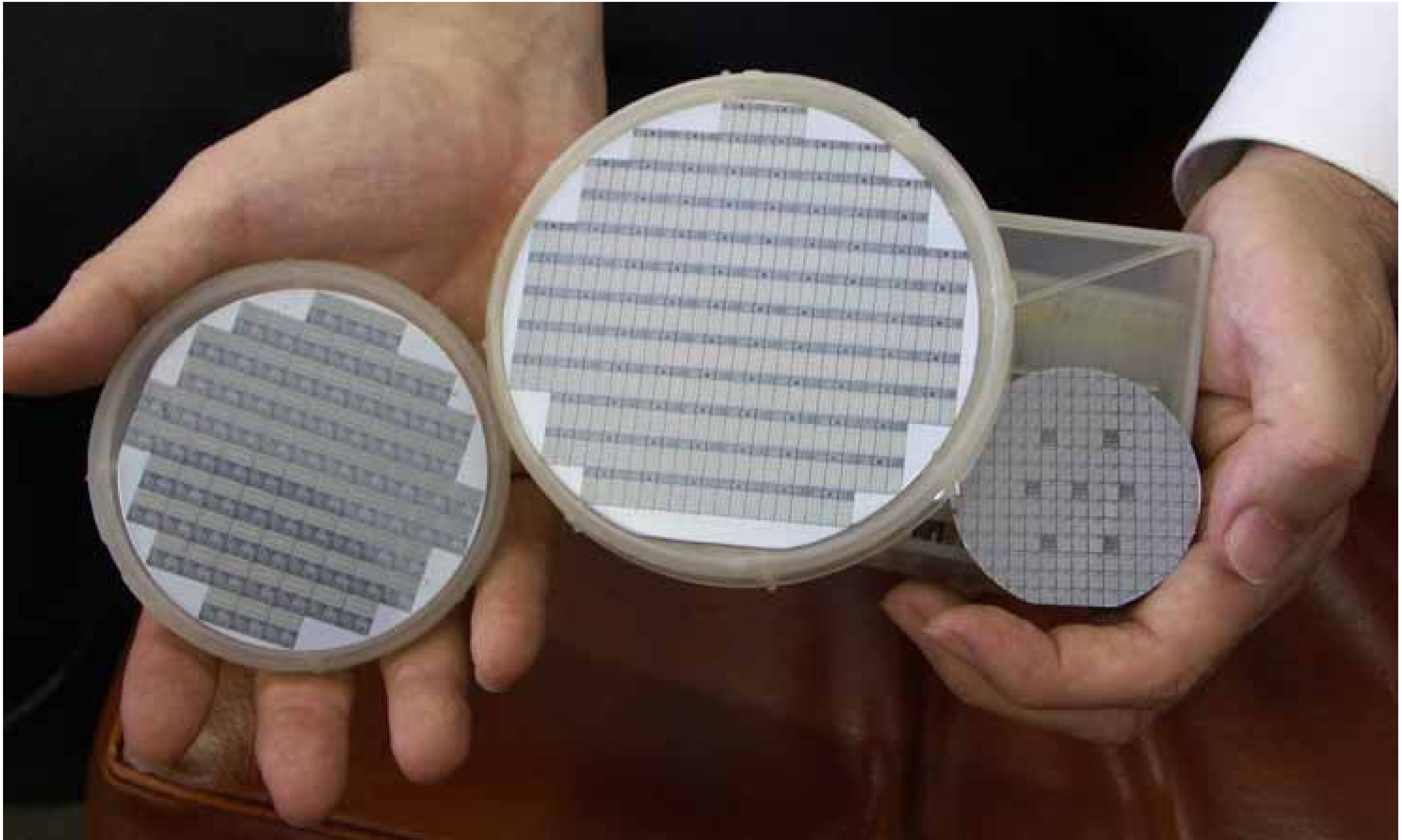
Fab 12i (Singapore)
- US$3.6 billion investment
- Production since 2004
- 25K wafers/month by E/06
- 130, 90nm in production
- Ready for 65nm pilot

UMC

Volume production with lager wafer is a solution

IC Fab Utility Usage : 12" vs. 8"

	12"-Fab	8"-Fab	12"/8" Usage ration	12"/8" Wafer size ratio
Power	1,100 KWH/Wafer	660 KWH/Wafer	1.7	2.25
Water	6.1 M3/Wafer	4.7 M3/Wafer	1.3	
Waste Water	3.8 M3/Wafer	2.9 M3/Wafer	1.3	
Waste Air	20,000 CMH/Wafer	13,000 CMH/Wafer	1.5	

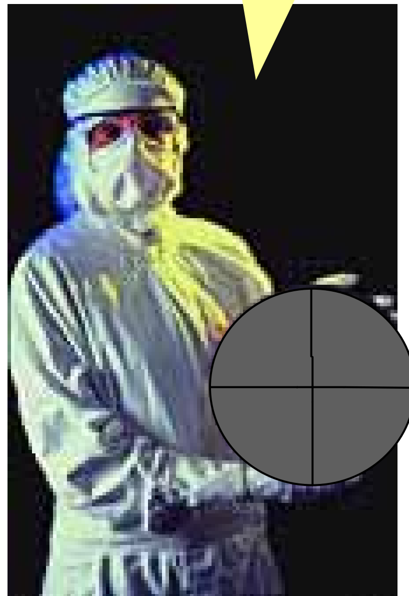


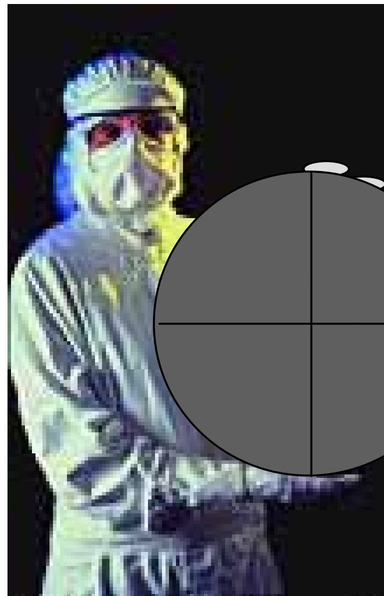
64k DRAM 3 inch wafer **1979**

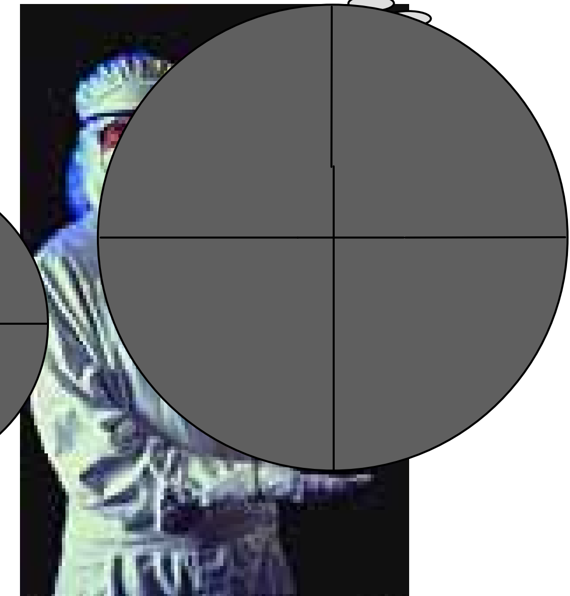
64k DRAM 4 inch wafer **1980**

1k SRAM 2 inch wafer **1974**

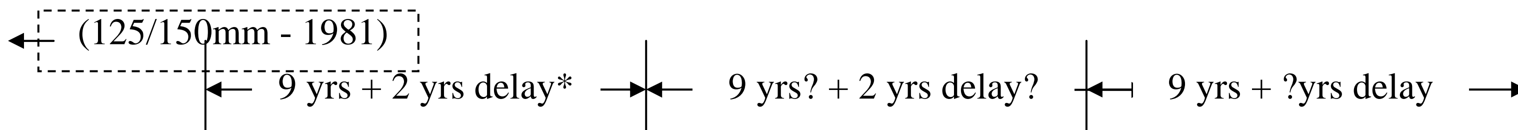
When do we start planning for next wafer size transition?







200mm/1990 300mm/2001 450mm/2012? 675mm/2021?



