Ultimate CMOS scaling

Hiroshi Iwai

Frontier Research Center
Tokyo Institute of Technology
1. Background for nano-electronics
1900 “Electronics” started.

Device: Vacuum tube
Device feature size: 10 cm
Major Appl.: Amplifier (Radio, TV, Wireless etc.)

→ Technology Revolution

1970 “Micro-Electronics” started.

Device: Si MOS integrated circuits
Device feature size: 10 μm
Major Appl.: Digital (Computer, PC, etc.)

→ Technology Revolution
2000 “Nano-Electronics” started.

Device: Still, Si CMOS integrated circuits
Device feature size: 100 nm
Major Appl.: Digital (\(\mu\)-processor, cell phone, etc.)

→Technology Revolution??

Maybe, just evolution or innovation!

But very important so many innovations!
Now, 2013 “Nano-Electronics” continued.

Device: Still, Si CMOS integrated circuits
Device feature size: near 10 nm
Major Appl.: Still Digital (μ-processor, cell phone, etc.)

Still evolution and innovation..

But, so many important emerging applications for smart society.
Questions for future

Future, “Nano-Electronics” still continued?

Device: Still, Si CMOS integrated circuits?

Device feature size: ? nm, what is the limit?

Application: New application?

→ Any Technology Revolution?
What is special or new for Nano-Electronics?

In 1990’s, people expected completely new mechanism or operational principle due to the nano size, like quantum mechanical effects.

However, no fancy new operational principle was found.

At least for logic application, there is no success story for “Beyond CMOS devices” to replace Si-CMOS.

Of course, I do not deny the importance of Beyond CMOS technology development. It is becoming very important as CMOS approach its limit.
Ballistic conduction will not happen even decreasing channel length.

![Diagram showing different transport mechanisms](image)

- **Diffusive transport**
  - $L >> \lambda$

- **Quasi-Ballistic transport**
  - $L \sim \lambda$

- **Ballistic transport**
  - $L < \lambda$

Ballistic transport will never happen for MOSFET because of back scattering at drain

With decreasing channel length, Drain current increase continue.

Also, 1D quantum conduction, or ballistic conduction will not happen.

(1D quantum conduction: $77.8\,\mu$S regardless of the length and material)
2. Importance of nano-electronics as integrated circuits
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
1960: First MOSFET by D. Kahng and M. Atalla

Silicon (Si)

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

DRAM  Intel 1103

MPU   Intel 4004
In 2012

Most Recent SD Card

128GB (Bite)
= 128G X 8bit
= 1T(Tera)bit

1T = 10^{12} = 1 Trillion

World Population: 7 Billion
Brain Cell: 10\sim 100 Billion
Stars in Galaxy: 100 Billion
128 GB = 1Tbit

2.4cm X 3.2cm X 0.21cm

Volume: 1.6 cm³  Weight: 2g

Voltage: 2.7 - 3.6V

Old Vacuum Tube:
5cm X 5cm X 10cm, 100g, 50W

What are volume, weight, power consumption for 1Tbit
Old Vacuum Tube: 5cm X 5cm X 10cm

1Tbit = 10,000 X 10,000 X 10,000 bit

Volume = (5cm X 10,000) X (5cm X 10,000) X (10cm X 10,000)
= 0.5km X 0.5km X 1km

Pingan International Finance Center
Shanghai, China (Year 2016)

Indian Tower
Mumbai, India (Year 2016)

Burj Khalifa
Dubai, UAE (Year 2010)

500 m

828 m

700 m

700 m

10,000 m

1Tbit
Old Vacuum Tube: 100W

1Tbit = $10^{12}$bit

Power = $0.05 \text{kW} \times 10^{12} = 50 \text{ TW}$

Nuclear Power Generator

1MkW = 1BW

We need 50,000 Nuclear Power Plant for just one 128 GB memory

In Japan we have only 54 Nuclear Power Generator

Last summer Tokyo Electric Power Company (TEPCO) can supply only 55BW.

We need 1000 TEPCO just one 128 GB memory

Imagine how many memories are used in the world!
So progress of integrated circuits is extremely important for power saving
Brain: Integrated Circuits

Ear, Eye: Sensor

Mouth: RF/Opto device

Stomach: PV device

Hands, Legs: Power device
Near future smart-society has to treat huge data.

Demand to high-performance and low power CMOS become much more stronger.
Semiconductor Device Market will grow 5 times in 12 years, even though, it is very matured market!!

2011 300B USD

→

2015 1,500B USD

Gartner: By K. Kim, CSTIC 2012
2. Current status of Si-CMOS device technologies
Downsizing

Important for

Decreasing cost, power

Increasing performance
Feature Size / Technology Node

(1970) 10 μm → 8 μm → 6 μm → 4 μm → 3 μm → 2 μm → 1.2 μm → 0.8 μm → 0.5 μm → 0.35 μm → 0.25 μm → 180 nm → 130 nm → 90 nm → 65 nm → 45 nm → 32 nm → 22 nm (2012)

From 1970 to 2013 (This year)

43 years 1 generation
18 generations 2.5 years
Line width: 1/450 Line width: 1/1.43 = 0.70
Area: 1/200,000 Area: 1/2 = 0.5
Problem for downsizing

Region governed by gate bias
Region governed by drain bias

Gate oxide
Gate metal

Source
Drain

Channel
Substrate

0V
1V

0V < $V_{dep} < 1V$

Depletion Region (DL)

No $t_{ox}$ thinning
$V_{dd}$

Large $I_{OFF}$

Large $I_{OFF}$

$0V < V_{dep} < 1V$

$V_{dd}$

$0V$

$0V$

$0V$

$t_{ox}$ and $V_{dd}$ have to be decreased for better channel potential control $\rightarrow I_{OFF}$ Suppression
$L_{\text{gate}}$ and $t_{\text{ox}}(\text{EOT})$ scaling trend

A. Toriumi (Tokyo Univ), IEDM 2006, Short Course
EOT = 0.40 nm
What would be the limit of downsizing!

Tunneling distance

Source

Channel

Drain

3 nm
Subthreshold leakage current of MOSFET

- Subthreshold leakage current of MOSFET

- Subthreshold leakage current is OK at Single Tr. level

- But not OK For Billions of Trs.

