Roadmap for 22nm Logic CMOS and Beyond

March 9, 2009

@Heritage Institute of Technology

Hiroshi Iwai,
Tokyo Institute of Technology
• There were many inventions in the 20\textsuperscript{th} century: Airplane, Nuclear Power generation, Computer, Space aircraft, etc
• However, everything has to be controlled by electronics
• Electronics
  Most important invention in the 20\textsuperscript{th} century
• What is Electronics: To use electrons, Electronic Circuits
Electronic Circuits started by the invention of vacuum tube (Triode) in 1906

Thermal electrons from cathode controlled by grid bias

Same mechanism as that of transistor
4 wives of Lee De Forest

1906 Lucille Sheardown
1907 Nora Blatch
1912 Mary Mayo, singer
1930 Marie Mosquini, silent film actress
First Computer Eniac: made of huge number of vacuum tubes 1946
Big size, huge power, short life time filament

→ dreamed of replacing vacuum tube with solid-state device

Today's pocket PC made of semiconductor has much higher performance with extremely low power consumption
History of Semiconductor devices

1947, 1st Point Contact Bipolar Transistor:
   Ge Semiconductor, Bardeen, Brattin
   → Nobel Prize

1948, 1st Junction Bipolar Transistor,
   Ge Semiconductor, Schokley
   → Nobel Prize

1958, 1st Integrated Circuits,
   Ge Semiconductor, J.Kilby → Nobel Prize

1959, 1st Planar Integrated Circuits,
   R.Noice

1960, 1st MOS Transistor, Kahng,
   Si Semiconductor

1963, 1st CMOS Circuits, C.T. Sah and F. Wanlass
J. E. LILIENFELD

DEVICES FOR CONTROLLED ELECTRIC CURRENT

Filed March 28, 1928
Electron Semiconductor Gate Electrode Gate Insulator Negative bias Capacitor structure with notch No current Electric field}

Positive bias Current flows
Surface Potential (Negative direction)

0 bias for gate

Positive bias for gate

Electron flow
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!
However, they found amplification phenomenon when investigating Ge surface when putting needles. This is the 1st Transistor: 

**Not Field Effect Transistor,**
**But Bipolar Transistor (another mechanism)**

1947: 1st transistor

J. Bardeen  
W. Bratten,  
W. Shockley
1958: 1st Integrated Circuit

Connect 2 bipolar transistors in the Same substrate by bonding wire.

Jack S. Kilby
1960: First MOSFET by D. Kahng and M. Atalla

Top View

Si
Source

Al Gate

Drain

Si

Si/SiO2 Interface is extraordinarily good
1970,71: 1st generation of LSIs

DRAM    Intel 1103

MPU     Intel 4004
MOS LSI experienced continuous progress for many years

<table>
<thead>
<tr>
<th>Name of Integrated Circuits</th>
<th>Number of Transistors</th>
</tr>
</thead>
<tbody>
<tr>
<td>1960s IC (Integrated Circuits)</td>
<td>~10</td>
</tr>
<tr>
<td>1970s LSI (Large Scale Integrated Circuit)</td>
<td>~1,000</td>
</tr>
<tr>
<td>1980s VLSI (Very Large Scale IC)</td>
<td>~10,000</td>
</tr>
<tr>
<td>1990s ULSI (Ultra Large Scale IC)</td>
<td>~1,000,000</td>
</tr>
<tr>
<td>2000s ?LSI (? Large Scale IC)</td>
<td>~1,000,000,000</td>
</tr>
</tbody>
</table>
MOSFET: Metal Oxide Semiconductor Field Effect Transistor

Use Gate Field Effect for switching
N-MOS (N-type MOSFET)  
Source  
Gate  
Drain  
Electron flow  
Current flow  

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

OFF

Gate

0 V

Source

0 V

Drain

1 V

Electrons

n-MOSFET

ON

Gate

1 V

Source

1 V

Drain

1 V

Electron flow

High Potential Region
n-MOSFET

Drain Current

Positive

Gate bias

Positive voltage

Threshold voltage

OFF

OFF

ON

p-MOSFET

Drain Current

Gate bias

Negative voltage

Threshold voltage

OFF

ON
CMOS

Complimentary MOS

When NMOS is ON, PMOS is OFF
When PMOS is ON, NMOS is OFF
CMOS: Low Power: No DC current from Power supply to the ground

\[ P = \frac{1}{2} CV_D^2 \]

