Past and future for micro- and nano-electronics, focusing on Si integrated circuits technology

June 2, 2008

@National Technical University of Athens

Hiroshi Iwai,
Toyo Institute of Technology
CMOS Technology: Indispensable for our human society

All 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
CMOS 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>~</td>
</tr>
<tr>
<td>1970s LSI (Large Scale Integrated Circuit)</td>
<td>~1,000,000</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>
Exponential Cost Reduction

Cost per Transistor

As the number of transistors goes UP

Cost per transistor goes DOWN

Sources: WSS/Intel

SPIE Microlithography 2006, San Jose, CA USA
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 MOSFE
   - 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.
First Computer Eniac: made of huge number of vacuum tubes. Big size, huge power, short life time filament.

Today's pocket PC has much higher performance with extremely low power consumption.
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 SiO</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 0.25 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

Potential Barrier

Wave function
Direct tunneling leakage was found to be OK! In 1994 MOSFETs with 1.5 nm gate oxide, the direct tunneling leakage current starts to flow when the thickness is 3 nm.
Do not believe a text book statement, blindly!

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, AMI
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

Atom distance
0.3 nm

Downsizing limit!
Channel length
Gate oxide thickness
Electron wave length 10 nm

Tunneling distance 3 nm

Atom distance 0.3 nm

MOSFET operation
Lg = 2 ~ 1.5 nm?

Below this, no one knows future!
How about the integration of such small-geometry MOSFETs in a chip?

1) Integration of huge number of the ultra-small MOSFETs would consume too huge power and thus, creates too huge heat?

2) Integration of such ultra-small MOSFETs causes too huge variations in the transistor characteristics, which could make the circuit design impossible?

3) There are too many number of transistors in a chip for the circuit designers to manipulate? (design crisis),

4) There would be no merit of transistor downsizing in performance and power, because of RC (resistance capacitance product) of interconnect cannot be reduced aggressively any more?

5) Who will pay the huge development and production costs for the integration of such ultra-small MOSFETs? Note that the prices for the recent process equipments and the lithography mask became extremely high.
These concerns have been argued in the past 15 years at every new generation of the products, like the wolf boy.

Fortunately, the wolf has not come, and the concerns have not come true.

It is expected that we can go with several more generations for the integration.

There will be still a room for squeezing the technologies to obtain the merit of the scaling-down for integration.
The continuous progress of CMOS technologies for
- high-performance
- low power
is very important because of the 3 reasons:

1) Rapid progress of aging population and falling birth rate

1) Global warming

1) Semiconductor industry and world economy
1) Rapid progress of aging population and falling birth rate:

Replacement of some of the human jobs by intelligent machines – such as human type robot for elderly-care, for example.

For, the daily family use, much higher intelligence and much lower power consumption than those of today are required.
Robot in 21c cannot be made without integrated circuits.

Robot (21C)

Karakuri (Windup Mechanical) doll (18C) in Japan
2) Recent Significant Global Warming
We need reduce CO2 generation!

Low power technology is urgent request

Carbon Dioxide Variations

The Industrial Revolution Has Caused A Dramatic Rise in CO₂

Bio Age Cycles
CO2 Emission per capita 2002

The bar chart shows the CO2 emission per capita for various countries in 2002. The x-axis represents different countries, and the y-axis represents the CO2 emission per capita in t/a (tons per annum). The countries are listed from left to right in descending order of CO2 emission per capita.

- India has the lowest CO2 emission per capita among the listed countries.
- The USA has the highest CO2 emission per capita among the listed countries.
- Canada and Australia also have relatively high CO2 emissions per capita.
- Several other countries fall in between, with Japan and Germany having notable emissions as well.

Overall, the chart illustrates the variation in CO2 emissions per capita across different countries, with some having significantly higher emissions than others.
3) Semiconductor industry, and world economy

If there is no more downsizing such as
45 $\rightarrow$ 32 nm Logic, 8 Gbit $\rightarrow$ 16 Gbit Memory

- LSIs will not be sold well, and semiconductor companies will face a disaster.

- Equipment and martial companies as well.

- There is no more R & D for semiconductors and many people will loose their jobs.

$\rightarrow$ World economy crisis!
History and future of Transistor

Shrinking, Shrinking, and Shrinking!

and then, Shrinking, Shrinking, and Shrinking

\[ C, V \propto \frac{1}{L^2} \]

Switching speed \( \frac{CV}{I} \) → Decrease

Power consumption \( CV^2/2 \) → Decrease

Integration density: \( 1/L^2 \) → Increase

1970 \hspace{1cm} 2007

Gate length \hspace{1cm} 10,000 nm \hspace{1cm} 25 nm

Gate Oxd Thickness \hspace{1cm} 100 nm \hspace{1cm} 1 nm
For the past 45 years, SiO2 and SiON were used for gate insulator. Today, EOT=1.0nm, and the EOT limit is 0.7~0.8 nm. For a 45nm node, L_g=22nm. A 22nm node, L_g=11nm, is now introduced.

One order of magnitude improvement is achieved by the introduction of High-k materials, which allows for a smaller EOT=0.5nm. High-k materials directly contact Si, with SiO2 or SiON used at the Si interface.