Fab in 1960s and 70s

Toshiba Corporation

300 mm Fab
TSMC

Toshiba Oita
Works

300 mm Super
clean room in
Tsukuba,Selete

300mm Global Fab Landscape

Europe
AMD Fab 36
Infineon SC300
Intel Fab 24
Crolles2 (TD)
(STM, Philips, Moto)
STM Catania

China
SMIC Fab 4

Korea
Samsung Line 12
Samsung Line 13
Samsung Line 14

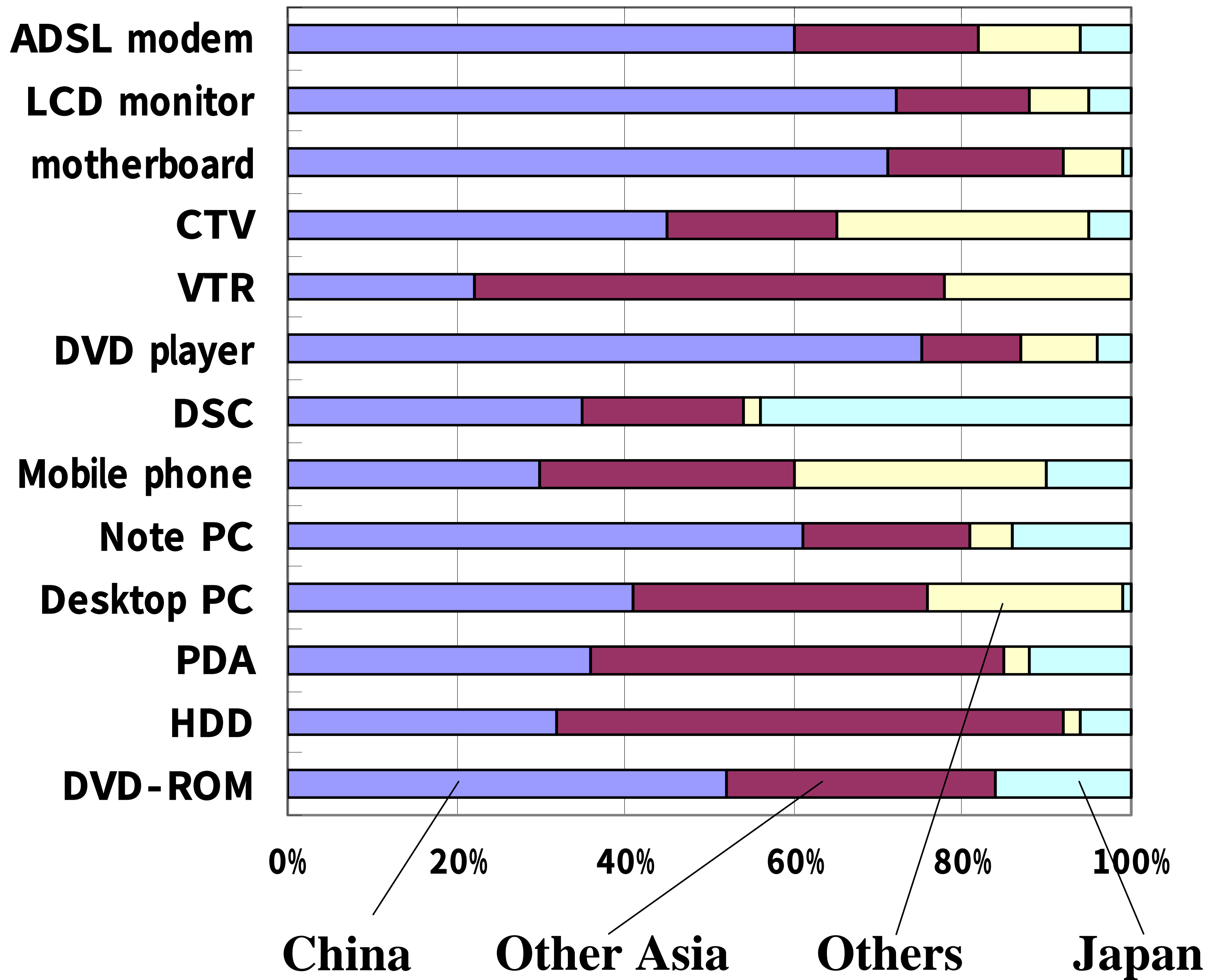
Japan
Elpida Fab 1
Elpida Fab 2
Renesas/ Trecenti
Toshiba Oita
Sony Kumamoto
Sony Fab 2 Nagasaki
Toshiba Yokkaichi City
NEC Elec Yamagata

United States
IBM Bldg 323
Infineon Richmond
Intel D1C
Intel D1D
Intel F11x
Intel F12C
Micron MTC
TI DMOS 6
TI DMOS 7

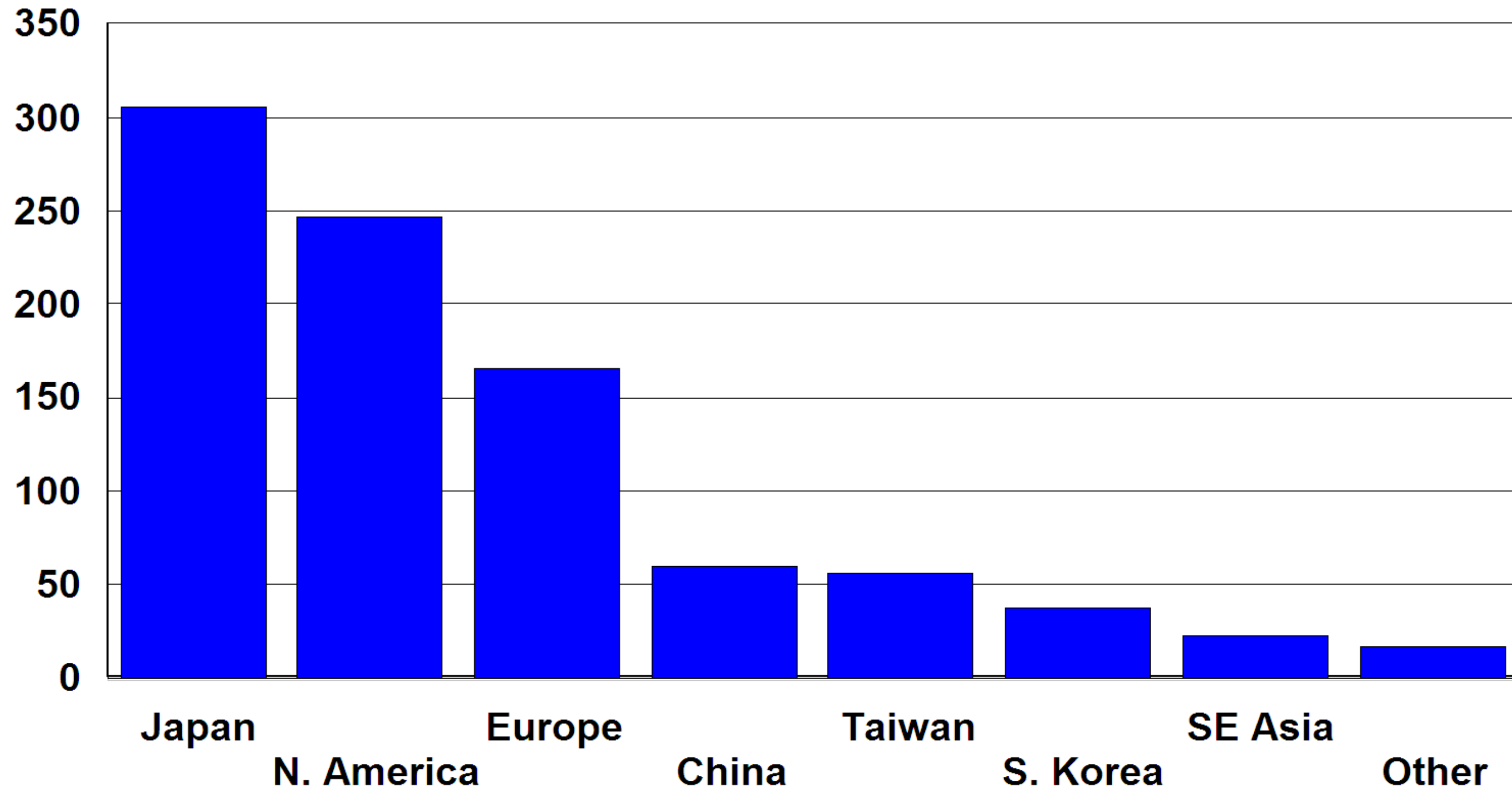
Singapore
UMC Singapore

Taiwan
Powerchip Fab 1
ProMOS Fab 2
TSMC Fab 12A & B
UMC Fab 12A
Inotera *(Nanya + Infineon)*
Powerchip Fab 1
Powerchip Fab 2
TSMC Fab 14

Sources: Silicon Strategies, EE Times, Nikkei Microelectronics, Industry Meetings