- Subthreshold region

- $V_{th}$ (Threshold Voltage)

- $V_g = 0V$

- $I_{on}$

- $I_{off}$

- OFF

- ON
Subthreshold leakage current will limit the downsizing

\[ I_d (A/\mu \text{m}) \]

\[ I_{on} \quad I_{off} \]

\[ 10^{-7} \quad 10^{-5} \quad 10^{-11} \quad 10^{-9} \]

\[ V_{th} \quad V_d \quad 0.15 \text{ V} \quad 0.3 \text{ V} \quad 0 \quad 0.5 \quad 1 \]

Electron Energy Injection

Boltzmann statics

\[ \text{Exp} (-qV/kT) \]

\[ S \]
The limit is different depending on application.
How far can we go for production?

**Past**

0.7 times per 2.5 years

In 43 years: 18 generations,
Size 1/450, Area 1/200,000

1970年

10µm → 8µm → 6µm → 4µm → 3µm → 2µm → 1.2µm → 0.8µm → 0.5µm →
0.35µm → 0.25µm → 180nm → 130nm → 90nm → 65nm → 45nm → 32nm

**Now**

→ 22nm → 16nm → 11.5 nm → 8nm → 5.5nm? → 4nm? → 2.9 nm?

**Future**

· At least 4.5 generations to 8 ~ 5 nm
Suppression of subthreshold leakage by surrounding gate structure

Planar

Surrounding gate
Because of off-leakage control,

Planar $\rightarrow$ Fin $\rightarrow$ Nanowire

Planar FET

Fin FET

Nanowire FET
Nanowire structures in a wide meaning

Fin

Tri-gate

Ω-gate

All-around
Our work at TIT: $\Omega$-gate Si Nanowire

S. Sato et al., pp.361, ESSDERC2010 (Tokyo Tech.)

- Conventional CMOS process
- High drive current
  \(1.32 \text{ mA/\mu m} @ I_{OFF}=117 \text{ nA/\mu m}\)
- DIBL of 62mV/V and SS of 70mV/dec for nFET
More Moore to More More Moore

Technology node |
<table>
<thead>
<tr>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>65nm</td>
</tr>
<tr>
<td>Lg 35nm</td>
</tr>
</tbody>
</table>

Now

Future

15nm, 11nm, 8nm, 5nm, 3nm

Main stream (Fin,Tri, Nanowire)

Si

Planar

Tri-Gate

Alternative (ETSOI)

Si is still main stream for future !!

M. Bohr, pp.1, IEDM2011 (Intel)
P. Packan, pp.659, IEDM2009 (Intel)
C. Auth et al., pp.131, VLSI2012 (Intel)
T. B. Hook, pp.115, IEDM2011 (IBM)
S. Bangsaruntip et al., pp.297, IEDM2009 (IBM)
High-k gate dielectrics

Continued research and development

SiO_2 IL (Interfacial Layer) is used at Si interface to realize good mobility

Technology for direct contact of high-k and Si is necessary

EOT=0.52 nm
Remote SiO_2-IL scavenging HfO_2 (IBM)

Hf-based oxides

| 45nm EOT:1nm | 32nm EOT:0.95nm | 22nm EOT:0.9nm | 15nm, 11nm, 8nm, 5nm, 3nm |

SiO_2 IL (Interfacial Layer) is used at Si interface to realize good mobility

Remote SiO_2-IL scavenging HfO_2 (IBM)

Continued research and development

K. Mistry, et al., p.247, IEDM 2007, (Intel)
T.C. Chen, et al., p.8, VLSI 2009, (IBM)
T. Ando, et al., p.423, IEDM2009, (IBM)
Supplementary text for graphs:

Supply voltage affects very much!

**ION** and **IOFF** benchmark

**NMOS**

- Intel [1] Bulk 32nm \(V_{DD}=0.8\) V
- Intel [1] Tri-Gate 22nm \(V_{DD}=0.8\) V
- Intel [2] Bulk 45nm \(V_{DD}=1\) V
- Samsung [3] Bulk 20nm \(V_{DD}=0.9\) V
- IBM [5] GAA NW \(V_{DD}=1\) V
- IBM [6] FinFET 25nm \(V_{DD}=1\) V
- STMicro. [8] GAA NW \(V_{DD}=0.9\) V
- Tokyo Tech. [9] \(\Omega\)-gate NW \(V_{DD}=1\) V

**PMOS**

- Intel [1] Bulk 32nm \(V_{DD}=0.8\) V
- Intel [1] Tri-Gate 22nm \(V_{DD}=0.8\) V
- Intel [2] Bulk 45nm \(V_{DD}=1\) V
- IBM [5] GAA NW \(V_{DD}=1\) V
- IBM [6] FinFET 25nm \(V_{DD}=1\) V
- STMicro. [8] GAA NW \(V_{DD}=1.1\) V
- Samsung [3] Tri-Gate NW \(V_{DD}=0.9\) V
- IBM [10] ETSOI \(V_{DD}=0.9\) V
- IBM [7] ETSOI \(V_{DD}=1\) V
- IBM [7] ETSOI \(V_{DD}=1\) V
- IBM [10] ETSOI \(V_{DD}=0.9\) V

References:

[10] K. Cheng et al., pp.419, IEDM2012 (IBM)
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Bulk Planar</td>
<td>Tri-Gate</td>
<td>Tri-Gate</td>
<td>ETSOI</td>
<td>Bulk Planar</td>
<td>GAA NW</td>
<td>GAA NW</td>
<td>Ω-gate NW</td>
</tr>
<tr>
<td></td>
<td>45nm</td>
<td>22nm</td>
<td>NW</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lg (nm)</td>
<td>35</td>
<td>30</td>
<td>30</td>
<td>14</td>
<td>22</td>
<td>20</td>
<td>35/25(nFET/pFET)</td>
<td>22/30(nFET/pFET)</td>
</tr>
<tr>
<td></td>
<td>32nm</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gate</td>
<td>Hf-based</td>
<td>Hf-based</td>
<td>SiO₂</td>
<td>HfO₂</td>
<td>HfO₂ ?</td>
<td>Hf-based</td>
<td>HfZrO₂</td>
<td>SiO₂</td>
</tr>
<tr>
<td>Dielectrics</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EOT (nm)</td>
<td>1</td>
<td>0.95</td>
<td>0.9</td>
<td>3</td>
<td>~1</td>
<td>-</td>
<td>1.5</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vth (V)</td>
<td>~0.4</td>
<td>~0.3</td>
<td>~0.2</td>
<td>~0.15(nFET)</td>
<td>0.3~0.4</td>
<td>~0.3</td>
<td>0.3~0.4</td>
<td>~0.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vdd (V)</td>
<td>1</td>
<td>1</td>
<td>0.8</td>
<td>1</td>
<td>0.9</td>
<td>0.9</td>
<td>1</td>
<td>1.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ion (mA/μm)</td>
<td>1.36/1.07</td>
<td>1.53/1.23</td>
<td>1.26/1.1</td>
<td>0.83(nFET)</td>
<td>0.59/0.62(I_{eff})</td>
<td>1.2/1.05</td>
<td>0.83/0.95</td>
<td>2.05/1.5</td>
</tr>
<tr>
<td>nFET/pFET</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DIBL (mV/V)</td>
<td>~150</td>
<td>~200</td>
<td>46/50</td>
<td>&lt;50</td>
<td>-</td>
<td>104/115</td>
<td>65/105</td>
<td>56/9</td>
</tr>
<tr>
<td>nFET/pFET</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SS (mV/dec)</td>
<td>-</td>
<td>~100</td>
<td>~70</td>
<td>&lt;80</td>
<td>-</td>
<td>87</td>
<td>85</td>
<td>&lt;80</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
**Ω-gate Si Nanowire**

S. Sato et al., pp.361, ESSDERC2010 (Tokyo Tech.)

- Conventional CMOS process
- High drive current
  \[ (1.32 \text{ mA/µm} \ @ \ I_{\text{OFF}} = 117 \text{ nA/µm}) \]
- DIBL of 62mV/V and SS of 70mV/dec for nFET
Effective Electron Mobility (cm²/Vs) vs. Inversion Carrier Density (cm⁻²)

- **A₁ (12x19)**: $h_{NW} \times w_{NW}$ (nm²)
- **A₂ (12x28)**
- **B (20x10)**

SOI Planar

$T_{SOI} = 28$ nm
Distance from SiNW Surface (nm)

Electron density (x10^{19} \text{cm}^{-3})

Electron Density

Edge portion

Flat portion

Distance from SiNW Surface (nm)
Tri-gate implementation for transistors

Tri-gate has been implemented since 22nm node, enabling further scaling

<table>
<thead>
<tr>
<th></th>
<th>HP</th>
<th>MP</th>
<th>SP</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T_{OX,E}$ (nm)</td>
<td>0.9</td>
<td>0.9</td>
<td>0.9</td>
</tr>
<tr>
<td>$L_{GATE}$ (nm)</td>
<td>30</td>
<td>34</td>
<td>34</td>
</tr>
<tr>
<td>$I_{OFF}$ (nA/um)</td>
<td>20-100</td>
<td>5-20</td>
<td>1-5</td>
</tr>
</tbody>
</table>

C. Auth et al., pp.131, VLSI2012 (Intel)
• $L_g$ and EOT are larger than ITRS requirements

• Implementation of Tri-gate and lower $V_{th}/V_{dd}$ since 22nm

K. Mistry et al., pp.247, IEDM2007 (Intel).
P. Packan et al., pp.659, IEDM2009 (Intel).
C. Auth et al., pp.131, VLSI2012 (Intel).
PMOS channel under the gate

S/D region showing the SiGe epitaxy

A fin width of 8nm to balance SCE and $R_{ext}$

A fin height of 34nm to balance drive current vs. capacitance
SS of 71 and 72 mV/dec for HP NMOS and PMOS, respectively

DIBL of 30 and 35 mV/V for NMOS and PMOS, respectively

$V_{th}$ of 22 nm is about 0.1 ~0.2 V lower than that of 32nm

Very good $V_{th}$ control!
**Extremely Thin SOI (ETSOI)**

K. Cheng et al., pp.419, IEDM2012 (IBM)

Also, ET-SOI works very good!

- Hybrid CMOS
  - Si Channel nFET
  - Strained SiGe Channel pFET
- RO delay improvement over FinFET with FO = 2

---

![Diagram showing STI-first and STI-last processes](image)

![Graph showing off-current comparison between Si CMOS and Hybrid CMOS](image)

![Graph showing on-current comparison between Si and SiGe](image)

![Graph showing RO delay vs. Vdd for FinFET and ETSOI with FO = 3](image)
• **L<sub>g</sub> = 25~35nm GAA NW**

• Hydrogen anneal provide smooth channel surface

• Competitive with conventional CMOS technologies

• Scaling the dimensions of NW leads to suppressed SCE
3. Problems
### Short-channel effect

**T. Skotnicki, IEDM 2009 Short Course (STMicroelectronics)**

- **Bulk**
- **PD SOI**
- **Thick BOX FD SOI**
- **UTBB (FD SOI or SON)**
- **FinFET (independent gates)**
- **FinFET (common gates)**

\[
\text{SCE} = 0.64 \frac{\varepsilon_s}{\varepsilon_{ox}} \times EI \times \Phi_d
\]

\[
\text{DIBL} = 0.8 \frac{\varepsilon_s}{\varepsilon_{ox}} \times EI \times V_{ds}
\]

<table>
<thead>
<tr>
<th><strong>EI</strong></th>
<th>Formula</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1\times)</td>
<td>(\left(1 + \frac{X_j^2}{L_{el}^2}\right) \times \frac{T_{ox}}{T_{dep}} \times \frac{T_{si}}{L_{el}})</td>
<td>if max ch. doping is (10^{19}) cm(^{-3}), (T_{ep} \approx 12) nm then:</td>
</tr>
<tr>
<td>(1\times)</td>
<td>(\left(1 + \frac{T_{si}^2}{L_{el}^2}\right) \times \frac{T_{ox}}{T_{dep}} \times \frac{T_{si} + \lambda T_{box}}{L_{el}})</td>
<td>Tbox = 145 nm (If Tsi (\approx 3) nm then:</td>
</tr>
<tr>
<td>(1\times)</td>
<td>(\left(1 + \frac{T_{si}^2}{L_{el}^2}\right) \times \frac{T_{ox}}{T_{si} + \lambda T_{box}})</td>
<td>Tbox = 10 nm (If Tsi (\approx 3) nm then:</td>
</tr>
<tr>
<td>(1\times)</td>
<td>(\left(1 + \frac{T_{si}^2}{L_{el}^2}\right) \times \frac{T_{ox}}{T_{si}})</td>
<td>Tbox = 0.46 nm (If Tsi (\approx 3) nm then:</td>
</tr>
</tbody>
</table>