1 cycle

Clock frequency \( f \)

\[ P = \frac{1}{2} fCV_D^2 \]
2 input NAND Circuit

NAND = NOT \cdot AND

<table>
<thead>
<tr>
<th>Input 1</th>
<th>Input 2</th>
<th>Output</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 1 0 0</td>
<td>1 0 1 0</td>
<td>1 0 0 0</td>
</tr>
</tbody>
</table>

AND

<table>
<thead>
<tr>
<th>Input 1</th>
<th>Input 2</th>
<th>Output</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 1 0 0</td>
<td>1 0 1 0</td>
<td>0 1 1 1</td>
</tr>
</tbody>
</table>
Needless to say, but....

**CMOS Technology:**
Indispensable for our human society

All the human activities are controlled by CMOS

living, production, financing, telecommunication, transportation, medical care, education, entertainment, etc.

**Without CMOS:**
There is no computer in banks, and

world economical activities immediately stop.

Cellarer phone dose not exists
Downsizing of the components has been the driving force for circuit evolution

<table>
<thead>
<tr>
<th>Year</th>
<th>Device</th>
<th>Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>1900</td>
<td>Vacuum Tube</td>
<td>10 cm</td>
</tr>
<tr>
<td>1950</td>
<td>Transistor</td>
<td>cm</td>
</tr>
<tr>
<td>1960</td>
<td>IC</td>
<td>mm</td>
</tr>
<tr>
<td>1970</td>
<td>LSI</td>
<td>10 µm</td>
</tr>
<tr>
<td>2000</td>
<td>ULSI</td>
<td>100 nm</td>
</tr>
</tbody>
</table>

- $10^{-1}$m
- $10^{-2}$m
- $10^{-3}$m
- $10^{-5}$m
- $10^{-7}$m

In 100 years, the size reduced by one million times. There have been many devices from stone age. We have never experienced such a tremendous reduction of devices in human history.
Downsizing

1. **Reduce Capacitance**
   - Reduce switching time of MOSFETs
   - Increase clock frequency
     - Increase circuit operation speed
2. **Increase number of Transistors**
   - Parallel processing
     - Increase circuit operation speed

Downsizing contribute to the performance increase in double ways

Thus, downsizing of Si devices is the most important and critical issue.
Scaling Method: by R. Dennard in 1974

**Wdep**: Space Charge Region (or Depletion Region) Width

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

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

K = 0.7 for example

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

- Good scaled I-V characteristics

\[ \text{Wdep} \propto \sqrt{V/Na} \]

\[ \text{I} : \text{K} \]
<table>
<thead>
<tr>
<th>Geometry &amp; Supply voltage</th>
<th>Lg, Wg, Tox, Vdd</th>
<th>K</th>
<th>Scaling K: K=0.7 for example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Drive current in saturation</td>
<td>I_d</td>
<td>K</td>
<td>I_d = v_sat Wg C_o (V_g - V_th) C_o: gate C per unit area</td>
</tr>
<tr>
<td></td>
<td>I_d/µm</td>
<td>1</td>
<td>I_d per unit W_g = I_d / W_g = 1</td>
</tr>
<tr>
<td>Gate capacitance</td>
<td>C_g</td>
<td>K</td>
<td>C_g = ε_o ε_ox L_g W_g / t_ox —— KK/K = K</td>
</tr>
<tr>
<td>Switching speed</td>
<td>τ</td>
<td>K</td>
<td>τ = C_g V_dd / I_d —— KK/K = K</td>
</tr>
<tr>
<td>Clock frequency</td>
<td>f</td>
<td>1/K</td>
<td>f = 1/τ = 1/K</td>
</tr>
<tr>
<td>Chip area</td>
<td>A_{chip}</td>
<td>α</td>
<td>α: Scaling factor —— In the past, α&gt;1 for most cases</td>
</tr>
<tr>
<td>Integration (# of Tr)</td>
<td>N</td>
<td>α/K^2</td>
<td>N —— α/K^2 = 1/K^2, when α=1</td>
</tr>
<tr>
<td>Power per chip</td>
<td>P</td>
<td>α</td>
<td>fNCV^2/2 —— K^{-1}(αK^{-2})K(K^1)^2 = α = 1, when α=1</td>
</tr>
</tbody>
</table>