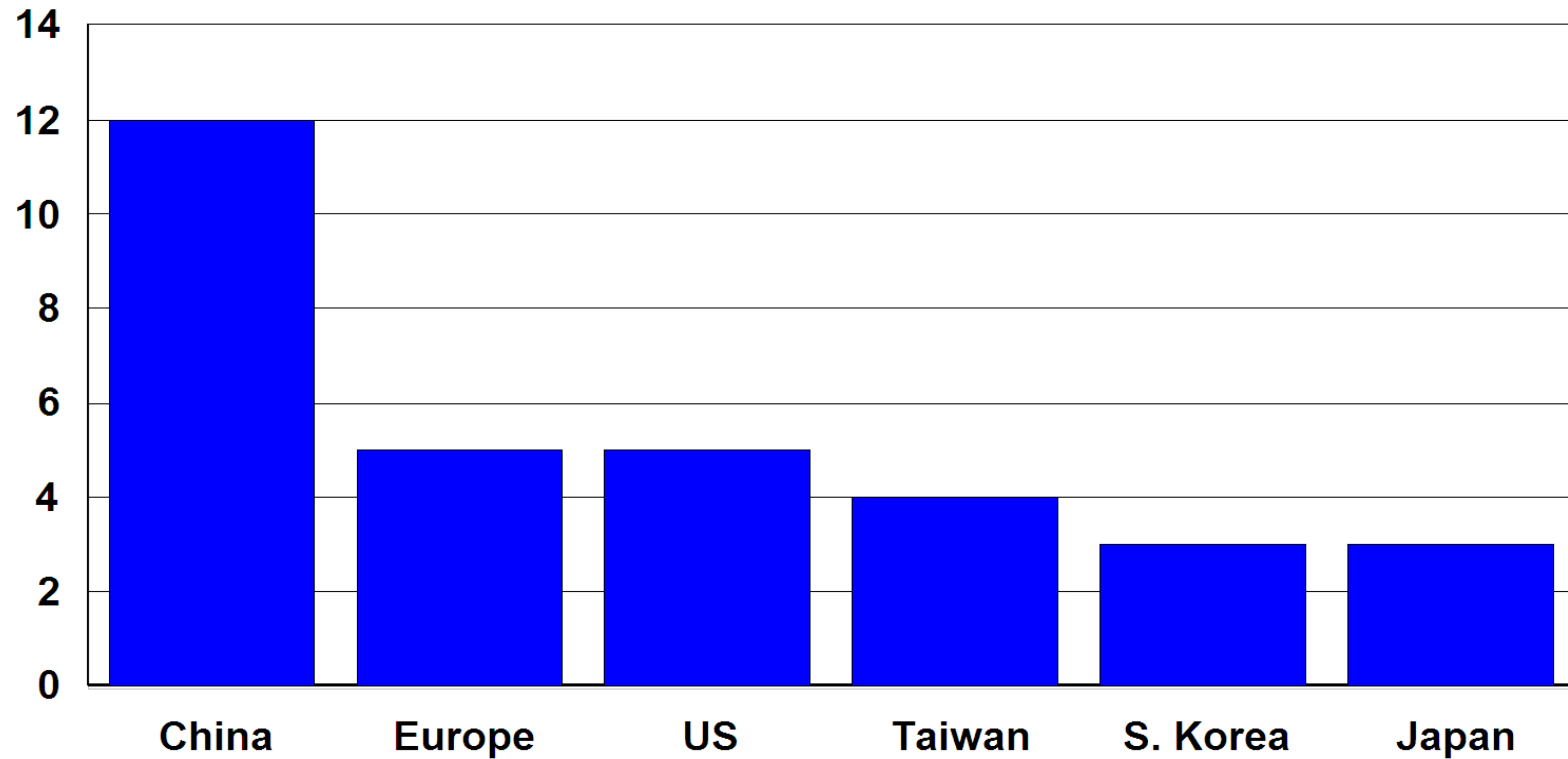

ADSL modem
LCD monitor
motherboard
CTV
VTR
DVD player
DSC
Mobile phone
Note PC
Desktop PC
PDA
HDD
DVD-ROM
0%
20%
40%
60%
80%
100%
China
Other Asia
Others
Japan

Number of Fabs in Operation by Region



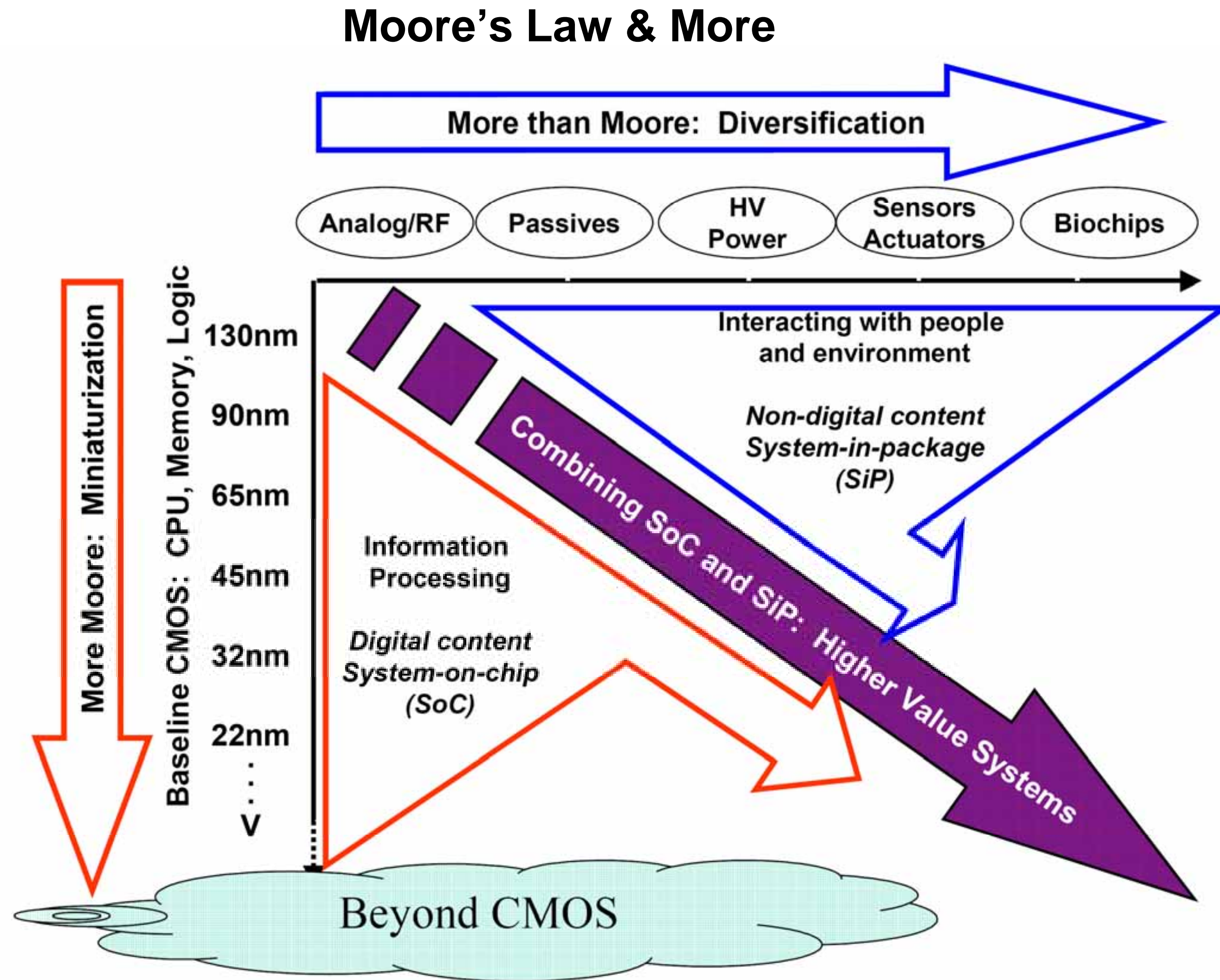
World Fab Watch (April 2004)

Number of Fabs Currently Under Construction or Equipping



What will be the roadmap after 2020?

More Moore and More than Moore



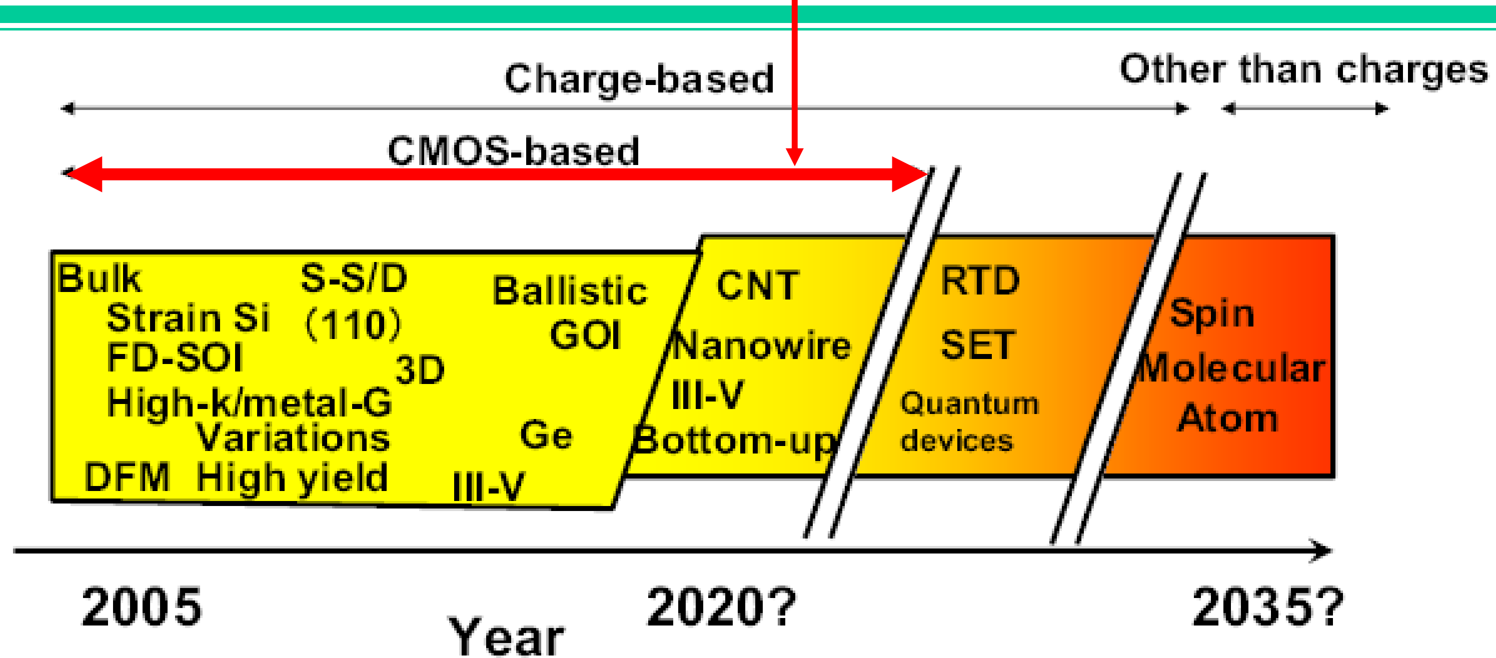
Question what is the other side of the cloud?