- **\(L_{el,\text{min}}\) (DIBL=100mV/V)**
  - Bulk: \(=34\) nm
  - PD SOI: \(=34\) nm
  - Thick BOX FD SOI: \(\approx 30\) nm
  - UTBB (FD SOI or SON): \(\approx 10\) nm
  - FinFET (independent gates): \(\approx 6\) nm
  - FinFET (common gates): \(\approx 2.7\) nm
Drain-induced barrier lowering

\[ DIBL = 0.80 \frac{\varepsilon_{Si}}{\varepsilon_{ox}} \left( 1 + \frac{X_j^2}{L_{el}^2} \right) \frac{T_{ox}}{L_{el}} \frac{T_{dep}}{L_{el}} V_{DS} \]

**BULK**

- \( T_{dep} = \frac{3}{4} L_{el} \)
- \( X_j = \frac{3}{4} L_{el} \)
- \( T_{ox} + 2\% \)
- \( \varepsilon_{Si} + 15\% \)

140mV/V

**III-V**

- \( T_{dep} = T_{Si} + \lambda T_{box} \)
- \( X_j = T_{Si} \geq 6nm \)

210mV/V

**FDSOI**

- \( T_{dep} = T_{Si}/2 \)
- \( X_j = T_{Si}/2 \)
- \( T_{Si} \geq 10nm \)

110mV/V

**ETSOI**

80mV/V

**UTBB**

70mV/V

**FinFET**
Sub-threshold Slope

\[ S = \frac{kT}{q} \ln(1 + \frac{C_{\text{dep}}}{C_{\text{ox}}} + \frac{\varepsilon_{\text{Si}}}{\varepsilon_{\text{ox}}} \frac{T_{\text{ox}}}{L_{\text{el}}} X_j \left(1 + \frac{3T_{\text{dep}}}{4L_{\text{el}}} \right) \sqrt{1 + \frac{2V_{\text{ds}}}{\Phi_d}}} \]

110 mV/dec

95 mV/dec

85 mV/dec

75 mV/dec

65 mV/dec

T_{\text{dep}} = T_{\text{si}} + \lambda T_{\text{box}}

X_j = T_{\text{si}}

T_{\text{dep}} = T_{\text{si}}/2

X_j = T_{\text{si}}/2

Log(\text{Id})

\[ S - \text{Subthreshold Slope} \]

V_{\text{g}}

BULK

T_{\text{dep}} = 3/4L_{\text{el}}

X_j = 3/4L_{\text{el}}

III-V

T_{\text{ox}} + 2A

\varepsilon_{\text{si}} + 15\%

ETSOI

FDSOI

UTBB

FinFET

RBB => Pstat
Problems in Multi-gate

Decreasing the diameter of NW

Improved short-channel control  \[\iffalse\text{Severe mobility degradation}\]  \fi

Need to decrease diameter for SCH

Significant $\mu$ degradation at diameter $< 10$ nm

---

S. Bangsaruntip et al., pp.297, IEDM2009 (IBM),
K. Tachi et al., pp.313, IEDM2009 (CEA-LETI)
Impact of Si thickness in FinFET

- Replacement Gate process
- TaN/HfO₂ gate stack
- Reduced $g_m$ and higher $V_{th}$ with decreasing Fin width
Problems in SOI

K. Uchida et al., pp.47, IEDM2002 (Toshiba)

Mobility is also decreased with decreasing the Si thickness of SOI transistor similar to the NW transistor.
Problem for nanowire

When wire diameter becomes less than 10 nm, sudden drop of $I_d$

1. Electron Scattering of every surface
2. Decrease of DOS

If diameter cannot be scaled, SCE cannot be suppressed.

Then, again aggressive EOT scaling of high-k is necessary.
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

300 K

$E_g = 1.12 \text{ eV}$
$E_i = 2.0 \text{ eV}$
$E_x = 1.2 \text{ eV}$
$E_o = 0.044 \text{ eV}$
$E_{T1} = 3.4 \text{ eV}$
$E_{T2} = 4.2 \text{ eV}$
Diameter dependence

By Profs. Oshiyama and Iwata, U. of Tokyo
Wire cross section dependence.

What cross section gives best solution for SCE suppression and drive current?

By Profs. Oshiyama and Iwata, U. of Tokyo
Significant increase in $V_{th}$ variability with decreasing Fin width
Smaller wire/fin width is necessary for SCE suppression

But mobility and $I_{ON}$ severely degrade with wire/fin width reduction

Therefore even in multi-gate structures, **EOT scaling should be accelerated** to provide SCE immunity
High-k beyond 0.5 nm
Gate oxide should be thicker than mono atomic layer

0.8 nm gate oxide thickness MOSFETs operate

0.8 nm $\rightarrow$ Distance of 3 Si atoms $\rightarrow$ 2 mono layers

*R. Chau, et al., (Intel) IWGI 2003*
Gate Leakage Power Density becomes significantly large with $L_g$ reduction, and thus, with $t_{ox}$ thinning!!
To use high-k dielectrics

**K**: Dielectric Constant

Thin SiO$_2$

K=4

Almost the same electric characteristics

K=20

Thick high-k dielectrics

5 times thicker

Small leakage Current

However, very difficult and big challenge!
Equivalent Oxide Thickness (EOT): gate dielectrics itself, \(C_{ox}\),

Capacitance Equivalent Thickness (CET): entire gate stack,

Metal gate can eliminate the poly-Si depletion.

Inversion CET = \(T_{inv}\) \(\approx\) EOT + 0.4nm with metal gate electrode.