Downscaling merit: Beautiful!
<table>
<thead>
<tr>
<th></th>
<th>$k = 0.7$ and $\alpha = 1$</th>
<th>$k = 0.7^2 = 0.5$ and $\alpha = 1$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Single MOFET</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$V_{dd}$</td>
<td>$0.7$</td>
<td>$0.5$</td>
</tr>
<tr>
<td>$L_g$</td>
<td>$0.7$</td>
<td>$0.5$</td>
</tr>
<tr>
<td>$I_d$</td>
<td>$0.7$</td>
<td>$0.5$</td>
</tr>
<tr>
<td>$C_g$</td>
<td>$0.7$</td>
<td>$0.5$</td>
</tr>
<tr>
<td>$P \text{ (Power)}/\text{Clock}$</td>
<td>$0.7^3 = 0.34$</td>
<td>$0.5^3 = 0.125$</td>
</tr>
<tr>
<td>$\tau \text{ (Switching time)}$</td>
<td>$0.7$</td>
<td>$0.5$</td>
</tr>
<tr>
<td><strong>Chip</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$N \text{ (# of Tr)}$</td>
<td>$1/0.7^2 = 2$</td>
<td>$1/0.5^2 = 4$</td>
</tr>
<tr>
<td>$f \text{ (Clock)}$</td>
<td>$1/0.7 = 1.4$</td>
<td>$1/0.5 = 2$</td>
</tr>
<tr>
<td>$P \text{ (Power)}$</td>
<td>$1$</td>
<td>$1$</td>
</tr>
</tbody>
</table>
Actual past downscaling trend until year 2000

Change in 30 years

<table>
<thead>
<tr>
<th></th>
<th>Ideal scaling</th>
<th>Real Change</th>
<th>Ideal scaling</th>
<th>Real Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>$L_g$</td>
<td>$K$</td>
<td>$10^{-2}$</td>
<td>$I_d$</td>
<td>$K(10^{-2})$</td>
</tr>
<tr>
<td>$t_{ox}$</td>
<td>$K(10^{-2})$</td>
<td>$10^{-2}$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$V_{dd}$</td>
<td>$K(10^{-2})$</td>
<td>$10^{-1}$</td>
<td>$I_d/\mu m$</td>
<td>$1$</td>
</tr>
<tr>
<td>$A_{chip}$</td>
<td>$\alpha$</td>
<td>$10^1$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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

Past 30 years scaling

Merit: N, f increase
Demerit: P increase

$V_{dd}$ scaling insufficient

Additional significant increase in $I_d$, f, P

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

<table>
<thead>
<tr>
<th>Period</th>
<th>Expected limit(size)</th>
<th>Cause</th>
</tr>
</thead>
<tbody>
<tr>
<td>Late 1970’s</td>
<td>1µm:</td>
<td>SCE</td>
</tr>
<tr>
<td>Early 1980’s</td>
<td>0.5µm:</td>
<td>S/D resistance</td>
</tr>
<tr>
<td>Early 1980’s</td>
<td>0.25µm:</td>
<td>Direct-tunneling of gate SiO₂</td>
</tr>
<tr>
<td>Late 1980’s</td>
<td>0.1µm:</td>
<td>‘0.1µm brick wall’ (various)</td>
</tr>
<tr>
<td>2000</td>
<td>50nm:</td>
<td>‘Red brick wall’ (various)</td>
</tr>
<tr>
<td>2000</td>
<td>10nm:</td>
<td>Fundamental?</td>
</tr>
</tbody>
</table>
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.
Finally, there appears to be a fundamental limit of approximately quarter micron channel length, where certain physical effects such as the tunneling through the gate oxide ..... begin to make the devices of smaller dimension unworkable.
Direct-tunneling effect