ITRS 2005 Edition

http://strj-jeita.elisasp.net/pdf_ws_2005nendo/9A_WS2005IRC_Ishiuchi.pdf

Question: Will CMOS end in 2020?

Three Stages in Silicon Nanoelectronics



1. CMOS Extension

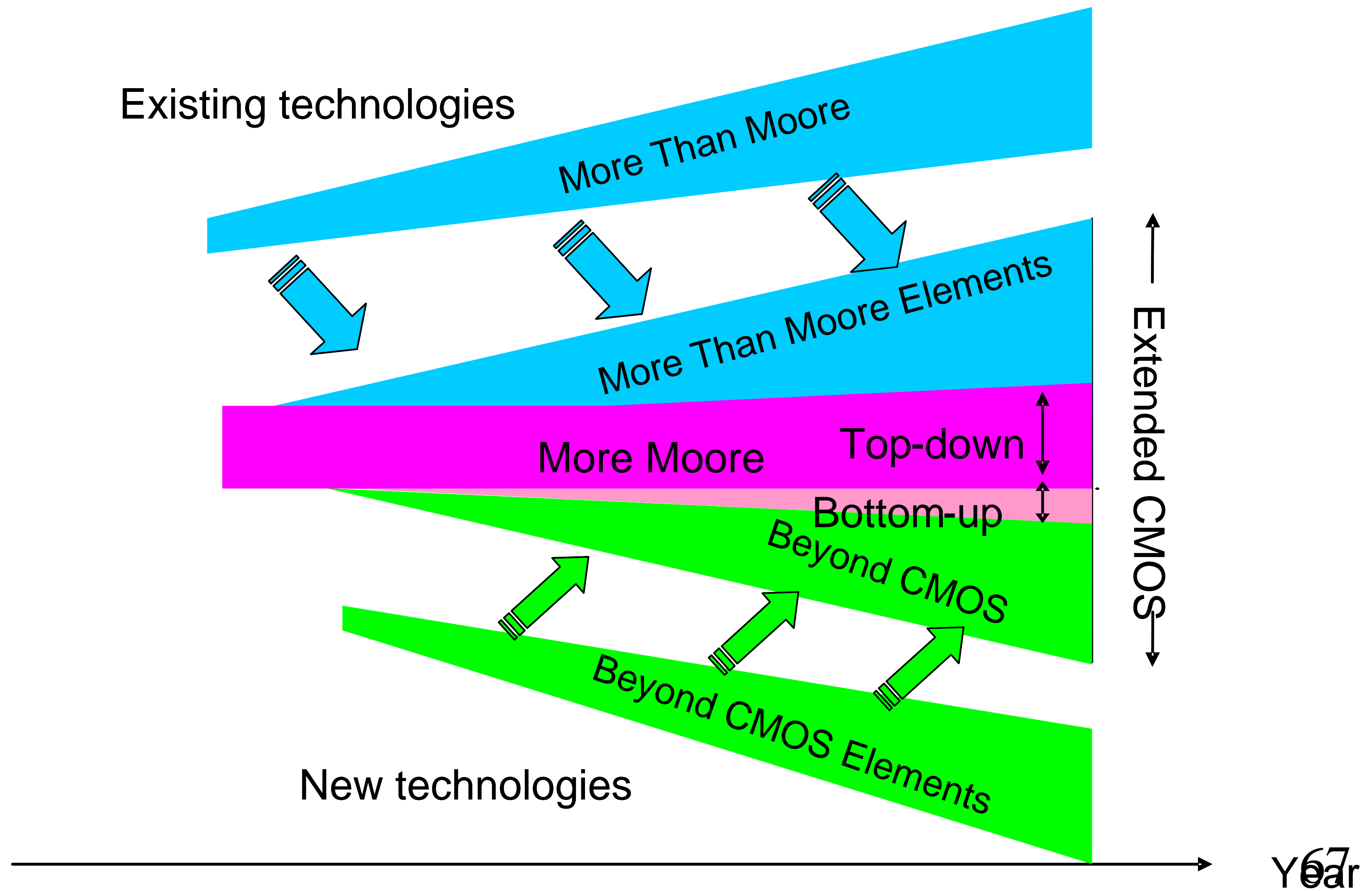
3. Beyond CMOS

2. New Functions Added to CMOS

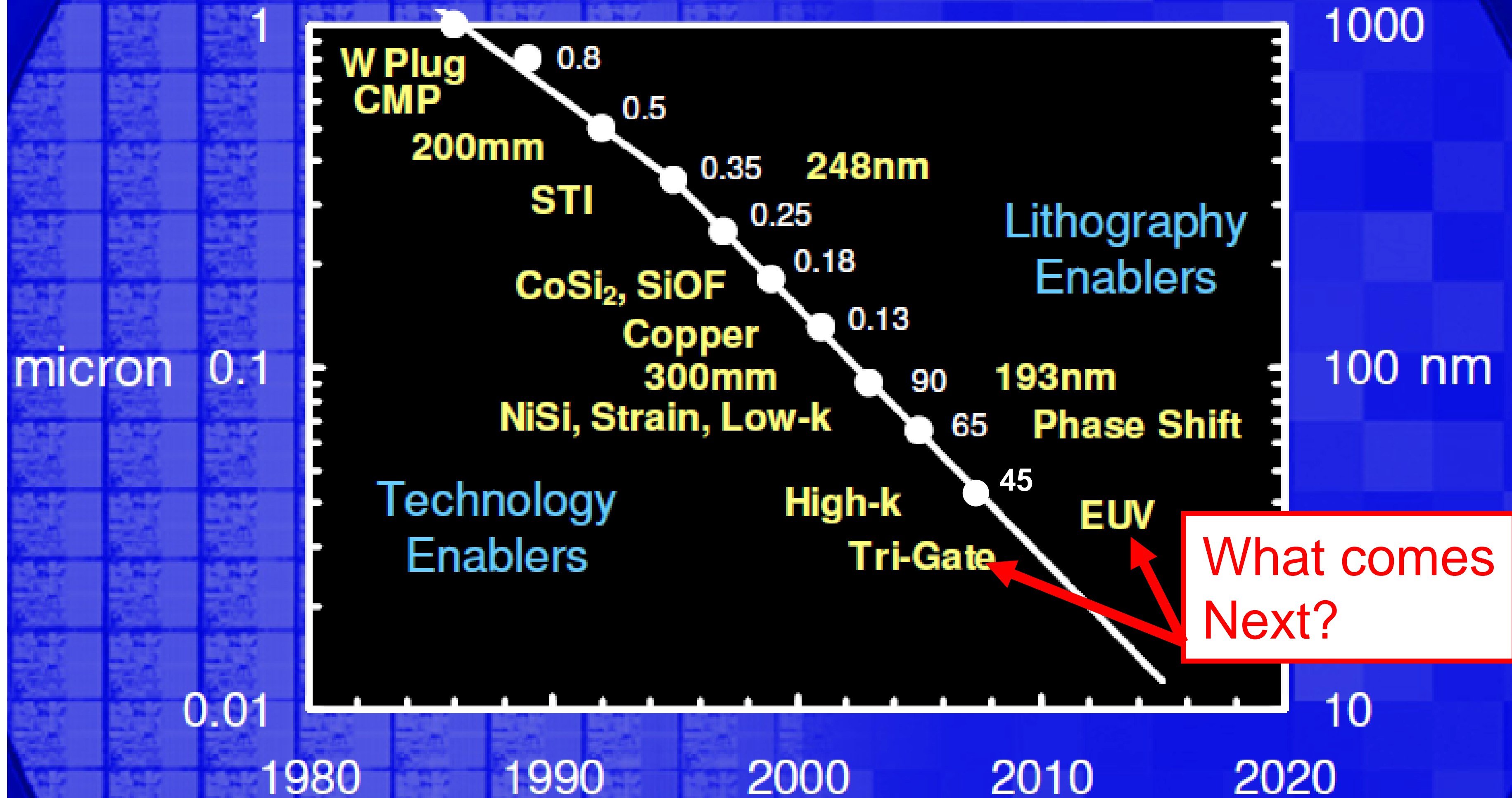
http://www.rcns.hiroshima-u.ac.jp/21coe/pdf/5th_WS/2-4-2_Hiramoto.pdf