K. Natori, et al., (Tsukuba Univ) SSDM 2005, p.286
Choice of High-k elements for oxide

<table>
<thead>
<tr>
<th>Candidates</th>
<th>Gas or liquid at 1000 K</th>
<th>Radio active</th>
</tr>
</thead>
<tbody>
<tr>
<td>H</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Li</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Be</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mg</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Na</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ca</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sc</td>
<td></td>
<td></td>
</tr>
<tr>
<td>K</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Y</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zr</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rh</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cs</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ba</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fr</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ra</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hf</td>
<td></td>
<td></td>
</tr>
<tr>
<td>La</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Unstable at Si interface:
1. \( Si + MO_x \rightarrow M + SiO_2 \)
2. \( Si + MO_x \rightarrow MSi_x + SiO_2 \)
3. \( Si + MO_x \rightarrow M + MSi_xO_y \)

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.
Issues in high-k/metal gate stack

Oxygen concentration control for prevention of EOT increase and oxygen vacancy formation in high-k

Suppression of gate leakage current

Endurance for high temperature process

Reliability: PBTI, NBTI, TDDB

Oxygen diffusion control for prevention of EOT increase and oxygen vacancy formation in high-k

Flat metal/high-k interface for better mobility

Suppression of metal diffusion

Suppression of oxygen vacancy formation

Small interfacial state density at high-k/Si

Control of interface reaction and Si diffusion to high-k

Workfunction engineering for $V_{th}$ control

Suppression of FLP

Interface dipole control for $V_{th}$ tuning

Remove contamination introduced by CVD

Thinning or removal of SiO$_2$-IL for small EOT

Oxygen concentration control for prevention of EOT increase and oxygen vacancy formation in high-k
Conduction band offset vs. Dielectric Constant

Leakage Current by Tunneling

Dielectric Constant $\varepsilon(0)$

Band Discontinuity [eV]

Si Band Gap

XPS measurement by Prof. T. Hattori, INFOS 2003
Direct high-k/Si by silicate reaction

HfO₂ case

La₂O₃ case

Our approach

Direct contact can be achieved with La₂O₃ by forming silicate at interface. Control of oxygen partial pressusre is the key for processing.

K. Kakushima, et al., VLSI2010, p.69
SiO_x-IL growth at HfO_2/Si Interface

XPS Si1s spectrum

500 °C

SiO_2
Hf Silicate
Si sub.

Binding energy (eV)

Intensity (a.u)

Phase separator

HfO_2 + Si + O_2 → HfO_2 + Si + 2O^* → HfO_2 + SiO_2

Oxygen supplied from W gate electrode

Oxygen control is required for optimizing the reaction

SiO_x-IL is formed after annealing

H. Shimizu, JJAP, 44, pp. 6131
D.J. Lichtenwalner, Trans. ECS 11, 319
La-Silicate Reaction at La$_2$O$_3$/Si

Direct contact high-k/Si is possible

La$_2$O$_3$ + Si + nO$_2$ → La$_2$SiO$_5$, La$_2$Si$_2$O$_7$, La$_{9.33}$Si$_6$O$_{26}$, La$_{10}$(SiO$_4$)$_6$O$_3$, etc.

La$_2$O$_3$ can achieve direct contact of high-k/Si
Cluster tool for HKMG Stack

- EB Deposition for HK
- Flash Lamp
- Sputter for MG
- Robot
- ALD
- Entrance
- RTA
Cluster Chambers for HKMG Gate Stack

- EB Deposition: HK
- Flash Lamp Anneal
- Sputter: MG
- ALD: HK
- Entrance
- Robot
- RTA
L=0.5~100 \mu m (8 kinds)

W=10, 20, 50, 100 \mu m (4 kinds)
Gate Leakage vs EOT, (Vg=|1|V)

Current density (A/cm²) vs EOT (nm)

Materials: Al₂O₃, HfAlO(N), HfO₂, HfSiO(N), HfTaO, La₂O₃, Nd₂O₃, Pr₂O₃, PrSiO, PrTiO, SiON/SiN, Sm₂O₃, SrTiO₃, Ta₂O₅, TiO₂, ZrO₂(N), ZrSiO, ZrAlO(N).
A fairly nice La-silicate/Si interface can be obtained with high temperature annealing. (800°C)
Physical mechanisms for small Dit

① silicate-reaction-formed fresh interface

② stress relaxation at interface by glass type structure of La silicate.

Fresh interface with silicate reaction

FGA800°C is necessary to reduce the interfacial stress


No interfacial layer can be confirmed with Si/TiN/W
La$_2$O$_3$/silicate/$n$-Si CV

W/La$_2$O$_3$(4nm)/$n$-Si
600°C, 30min

Capacitance density (µF/cm$^2$)

Gate voltage (V)
Annealing temperature (°C)

$D_{it}, D_{slow}$

$G_p/\omega$ (F/cm²)

$\omega$ (rad/s)

$10^{11}$ $10^{12}$ $10^{13}$ $10^{14}$

Annealing temperature (°C)

$D_{slow}$

$D_{it}$
It is important to change the La2O3 to La-silicate completely.
TiN(45nm)/W(6nm)

Annealed for 2 s
La$_2$O$_3$(3.5 nm)

EOT (nm)

Annealing temperature (°C)

EOT=0.55nm

TaN/(45nm)/W(3nm)
900°C, 30min

EOT=0.55nm

C$_g$ (uF/cm$^2$)

V$_g$ (V)

Experiment
Theory
TaN(45nm)/W(3nm)

\( Q_{\text{fix}} = 1 \times 10^{11} \text{ cm}^{-2} \)

900°C, 30min

Fixed Charge density: \( 1 \times 10^{11} \text{ cm}^{-2} \)
EOT = 0.53 nm

- L/W = 20/20 µm
- T = 300 K
- $N_{sub} = 3 \times 10^{16} \text{cm}^{-3}$

Electron Mobility [cm²/Vsec]

- Drain Voltage (V): 0 0.2 0.4 0.6 0.8 1.0
- Drain Current (µA): 0 50 100 150 200 250 300

- Vg = 0.2 V
- Vg = 0.4 V
- Vg = 0.6 V
- Vg = 0.8 V
- Vg = 1.0 V

- E_{eff} = 10 MV/cm
- $151 \text{cm}^2/\text{Vs}$
- $150 \text{cm}^2/\text{Vs}$
TaN/W/LaSiOₓ/nFET

$W_g/L_g = 20/20 \mu m$

EOT = 0.55nm
Recent results by my group.

L/W = 20/20μm
T = 300K
N_{sub} = 3 \times 10^{16} \text{cm}^{-3}
EOT = 0.53\text{nm}

Electron Mobility [cm^2/Vsec]
Benchmark of La-silicate dielectrics

Gate leakage is one order of magnitude lower than that of ITRS.