Potential Barrier

Wave function

Direct tunneling current

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 = 0.1 µm
Gate leakage: \( I_g \propto \text{Gate Area} \propto \text{Gate length (Lg)} \)

Drain current: \( I_d \propto 1/\text{Gate length (Lg)} \)

\( \text{Lg} \to \text{small}, \)  
Then, \( I_g \to \text{small}, I_d \to \text{large}, \) Thus, \( I_g/I_d \to \text{very small} \)
Never Give Up!

No one knows future!

There would be a solution!

Think, Think, and Think!

Or, Wait the time!

Some one will think for you
Transistor Scaling Continues

- 90nm node: Lg=50nm
- 65nm node: Lg=35nm
- 45nm node: Lg=25nm
- 32nm node: Lg=15nm
- 22nm node: Lg=10nm

Approximately 30% every two years.

ITRS 2003 (HP)

Qi Xinag, ECS 2004, AMD
Downsizing limit? Channel length?

Electron wave length

10 nm
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

Downsizing limit!

Channel length

Gate oxide thickness

Gate Oxd

Channel
Prediction now!

Electron wave length
10 nm

Tunneling distance
3 nm

Atom distance
0.3 nm

MOSFET operation
$Lg = 2 \sim 1.5 \text{ nm}$?

Below this, no one knows future!
Maybe, practical limit around 5 nm.

When Gate length Smaller,

→ Subthreshold Leakage Current Larger

Subthreshold Leakage Current

Subthreshold Current Is OK at Single Tr.

But not OK For Billions of Trs.

Id

OFF

ON

Vg

Vth (Threshold Voltage)

Vg=0V

44
We have to reduce the Supply voltage.

Then Vth should be lowered.

Subthreshold leakage current increase

\[ Vg = 0V \]

\[ V_{th} \]
Electron wave length: 10 nm
Tunneling distance: 3 nm
Atom distance: 0.3 nm

Prediction now!

Practical limit for integration:
Lg = 5 nm?

MOSFET operation:
Lg = 2 ~ 1.5 nm?

Below this, no one knows future!
However, Gate oxide thickness 2 orders magnitude smaller Close to limitation!!

ULTIMATE LIMIT

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
So, we are now in the limitation of downsizing?

Do you believe this or do not?
There is a solution! To use high-k dielectrics

Thin gate SiO₂

Almost the same electric characteristics

Thick gate high-k dielectrics

Thick
Small leakage Current

However, very difficult and big challenge!

Remember MOSFET had not been realized without Si/SiO₂!
Choice of High-k elements for oxide

Candidates

Unstable at Si interface

- Si + MO\textsubscript{X} M + SiO\textsubscript{2}
- Si + MO\textsubscript{X} MSi\textsubscript{X} + SiO\textsubscript{2}
- Si + MO\textsubscript{X} M + MSi\textsubscript{X}O\textsubscript{Y}

Gas or liquid at 1000 K

Radio active

He

B C N O F Ne

Al Si P S Cl Ar

K Ca Sc Ti V Cr Mn Fc Co Ni Cu Zn Ga Ge As Se Br Kr

Sr Y Zr Nb Mo Tc Ru Rb Pd Ag Cd In Sn Sb Te I Xe

Cs Ba Hf Ta W Re Os Ir Pt Au Hg Ti Pb 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

La\textsubscript{2}O\textsubscript{3} based dielectrics are thought to be the next generation materials, which may not need a thicker interfacial layer.

HfO\textsubscript{2} based dielectrics are selected as the first generation materials, because of their merit in 1) band-offset, 2) dielectric constant 3) thermal stability.