New Concept for Roadmap

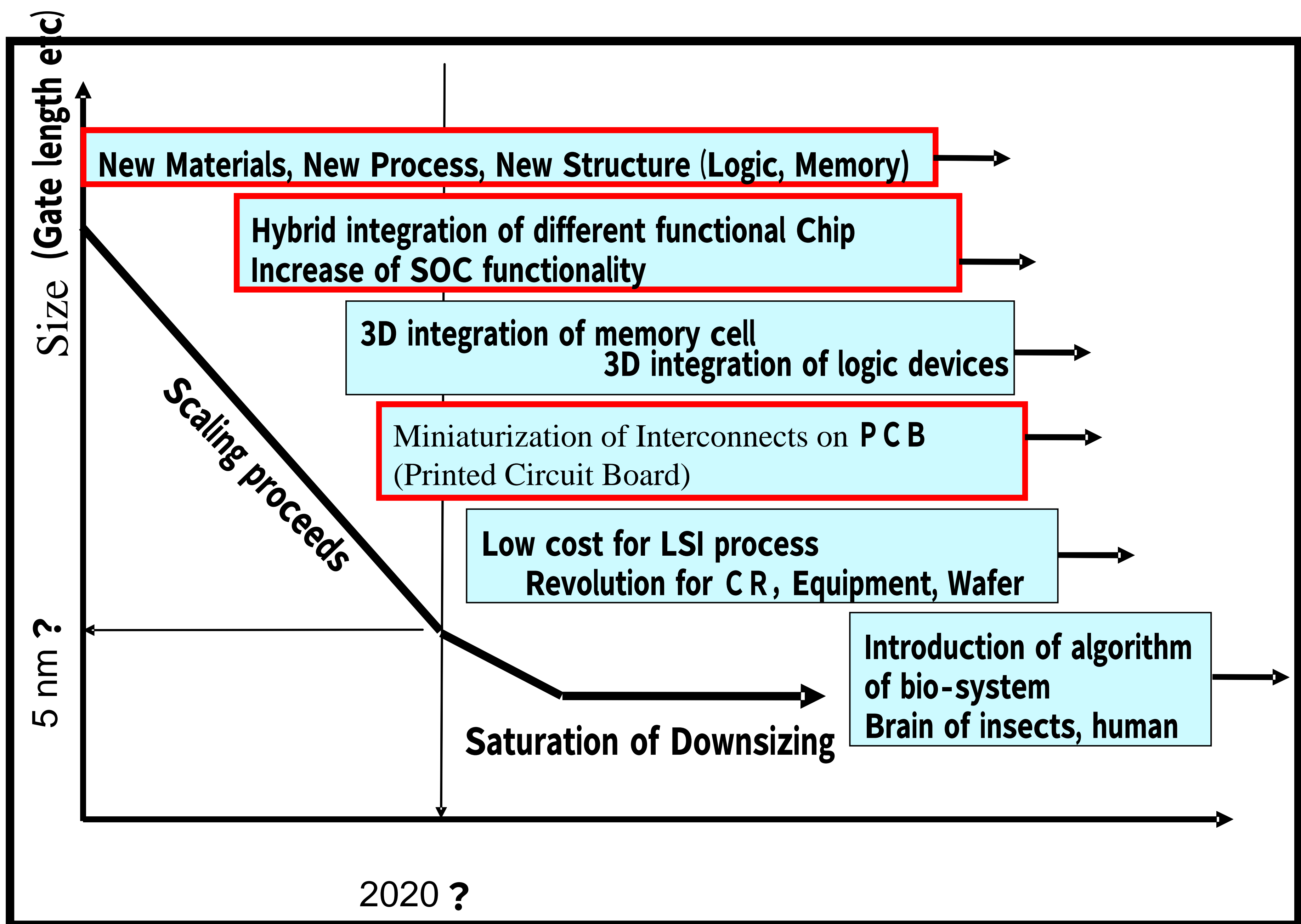
Evolution of Extended CMOS Continues!!



Scaling Gets Tougher at Smaller Dimensions



Intel continues to develop and implement new materials and structures to meet the challenge


Size (Gate length etc)
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
Scaling proceeds
Miniaturization of Interconnects on PCB
(Printed Circuit Board)
Low cost for LSI process
Revolution for CR, Equipment, Wafer
5 nm ?
Introduction of algorithm
of bio-system
Brain of insects, human
Saturation of Downsizing
2020 ?

After 2020

There is no decrease in gate length around at 5 nm due to subthreshold leakage increase.

It is not useful to reduce the gate length any more for increasing the drain current, because the conduction is already semi-ballistic.

Ballistic conduction is that with no scattering of the carriers.

In the ballistic conduction, there is no increase of drain current with decreasing the channel length.

What is important for keeping Moore Law, then?

→ Increase drain current drive under low drain voltage in order to reduce the power consumption.

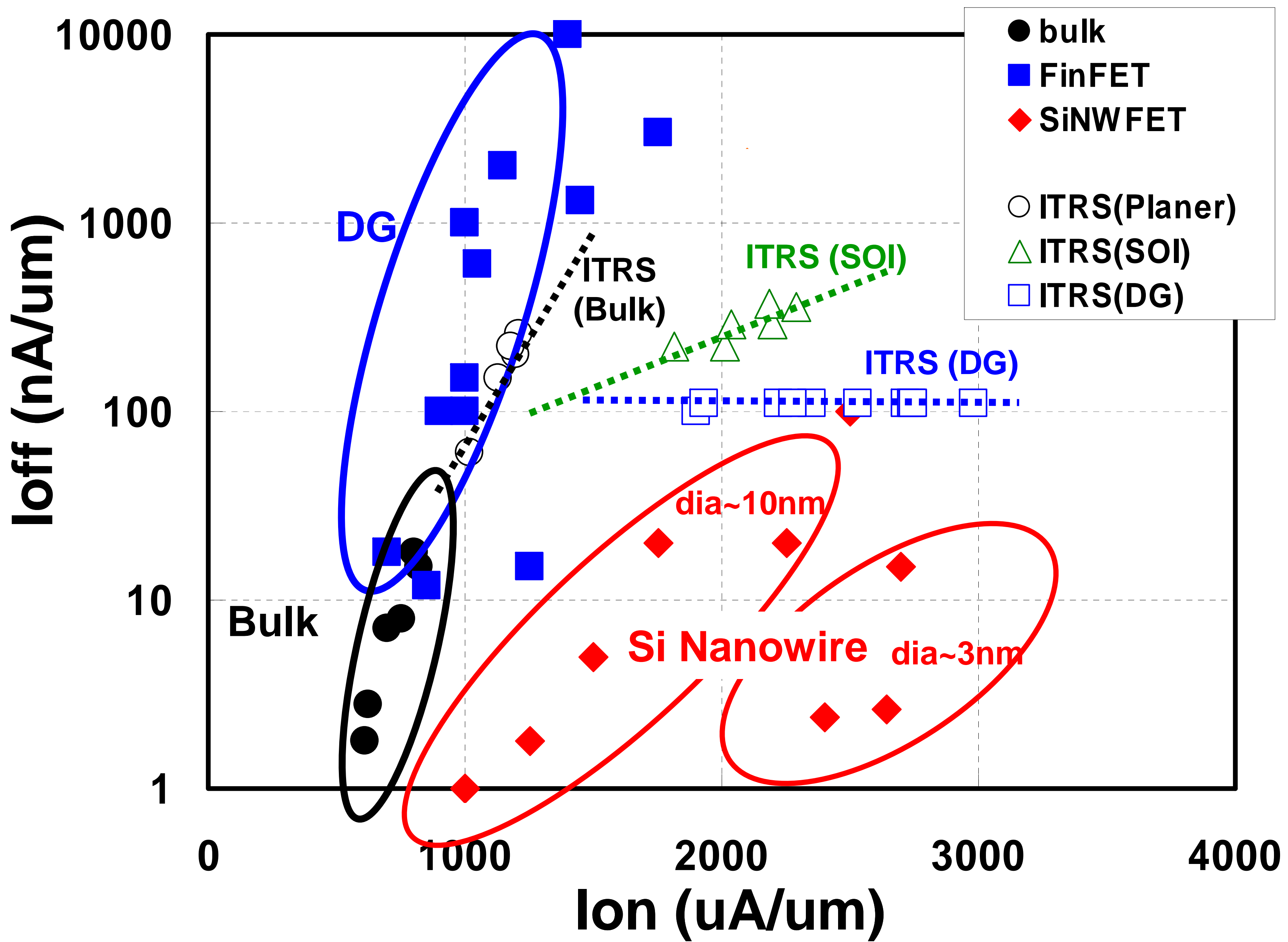
Selection of MOSFET structure

For suppression of Ioff, the structure will be Fin-FET type

High-conduction at low voltage

1. 1D conduction → Nano-wire, Nano-tube FET
2. Increase number of quantum channel →
   Band engineering, CNT(Carbon Nano-tube Tnansistor)
3. Increase the number of wire or tube →
   3D integration of wires and tubes