Metal layers are used, with High-k materials such as HfO2 replacing SiO2/SiON for better performance.
Choice of High-k elements for oxide

Candidates

Unstable at Si interface
- Si + MOₓ M + SiO₂
- Si + MOₓ MSiₓ + SiO₂
- Si + MOₓ M + MSiₓOᵧ

Gas or liquid at 1000 K
- Radioactive materials

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
EOT = 0.48 nm  
Our results
Transistor with La2O3 gate insulator
CMOS downsizing is critically important

However now, many people expect that we will reach limit in 2020.

Totally, new paradigm after reaching the downsizing limit.

What will be?
After 2020

There is no decrease in gate length around at 10 ~ 5 nm.

4 reasons.
After 2020

4 reasons for no downsizing anymore or No decrease in gate length

1. No increase of On-current (Drain current) because of already semi-ballistic conduction.
   Ballistic $\iff$ No scattering of carriers in channel
   Thus, all the carrier from the source reach drain

2. Increase of Off-current (Subthreshold current)

3. No decrease of Gate capacitance by parasitic components

4. Increase in production cost.
After 2020

What will be the world with no gate length reduction?
Question what is the other side of the cloud?
### Table 1: Device Characteristics

<table>
<thead>
<tr>
<th>Device</th>
<th>FET</th>
<th>RSFQ</th>
<th>1D structures</th>
<th>Resonant Tunneling Devices</th>
<th>SET</th>
<th>Molecular</th>
<th>QCA</th>
<th>Spin transistor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cell Size</td>
<td>100 nm</td>
<td>0.3 µm</td>
<td>100 nm</td>
<td>100 nm</td>
<td>40 nm</td>
<td>Not known</td>
<td>60 nm</td>
<td>100 nm</td>
</tr>
<tr>
<td>Density ((\text{cm}^{-2}))</td>
<td>3E9</td>
<td>1E6</td>
<td>3E9</td>
<td>3E9</td>
<td>6E10</td>
<td>1E12</td>
<td>3E10</td>
<td>3E9</td>
</tr>
<tr>
<td>Switch Speed (\text{GHz})</td>
<td>700 GHz</td>
<td>1.2 THz</td>
<td>Not known</td>
<td>1 THz</td>
<td>1 GHz</td>
<td>Not known</td>
<td>30 MHz</td>
<td>700 GHz</td>
</tr>
<tr>
<td>Circuit Speed (\text{GHz})</td>
<td>30 GHz</td>
<td>250–800 GHz</td>
<td>30 GHz</td>
<td>30 GHz</td>
<td>1 GHz</td>
<td>&lt;1 MHz</td>
<td>1 MHz</td>
<td>30 GHz</td>
</tr>
<tr>
<td>Switching Energy, (\text{J})</td>
<td>(2\times10^{-18})</td>
<td>(&gt;1.4\times10^{-17})</td>
<td>(2\times10^{-18})</td>
<td>(&gt;2\times10^{-18})</td>
<td>(&gt;1.5\times10^{-17})</td>
<td>(1.3\times10^{-16})</td>
<td>(&gt;1\times10^{-18})</td>
<td>(2\times10^{-18})</td>
</tr>
<tr>
<td>Binary Throughput, (\text{GBit/\text{ns/cm}^2})</td>
<td>86</td>
<td>0.4</td>
<td>86</td>
<td>86</td>
<td>10</td>
<td>N/A</td>
<td>0.06</td>
<td>86</td>
</tr>
</tbody>
</table>

We HAVE IDENTIFIED NO VIABLE EMERGING LOGIC TECHNOLOGIES for Information Processing beyond CMOS.
We could keep the Moore’s law after 2020
Without downswing the gate length

What is Moore’s law.
Keep increase of the number of components.
Cost per components decreases!

Gordon Moore

We could keep the Moore’s law after 2020
Without downswing the gate length

What is Moore’s law.
\[ \rightarrow \text{to increase the number (\#) of Tr. In a chip} \]

Now, \# of Tr. in a chip is limited by power.
\[ \rightarrow \text{key issue is to reduce the power.} \]
\[ \rightarrow \text{to reduce the supply voltage is still effective} \]

To develop devices with sufficiently high drain current under low supply voltage is important.
FinFET to Nanowire

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

F.-L.Yang, VLSI2004
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:
\[ G = \frac{2e^2}{h} = 77.5 \, \mu S/wire \text{ or tube} \]
regardless of gate length and channel material

That is 77.5 mA/wire at 1V supply

This an extremely high value

However, already 20mA/wire was obtained experimentaly by Samsung
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/0.5Ge0.5
6 nm
Si(100)

(c) Selective Etching
(b) Dry Etching
(a) Si/SiGe/Si epitaxial wafer
(d) H₂ Annealing
(e) Gate Oxide
(f) Gate, S/D Formation

Si/SiGe multi stacked wafer
Dry Etching
Selective Etching
H₂ Annealing
Our new roadmap

Extended CMOS: More Moore + CMOS logic

Beyond the horizon

Natural direction of downsizing

Si Channel

Diameter = 2nm

Nanowire

Diameter = 10nm

Graphene

Si Fin, Tri-gate

Si Nano wire

III-V及Ge Nano wire

Tube

CNT

Tube, Ribbon

Selection

ITRS Beyond CMOS

High conduction

By 1D conduction

Extended CMOS

More Moore

More Moore ??

ITRS

Cloud
Miniaturization of Interconnects on PCB (Printed Circuit Board)
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

Brain

System and Algorism becomes more important!

But do not know how?

Dragonfly is further high performance
Thank you for your attention!