Choice of High-k elements for oxide are selected as the first generation materials, because of their merit in 1) band-offset, 2) dielectric constant 3) thermal stability R. Hauser, IEDM Short Course, 1999
Conduction band offset vs. Dielectric Constant

Leakage Current by Tunneling

XPS measurement by Prof. T. Hattori, INFOS 2003
High-k gate insulator MOSFETs for Intel: EOT=1nm

EOT: Equivalent Oxide Thickness
For the past 45 years, SiO₂ and SiON have been used for gate insulators. Today, EOT = 1.0 nm. The EOT limit is 0.7~0.8 nm. One order of magnitude improvement is desired.

EOT can be reduced further beyond 0.5 nm by using direct contact to Si, by choosing appropriate materials and processes.
### Choice of High-k elements for oxide

**Candidates**

<table>
<thead>
<tr>
<th>Elements</th>
<th>Status</th>
<th>Conditions</th>
</tr>
</thead>
<tbody>
<tr>
<td>H</td>
<td></td>
<td>Unstable at Si interface</td>
</tr>
<tr>
<td>Li, Be, Mg</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Na</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ca, Sc, Sr, Y, Zr</td>
<td>Stable</td>
<td>HfO₂ based dielectrics are selected as the first generation materials, because of their merit in 1) band-offset, 2) dielectric constant, 3) thermal stability.</td>
</tr>
<tr>
<td>K, Ca, Sc, Sr, Y, Zr, Cs, Ba, Hf</td>
<td>Al, Si</td>
<td>Si + MOₓ M + SiO₂</td>
</tr>
<tr>
<td>K, Ca, Sc, Sr, Y, Zr, Cs, Ba, Hf, La</td>
<td>Al, Si, P, S, Cl, Ar</td>
<td>Si + MOₓ M + SiO₂</td>
</tr>
</tbody>
</table>

HfO₂ based dielectrics are selected as the first generation materials, because of their merit in 1) band-offset, 2) dielectric constant, 3) thermal stability.

La₂O₃ based dielectrics are thought to be the next generation materials, which may not need a thicker interfacial layer.

R. Hauser, IEDM Short Course, 1999
Gate Leakage vs EOT, \((V_g = |1| V)\)
EOT = 0.48 nm

Transistor with La2O3 gate insulator

Our results
EOT=0.37nm

La2O3

EOT=0.37nm    EOT=0.40nm    EOT=0.48nm

0.48 → 0.37nm Increase of Id at 30%
New material research will give us many future possibilities and the most important for Nano-CMOS!

Not only for high-k!

- New material for Metal gate electrode
- New material for High-k gate dielectric
- New channel material
- New material For Metal S/D
6 µm NMOS LSI in 1974

Layers
1. Si substrate
2. Field oxide
3. Gate oxide
4. Poly Si
5. S/D
6. Interlayer
7. Aluminum
8. Passivation

Materials
1. Si
2. SiO₂
3. BPSG
4. Al
5. PSG

Atoms
1. Si
2. O
3. P
4. B
5. Al
(H, N, Cl)
New materials

Just examples!
Many other candidates

Y. Nishi, Si Nano Workshop, 2006,
(S. Sze, Based on invited talk at Stanford Univ., Aug. 1999)
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)
Now: After 45 Years from the 1st single MOSFETs

32 Gb and 16Gb NAND, SAMSUNG
### NAND flash trend

<table>
<thead>
<tr>
<th>Capacity</th>
<th>Node</th>
<th>1&lt;sup&gt;st&lt;/sup&gt; Fabrication</th>
<th>Production</th>
</tr>
</thead>
<tbody>
<tr>
<td>512Mbit</td>
<td>120nm</td>
<td>2000</td>
<td>2001</td>
</tr>
<tr>
<td>1Gbit</td>
<td>100nm</td>
<td>2001</td>
<td>2002</td>
</tr>
<tr>
<td>2Gbit</td>
<td>90nm</td>
<td>2002</td>
<td>2003</td>
</tr>
<tr>
<td>4Gbit</td>
<td>70nm</td>
<td>2003</td>
<td>2004</td>
</tr>
<tr>
<td>8Gbit</td>
<td>60nm</td>
<td>2004</td>
<td>2005</td>
</tr>
<tr>
<td>16Gbit</td>
<td>50nm</td>
<td>2005</td>
<td>2006</td>
</tr>
<tr>
<td>32Gbit</td>
<td>40nm</td>
<td></td>
<td></td>
</tr>
<tr>
<td>256Gbit</td>
<td>20nm</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Even Tbit</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Even Tbit would be possible in future!
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
$8\mu m \rightarrow 6\mu m \rightarrow 4\mu m \rightarrow 3\mu m \rightarrow 2\mu m \rightarrow 1.2\mu m \rightarrow 0.8\mu m \rightarrow 0.5\mu m$