Electron mobility is comparable to record mobility with Hf-based oxides.


T. Ando, et al., (IBM) IEDM 2009, p.423
<table>
<thead>
<tr>
<th>Gate stack</th>
<th>EOT</th>
<th>Mobility</th>
<th>$V_{th}$</th>
<th>SS</th>
<th>DIBL</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>TiN/Cap/HfO$_2$</td>
<td>0.52nm</td>
<td>110cm$^2$/Vs</td>
<td>$\sim$0.4V</td>
<td>90mV/dec</td>
<td>147mV/V</td>
<td>IBM VLSI2011</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(at 1x10$^{13}$cm$^{-2}$)</td>
<td>(L$_g$=24nm)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TiN/Cap/HfO$_2$</td>
<td>0.55nm</td>
<td>140cm$^2$/Vs</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>IBM VLSI2009</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(at 1MV/cm)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TiN/Cap/HfO$_2$</td>
<td>0.45nm</td>
<td>115cm$^2$/Vs</td>
<td>0.3V</td>
<td>—</td>
<td>—</td>
<td>IMEC MEE2011</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(at 1x10$^{13}$cm$^{-2}$)</td>
<td>(L$_g$=10um)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Metal/HfO$_2$</td>
<td>0.59nm</td>
<td>130cm$^2$/Vs</td>
<td>0.45V</td>
<td>75mV/dec</td>
<td>—</td>
<td>Sematech VLSI2009</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(at 1MV/cm)</td>
<td>(L$_g$=1um)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Metal/Hf-based</td>
<td>0.65nm</td>
<td>—</td>
<td>0.3$\sim$0.4V</td>
<td>90mV/dec</td>
<td>100mV/V</td>
<td>Samsung VLSI2011</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(L$_g$=~30nm)</td>
<td>(L$_g$=~30nm)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Metal/Hf-based</td>
<td>0.95nm</td>
<td>—</td>
<td>$\sim$0.3V</td>
<td>100mV/dec</td>
<td>$\sim$200mV/V</td>
<td>Intel IEDM2009</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(L$_g$=30nm)</td>
<td>(L$_g$=30nm)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>W/La-silicate 92</td>
<td>0.62nm</td>
<td>155cm$^2$/Vs</td>
<td>$\sim$0.08V</td>
<td>$\sim$70mV/dec</td>
<td>—</td>
<td>Tokyo Tech. T-ED2012</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(at 1MV/cm)</td>
<td>(L$_g$=10um)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
ALD of La2O3


Precursor (ligand)

La(iPrCp)₃

La(FAMD)₃

ALD is indispensable from the manufacturing viewpoint

- precise control of film thickness and good uniformity

La gas feed

Ar purge

H₂O feed

Ar purge

1 cycle
Advantages of metal S/D
- atomically abrupt junction
- low parasitic resistance
- reduced channel dopant concentration

Issues in metal S/D
- two different $\phi_B$ for $p/n$-ch FETs
- underlap/overlap to the gate
- narrow process temperature window

Metal S/D is considered for alternative channel material such as InGaAs and Ge

Ni is used both on InGaAs and Ge to form alloy.

S.-H. Kim, IEDM (2010) 596
K. Ikeda, VLSI (2012) 165
4. Alternative channel devices
### Ge, III-V bulk properties

<table>
<thead>
<tr>
<th></th>
<th>Si</th>
<th>Ge</th>
<th>GaAs</th>
<th>InP</th>
<th>InAs</th>
<th>InSb</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>electron mob.</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(cm²/Vs)</td>
<td>1600</td>
<td>3900</td>
<td>9200</td>
<td>5400</td>
<td>40000</td>
<td>77000</td>
</tr>
<tr>
<td><strong>electron effective mass (/m₀)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mᵢ: 0.19 mᵢ: 0.916</td>
<td>0.067</td>
<td>0.08</td>
<td>0.026</td>
<td>0.0135</td>
<td></td>
<td></td>
</tr>
<tr>
<td>mᵢ: 0.82 mᵢ: 1.467</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>hole mob.</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(cm²/Vs)</td>
<td>430</td>
<td>1900</td>
<td>400</td>
<td>200</td>
<td>500</td>
<td>850</td>
</tr>
<tr>
<td><strong>hole effective mass (/m₀)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mᵢ₡: 0.49 mᵢ₡: 0.16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mᵢ₡: 0.28 mᵢ₡: 0.044</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mᵢ₨: 0.45 mᵨₙ: 0.082</td>
<td>mᵢ₨: 0.45 mᵨₙ: 0.12</td>
<td>mᵨₙ: 0.57 mᵨ₉: 0.35</td>
<td>mᵨ₉: 0.44 mᵨ₉: 0.016</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>band gap (eV)</strong></td>
<td>1.12</td>
<td>0.66</td>
<td>1.42</td>
<td>1.34</td>
<td>0.36</td>
<td>0.14</td>
</tr>
<tr>
<td><strong>permittivity</strong></td>
<td>11.8</td>
<td>16</td>
<td>12</td>
<td>12.6</td>
<td>14.8</td>
<td>17</td>
</tr>
</tbody>
</table>

- Ge ⇒ lightest hole $m^*$ (light electron $m^*$) ⇒ pMOS (CMOS)
- III-V ⇒ light electron $m^*$ ⇒ nMOS
- GaAs, InP ⇒ $E_g$ higher than that in Si ⇒ low power
Multi-gate III-V and Si benchmark

**nMOS**

- **InGaAs GAA**
  - $L_{ch}=50\text{nm}$, Dielectric: $10\text{nm} \text{Al}_2\text{O}_3$
  - $V_{DS}=0.5\text{V}$ (Purdue Uni.) [1]

- **InGaAs Nanowire**
  - $L_g=200\text{nm}$, $T_{ox}=14.8\text{nm}$
  - $V_{DS}=0.5\text{V}$ (Hokkaido Uni.) [4]

- **InGaAs FinFET**
  - $L_{ch}=130\text{nm}$
  - $EOT=3.8\text{nm}$
  - $V_{DS}=0.5\text{V}$ (NUS) [3]

- **InGaAs Tri-gate**
  - $L_g=60\text{nm}$, $EOT=12\text{A}$
  - $V_{DS}=0.5\text{V}$ (Intel) [2]