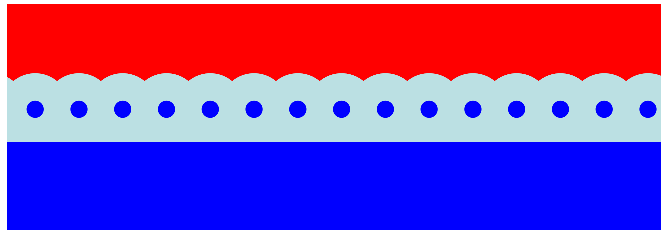
High-integration, low cost production,
no-small-geometry lithography → CNT

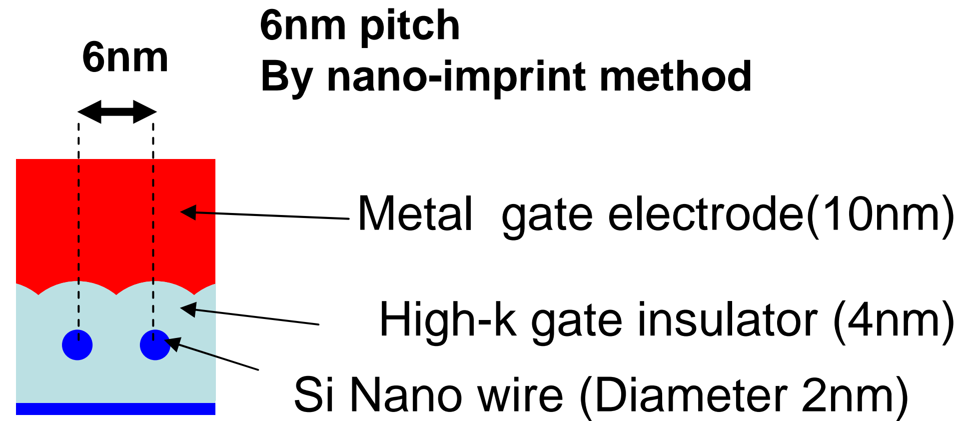


bulk
FinFET
SiNWFET
ITRS(Planer)
ITRS(SOI)
ITRS(DG)
DG
ITRS
(Bulk)
ITRS (SOI)
ITRS (DG)
Bulk
dia~10nm
Si Nanowire
dia~3nm
Ioff (nA/um)
Ion (uA/um)
10000
1000
100
10
1
0
1000
2000
3000
4000