$\rightarrow 350\text{nm} \rightarrow 250\text{nm} \rightarrow 180\text{nm} \rightarrow 130\text{nm} \rightarrow 90\text{nm} \rightarrow 65\text{nm} \rightarrow 45\text{nm}$

$\rightarrow 32\text{nm} \rightarrow 22\text{nm} \rightarrow 16\text{nm} \rightarrow 11\text{nm} \rightarrow 8\text{nm}?? \rightarrow 5.5\text{nm} ??$
Clock frequency does not increase aggressively anymore.

Advantage in RISC
Simple configuration

Advantage in SISC
Era for ‘out of order’

Even decreased!

Multi Core
Clock
Performance

Source: Mitsuo Saito, Toshiba
MPU “GHz” by “Cores” ITRS2007

Cell Broadband Engine

6GHz capability for SRAM

Source: IBM, Toshiba, Sony ISSCC2007 and 08

Source: 2007 ITRS Winter Public Conf.
Structure and technology innovation (ITRS 2007)

Source: 2008 ITRS Summer Public Conf.
## Intel’s SRAM test chip trend

**Source:** B. Krzanich, S. Natrajan, Intel Developer’s Forum 2007


<table>
<thead>
<tr>
<th>Process name</th>
<th>Lithography</th>
<th>1st production</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1264</td>
<td>65nm</td>
<td>2005</td>
</tr>
<tr>
<td>P1266</td>
<td>45nm</td>
<td>2007</td>
</tr>
<tr>
<td>P1268</td>
<td>32nm</td>
<td>2009</td>
</tr>
<tr>
<td>P1270</td>
<td>22nm</td>
<td>2011</td>
</tr>
</tbody>
</table>

*Only schedule has been published*

---

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

- **Technology**: 90 nm Process, 65 nm Process, 45 nm Process, 32 nm Process
- **Cell size**: 1.0 $\mu$m$^2$ cell, 0.57 $\mu$m$^2$ cell, 0.346 $\mu$m$^2$ cell, 0.182 $\mu$m$^2$ 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

---

*Graph showing the down-scaling trend from 180nm to 32nm with a 0.5x every 2 years schedule.*
Cell size reduction trends

1/2 or 2/3 per cycle?

- Intel
  - 65nm Apr. 2004
  - 45nm Jan. 2006
  - 32nm Sep. 2007

- TSMC
  - 45nm Dec. 2007
  - 32nm Dec. 2007

- IBM Alliance (Consortium)
  - 32nm Dec. 2007

- Press release
  - 22nm Aug. 2008
Most Difficult part of SRAM down-scaling is Vdd down-scaling

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

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

Thus, Vdd down-scaling is more severe in SRAM than logic part of the circuits
Intel® Xeon® 7400 Series (Dunnington)

45 nm high-k6 cores
16MB shared L3 cache

Source: Intel Developer Forum 2008

Cache occupies huge area

→ Cell size of SRAM should be minimized
→ Isd-leak should be minimized
  → Vth are often designed to be higher than Min. logic Vth
  → Lg are often designed to be larger than Min. logic Lg
Nehalem (Intel) 2, 4 or 8 Cores

Voltage/Frequency Partitioning

- DDR Vcc
- Core Vcc
- Uncore Vcc

Dynamic Power Management

8T SRAM Cell

- 32kB L1 I-cache
- 32kB L1 D-cache
- 256kB L2 cache

6T SRAM Cell

- 8 MB L3 cache

Source: Intel Developer Forum 2008
6T and 8T Cell

6T Cell

Cell size is small
For high density use

8T Cell

Add separate read function
Cell size increase 30%
For low voltage use

Source: Morita et. al, Symp. on VLSI Circ. 2007
More Moore and More than Moore

Moore’s Law & More

More than Moore: Diversification

Analog/RF  Passives  HV  Power  Sensors  Actuators  Biochips

Interacting with people and environment

Non-digital content System-in-package (SiP)