- **Si-FinFET 32nm**
  - $L_{ch}=50\text{nm}$
  - $EOT=3.5\text{nm}$
  - $V_{DS}=0.5\text{V}$ (Tokyo Uni.) [5]

- **Si-FinFET 22nm**
  - $L_{ch}=55\text{nm}$
  - $EOT=3.5\text{nm}$
  - $V_{DS}=0.5\text{V}$ (Intel) [10]

- **Si-bulk 45nm**
  - $L_{ch}=55\text{nm}$
  - $EOT=3.5\text{nm}$
  - $V_{DS}=0.5\text{V}$ (Intel) [11]

**pMOS**

- **GOI Tri-gate**
  - $L_g=65\text{nm}$, $EOT=3.0\text{nm}$
  - $V_D=-1\text{V}$ (AIST Tsukuba) [6]

- **Ge FinFET**
  - $L_g=4.5\text{mm}$
  - Dielectric: $\text{SiON}$, $V_{DS}=-1\text{V}$
  - (Stanford Uni.) [7]

- **Si-FinFET 32nm**
  - $L_{ch}=50\text{nm}$
  - $EOT=3.5\text{nm}$
  - $V_{DS}=0.5\text{V}$ (Purdue Uni.) [1]

- **Si-FinFET 22nm**
  - $L_{ch}=55\text{nm}$
  - $EOT=3.5\text{nm}$
  - $V_{DS}=0.5\text{V}$ (Intel) [10]

- **Si-bulk 45nm**
  - $L_{ch}=55\text{nm}$
  - $EOT=3.5\text{nm}$
  - $V_{DS}=0.5\text{V}$ (Intel) [11]

- **Ge GAA**
  - $L_g=300\text{nm}$
  - Dielectric: $\text{GeO}_2(7\text{nm})-\text{HfO}_2(10\text{nm})$
  - $V_D=-0.8\text{V}$ (ASTAR Singapore) [8]

- **Ge Tri-gate**
  - $L_g=183\text{nm}$, $EOT=5.5\text{nm}$
  - $V_D=-1\text{V}$ (NNDL Taiwan) [9]

---

ION/IOFF Benchmark of Ge pMOSFET

K. J. Kuhn, ECS Transactions, 33 (6) 3-17 (2010)

Optimization of short channel Ge p-MOSFETs is still under investigation
## III-V/Ge benchmark for various structures

<table>
<thead>
<tr>
<th>Material</th>
<th>Planar (metal S/D, Strain, Buffer...)</th>
<th>FinFET</th>
<th>Tri-gate</th>
<th>Gate-all-around MOSFET</th>
<th>Nanowire</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mobility</strong></td>
<td>-</td>
<td>~600 cm²/Vs</td>
<td>~700 μS/μm</td>
<td>-</td>
<td>701 μS/μm</td>
</tr>
<tr>
<td><strong>Lch (nm)</strong></td>
<td>55</td>
<td>50 μm</td>
<td>100 4.5 μm</td>
<td>50</td>
<td>200</td>
</tr>
<tr>
<td><strong>DIBL (mV/V)</strong></td>
<td>84</td>
<td>180</td>
<td>~50</td>
<td>210</td>
<td>-</td>
</tr>
<tr>
<td><strong>SS (mV/dec)</strong></td>
<td>105</td>
<td>145</td>
<td>90</td>
<td>130</td>
<td>150 160</td>
</tr>
<tr>
<td><strong>I\text{ON} (μA/μm)</strong></td>
<td>278 (V\text{D}=0.5V)</td>
<td>3 (V\text{D}=-0.2V)</td>
<td>4 (n,p) (V\text{D}=0.5V)</td>
<td>10 (V\text{D}=0.5V)</td>
<td>180 (V\text{D}=0.5V)</td>
</tr>
</tbody>
</table>

### Dielectric /EOT
- **Dielectric**: Al₂O₃ / HfO₂ + Al₂O₃ + GeO₂
- **EOT**: 3.5 nm
- **Material**: InGaAs, Ge, InGaSn, InGaAs, Ge

### Characteristics
- **Mobility**: ~600 cm²/Vs
- **Lch**: 55 nm
- **W/L**: 30/5 μm
- **DIBL (mV/V)**: 84
- **SS (mV/dec)**: 105
- **I\text{ON} (μA/μm)**: 278 (V\text{D}=0.5V)

### Devices
- **FinFET**: SiON 1.2 nm
- **Tri-gate**: 5.5 nm (Al₂O₃ + GeO₂)
- **Gate-all-around MOSFET**: HfO₂: 11nm
- **Nanowire**: HfAlO 14.8 nm

### Additional Information
InGaSb as channel material (stanford)

Z. Yuan et al., pp.185, VLSI2012 (Stanford Uni)

Achieving both N- and P-type MOSFET on a single channel is possible

- In-content of 20-40% improves performance
- electron/hole mobility > 4000/900cm²/Vs was gained in a single channel material

\[ I_{ON} \text{ at } L_G = 50 \mu m \quad pMOS: 4 \mu A/\mu m \quad nMOS: 3.8 \mu A/\mu m \]
Metal S/D and InAs buffer layer are used as performance boosters.

- DIBL=84 mV/V and SS=105 mV/V was shown for $L_{ch} = 55$ nm when In-content was higher.

S. H. Kim et al., pp.177, VLSI2012 (Tokyo Uni)
Common InGaAs-GeSn gate stack (NUS)

- Common gate stack (gate metal and dielectric) were used for both p- and n-type
- Si$_2$H$_6$ plasma passivation is employed which creates Si layer at interface.
  - SS: nMOS: 90 (mV/decade)
  - pMOS: 190 (mV/decade)
- High intrinsic peak $G_{M,\text{Sat}}$ of ~465 $\mu$S/$\mu$m at $V_{DS}=-1.1$ V was achieved for $L_G = 250$ nm.