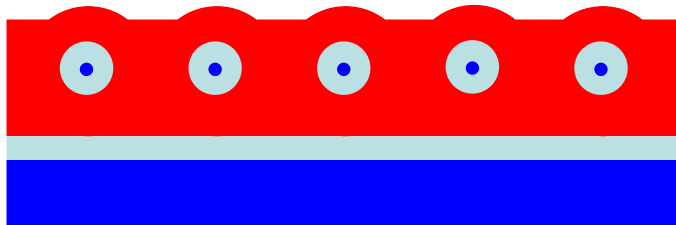
Maximum number of wires per 1 μm

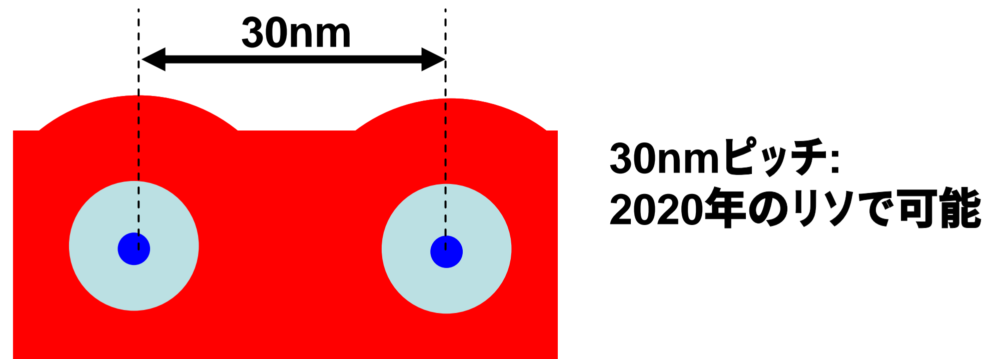
Front gate type MOS 165 wires /μm

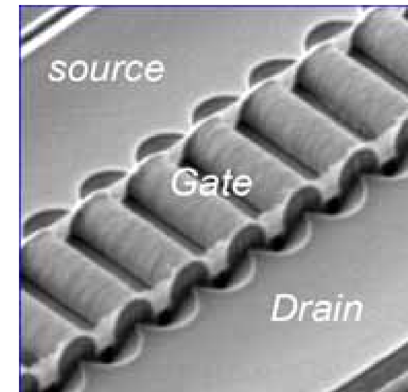




Surrounded gate type MOS 33 wires/μm

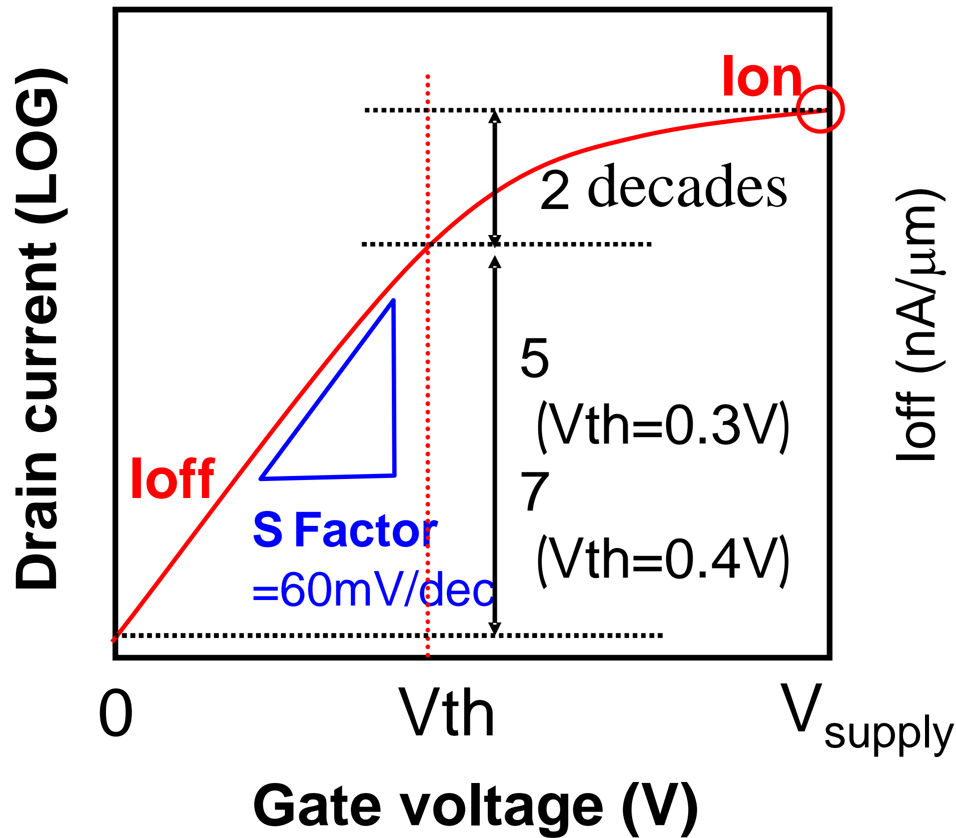


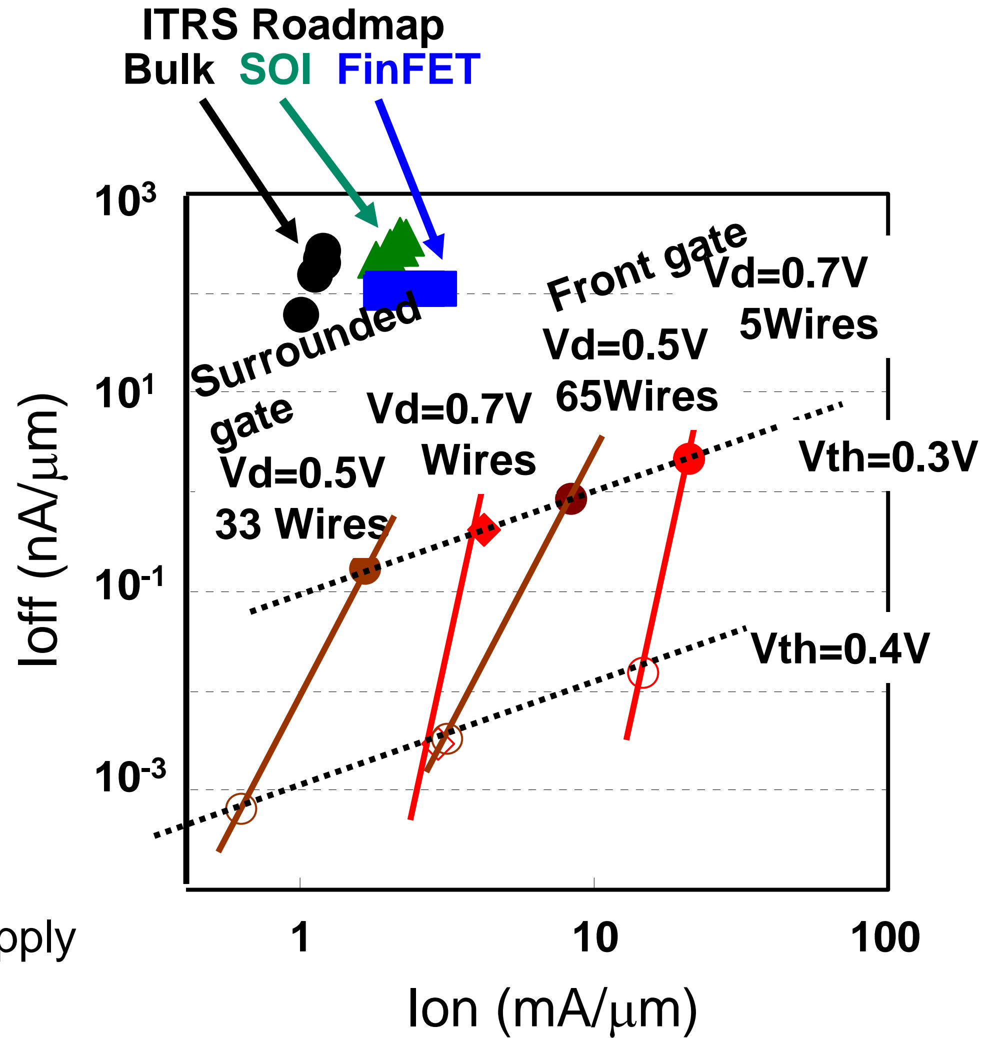




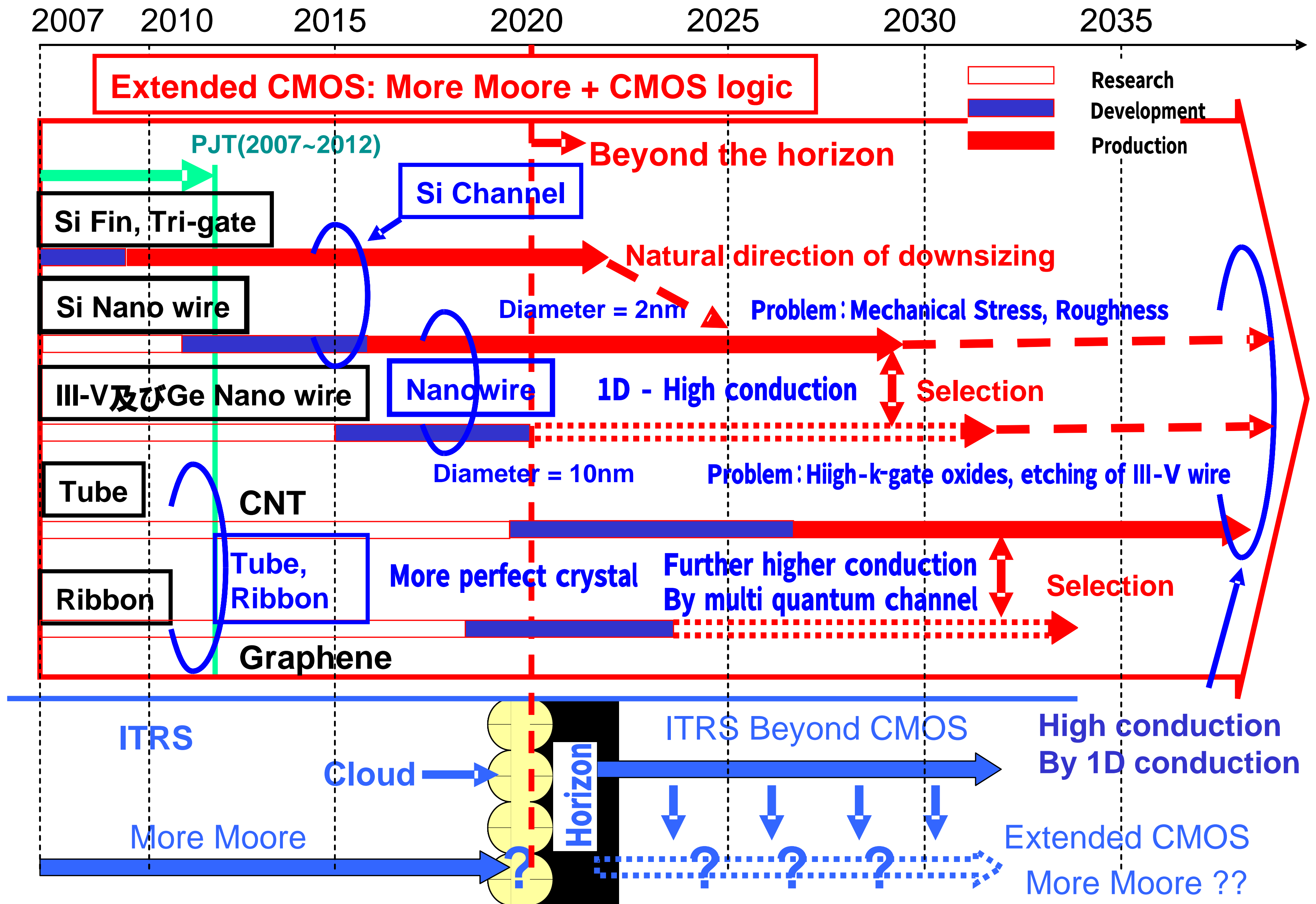
Surrounded gate MOS

Estimation of Ion and Ioff of 1D conduction MOFETs For optimum prediction





Our new roadmap

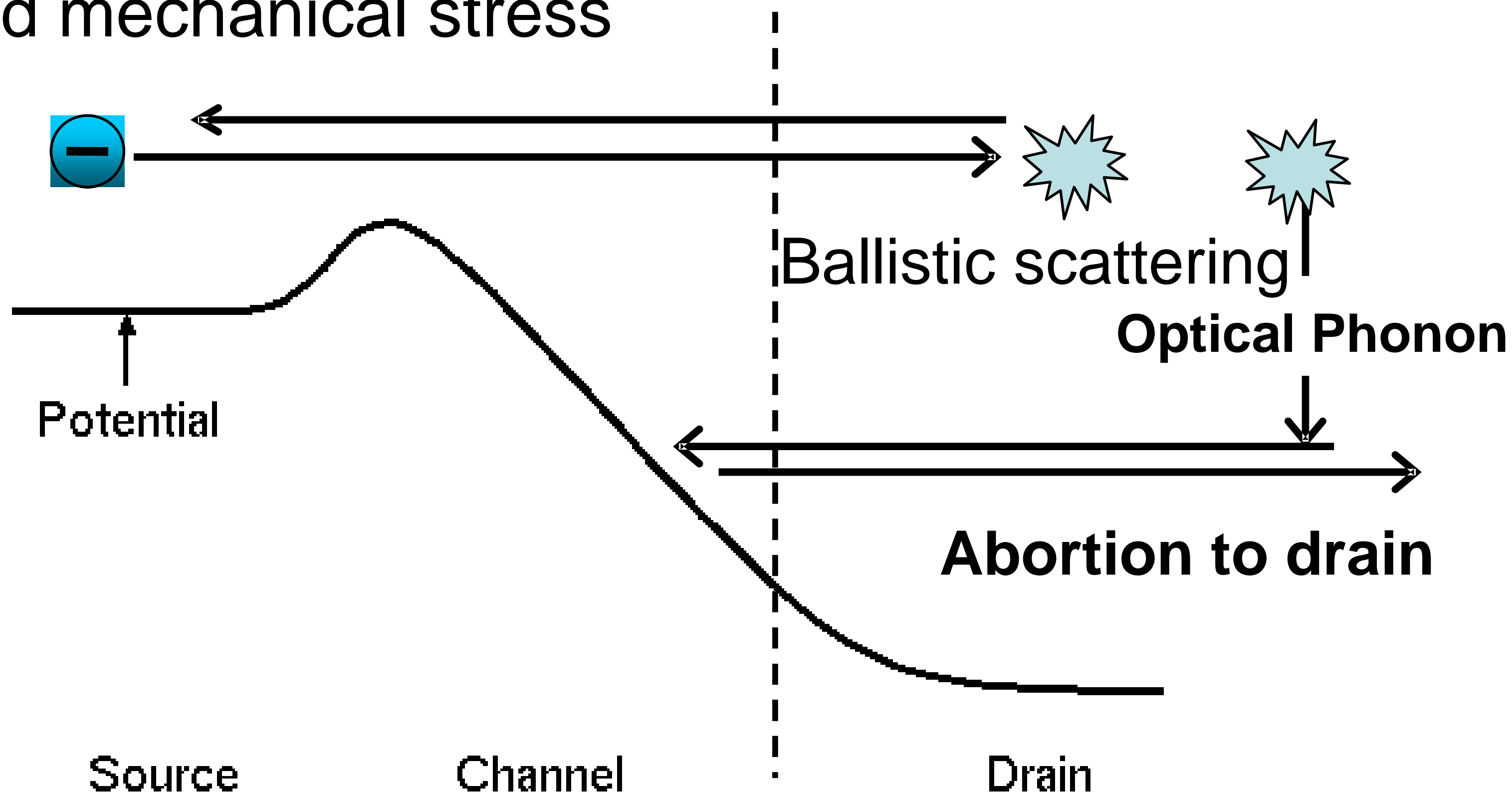


1D conduction per one quantum channel:

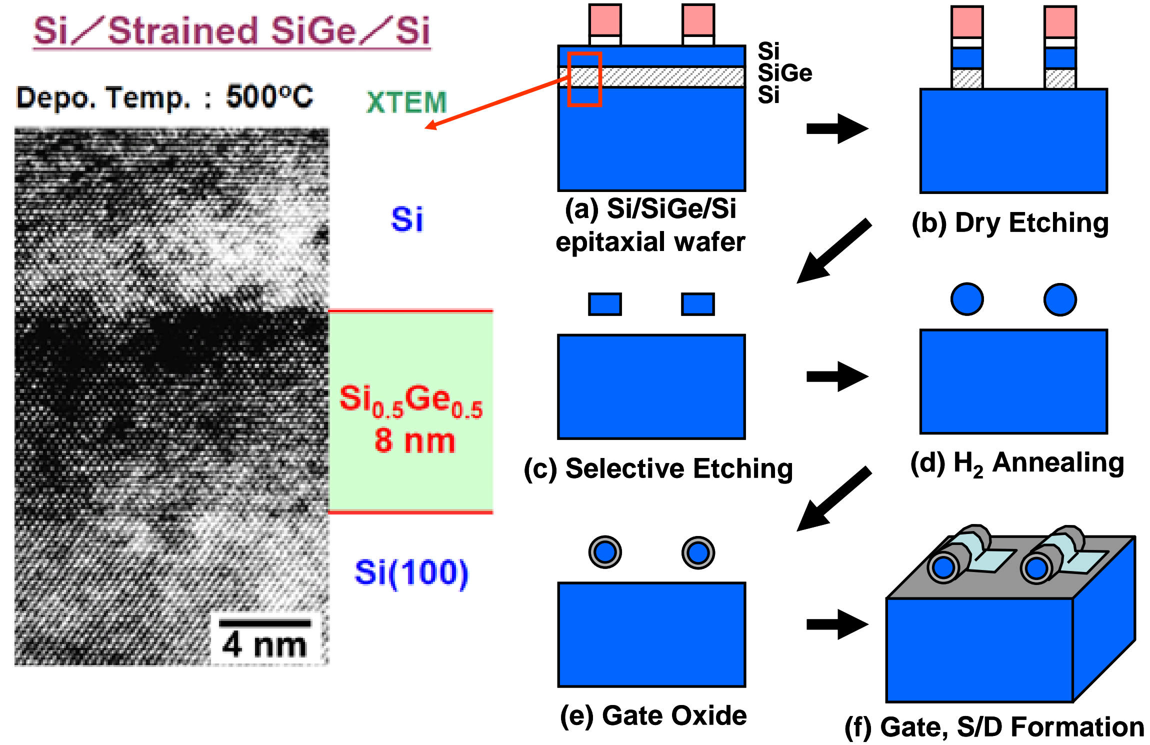
$G = 2e^2/h = 77.5\ \mu S$/wire or tube

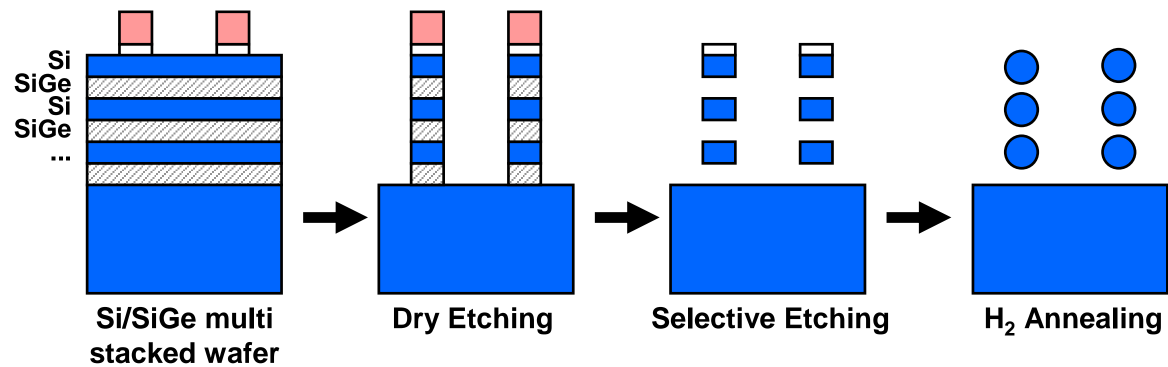
regardless of gate length and channel material

But this value has not been obtained yet, due to reflection of carriers from drain, surface roughness and mechanical stress

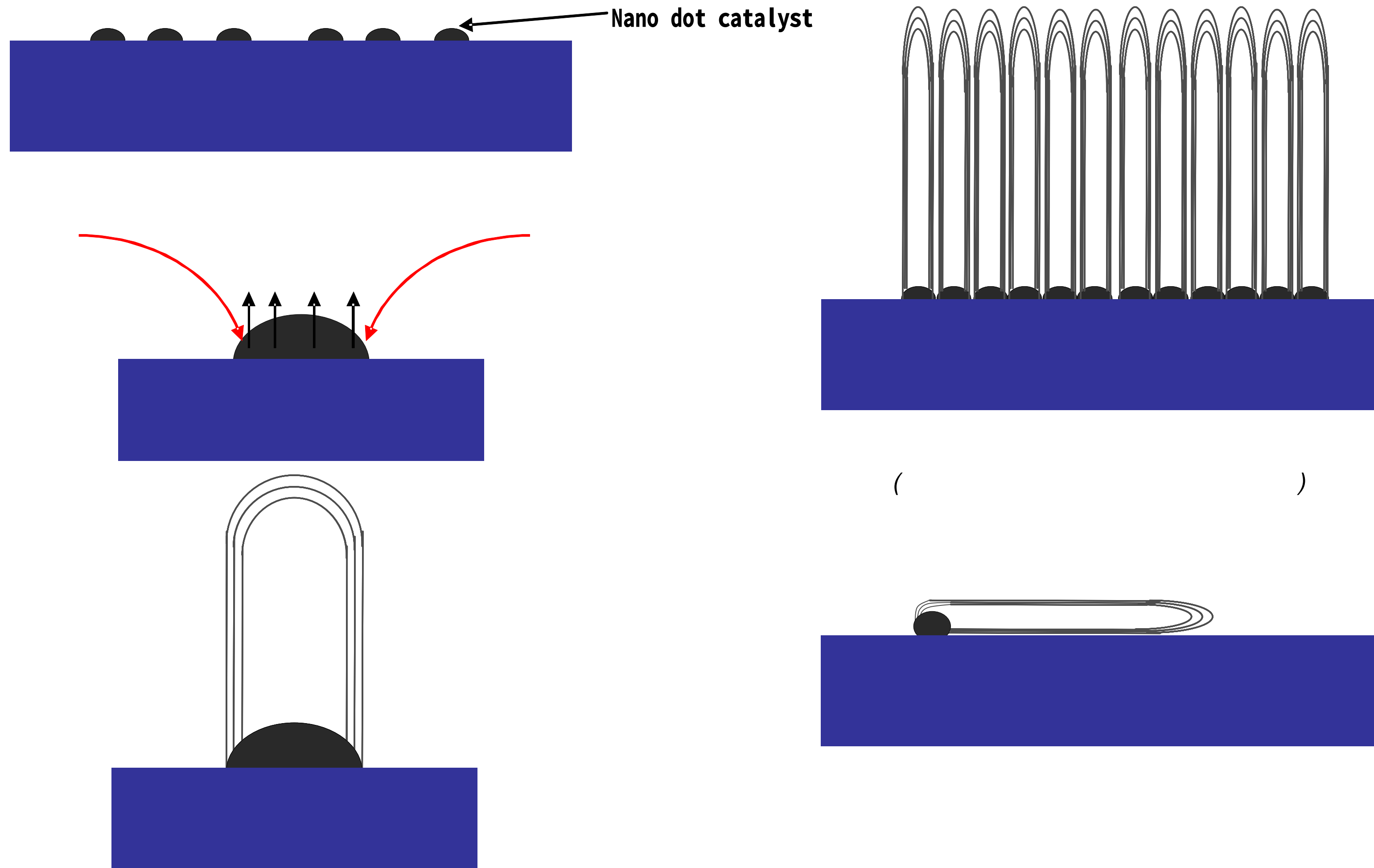


Increase the number of wires





Lateral growth of CNT



Thank you
for your attention!