Combining SoC and SiP: Higher Value Systems

Information Processing

Digital content System-on-chip (SoC)

Baseline CMOS: CPU, Memory, Logic

130nm
90nm
65nm
45nm
32nm
22nm

Beyond CMOS

More Moore: Miniaturization

Question what is the other side of the cloud?

ITRS 2005 Edition

FinFET to Nanowire

Channel conductance is well controlled by Gate even at L=5nm
- 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.
Si nanowire FET with Semi-1D Ballistic Transport

Merit of Si-nanowire


Trade off

Carrier scattering probability

Small Large

# of quantum channel

Small Large

E-k band

High Conduction (1D)

Go=77.8 µS/wire

Multiple quantum channel (QC) used for conduction

High-density lateral and vertical integration

Selection of MOSFET structure for high conduction: Nano-wire or Nano-tube FETs is promising

3 methods to realize High-conduction at Low voltage
M1. Use 1D ballistic conduction
M2. Increase number of quantum channel
M3. Increase the number of wire or tube per area
   3D integration of wire and tubes

For suppression of Ioff, the Nanowire/tube is also good.
1D conduction per one quantum channel:
\[ G = \frac{2e^2}{h} = 77.8 \, \mu\text{S/wire or tube} \]
regardless of gate length and channel material

That is 77.8 \( \mu \text{A/wire} \) at 1V supply

This an extremely high value
Increase the Number of quantum channels

By Prof. Shiraishi of Tsukuba univ.

Energy band of Bulk Si

Energy band of 3 x 3 Si wire

4 channels can be used
Maximum number of wires per 1 µm

Front gate type MOS 165 wires /µm

Surrounded gate type MOS 33 wires/µm

6nm pitch
By nano-imprint method

Metal gate electrode (10nm)
High-k gate insulator (4nm)
Si Nano wire (Diameter 2nm)

30nm pitch: EUV lithography

Surrounded gate MOS
Increase the number of wires towards vertical dimension

Si/Strained SiGe/Si

Depo. Temp.: 500°C

Si

Si$_{0.5}$Ge$_{0.5}$

6 nm

Si(100)

(a) Si/SiGe/Si epitaxial wafer
(b) Dry Etching
(c) Selective Etching
(d) H$_2$ Annealing
(e) Gate Oxide
(f) Gate, S/D Formation

Si/SiGe multi stacked wafer
Dry Etching
Selective Etching
H$_2$ Annealing
Our new roadmap

Extended CMOS: More Moore + CMOS logic

Beyond the horizon

Natural direction of downsizing

Diameter = 2nm

Si Channel

Si Fin, Tri-gate

Si Nano wire

III-V & Ge Nano wire

Tube

CNT

Nanowire

Tube, Ribbon

Graphene

Diameter = 10nm

Selection

High conduction

By 1D conduction

ITRS

Beyond CMOS

Extended CMOS

More Moore

More Moore

Cloud

ITRS Beyond CMOS

More Moore ??

Extended CMOS

More Moore ??

ITRS

2007 2010 2015 2020 2025 2030 2035

PJT (2007~2012)
Our roadmap for R & D
Source: H. Iwai, IWJT 2008

Current Issues

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

III-V & Ge Nanowire
- High-k gate insulator
- Wire formation technique

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

Graphene:
- Graphene formation technique
- Suppression of off-current
- Very small bandgap or no bandgap (semi-metal)
- Control of ribbon edge structure which affects bandgap
Miniaturization of Interconnects on PCB (Printed Circuit Board)

5 nm?

2020?
Brain

Ultra small volume
Small number of neuron cells
Extremely low power

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

Sensor
Infrared
Humidity
CO₂

Mosquito

System and Algorism becomes more important!

But do not know how?

Dragonfly is further high performance
Thank you for your attention!