V$_{GS}$-V$_{TH}$ = 0~2.0V
$L_G = 5\mu$m

Core-multishell InGaAs nanowires grown without buffer layer on Si substrate (bottom up approach)

At $V_d = 1$ V peak transconductance of 500 mS/mm is achieved (roughly $x3$ InGaAs nanowire)
Tri-gate InGaAs QW-FET (Intel)


- Tri-gate structure has superiority electrostatic controllability compared to ultra-thin body planar structure

- Steepest SS and smallest DIBL ever reported ($W_{\text{fin}} = 30\text{nm}$)
Gate all around InGaAs MOSFET (Purdue)


- Inversion mode In$_{0.53}$Ga$_{0.47}$As MOSFET with ALD Al$_2$O$_3$/WN with well electrostatic properties

- DIBL was suppressed down to $L_{ch} = 50\text{nm}$ and
  
  $G_{m,max} = 701\text{mS/mm at } V_{ds} = 1\text{V}$
DIBL = 135 mV/V and drive current over 840 µA/µm at $L_{ch} = 130$nm and $V_{ds} = 1.5$V was achieved.
Ge-nanowire pMOSFET (AIST, Tsukuba)

Using Ni-Ge alloy as metal S/D

Significantly reduces contact resistance

High saturation current and high mobility

\[ \mu_{\text{eff}} = 855 \text{ cm}^2/\text{Vs} \] at \( N_s = 5 \times 10^{12} \text{cm}^{-2} \)

and saturation drain current of \( 731 \mu\text{A}/\mu\text{m at } V_d = -1\text{V} \)
Selective etching of high defect Ge near Ge/Si interface is used which improves gate controllability.

- $I_{ON}/I_{OFF} = 10^5$ and $SS = 130$ mV/dec
- $I_{ON} = 235 \mu m/\mu m$ at $V_D = -1V$
Implementing high-k material to III-V, Ge

III-V (InGaAs, InAs, InGaSb, …)

- ALD-Al₂O₃ is most commonly used as gate dielectric in planar or Multi-gate
- HfO₂-only stacks have high D_it (combination of Al₂O₃ or Al or Si is used)

- In₀.₅₃Ga₀.₄₇As
- In₀.₇Ga₀.₃As
- Ge

Ge

- By controlling the formation of GeOₓ at the interface, HfO₂ and Al₂O₃ show good results.
5. Emerging devices
Emerging devices (future scaling trends)

Carbon-based FET

Carbon nanotube

Graphene

Junctionless Transistor

GaAs mHEMT

SiMOSFET

GaAs pHEMT

CNT

Graphene

All-spin logic device

Input and output related via Spin-coherent channel


J. P. Colinge et al., Nature Nano. 5(2010)225

J. P. Colinge et al., Nature Nano. 5(2010)266


A. D. Franklin et al., pp.525, IEDM2011 (IBM)
Tunnel FET

Low $I_{OFF}$, Low $V_{DD}$, $SS<60mV/\text{decade}$
TFET vs. MOSFET at low $V_{DD}$

A. Seabaugh, IEDM 2011 Short Course (University of Notre Dame)

EOT = 1 nm  \( \Rightarrow \)  \( L_G = 20 \text{ nm} \)

\( P^+ \) Undoped \( \leftrightarrow \) \( N^+ \)

\( t_{lnAs} = 5 \text{ nm} \)

![TFET Graph](image)

<table>
<thead>
<tr>
<th>TFET @0.35V</th>
<th>Stage Delay (ps)</th>
<th>Switching Energy (fJ)</th>
<th>Leakage Power (nW)</th>
</tr>
</thead>
<tbody>
<tr>
<td>INV (FO=3)</td>
<td>105</td>
<td>0.26</td>
<td>0.37</td>
</tr>
<tr>
<td>NAND (FO=3)</td>
<td>149</td>
<td>0.32</td>
<td>0.57</td>
</tr>
<tr>
<td>NOR (FO=3)</td>
<td>158</td>
<td>0.34</td>
<td>0.54</td>
</tr>
<tr>
<td>Combined</td>
<td>129</td>
<td>0.29</td>
<td>0.46</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CMOS @0.3V</th>
<th>Stage Delay (ps)</th>
<th>Switching Energy (fJ)</th>
<th>Leakage Power (nW)</th>
</tr>
</thead>
<tbody>
<tr>
<td>INV (FO=3)</td>
<td>830</td>
<td>0.25</td>
<td>0.37</td>
</tr>
<tr>
<td>NAND (FO=3)</td>
<td>1364</td>
<td>0.32</td>
<td>0.56</td>
</tr>
<tr>
<td>NOR (FO=3)</td>
<td>1389</td>
<td>0.33</td>
<td>0.53</td>
</tr>
<tr>
<td>Combined</td>
<td>1103</td>
<td>0.29</td>
<td>0.46</td>
</tr>
</tbody>
</table>

$V_{DD}$ 0.3~0.35V

TFET 8x faster at the same power

“parameter variation is not a significant factor for differentiation between MOSFET and TFET”
X in $\text{Si}_{1-x}\text{Ge}_x$ is optimized to allow for efficient BTBT.

$V_G = 200\, \text{nm}$

$I_{\text{ON}} / I_{\text{OFF}} \approx 10^5$

Reducing SiGe Body thickness improves Subthreshold swing.

- 130mV/dec
- 190mV/dec
SS of TFET is function of $V_G$ due to Zener tunnel current

Minimum SS = 21 mV/dec is reached due to optimized series resistance of contact, undoped InAs and InAs/Si

$I_{ON}/I_{OFF} \sim 10^6$ at $V_{DS} = 1.0V$ ($I_{ON} = 1\mu A/\mu m$)
Device structure

A. Seabaugh, IEDM 2011 Short Course (University of Notre Dame)

- p+AlGaSb 40 nm
- Al₂O₃ 6 nm
- n+InAs 30 nm

- Tunneling // to the gate
- Oblique to the gate field

- Tunneling ⊥ to the gate
- In-line with the gate field

- Nanowire TFET
- Gate-all-around - best electrostatics

K. Tomioka et al., pp.47, VLSI2012 (Hokkaido University)
## Tunnel FET performance comparison

A. Seabaugh, IEDM 2011 Short Course (University of Notre Dame)

Measured III-V channel TFETs

<table>
<thead>
<tr>
<th>Reference</th>
<th>Source</th>
<th>Channel</th>
<th>Dielectric</th>
<th>EOT nm</th>
<th>$I_{ON}$ $\mu A/\mu m$</th>
<th>$V_{DS}$ V</th>
<th>$V_{GS}=V_{ON}$ V</th>
<th>$V_{ON-V_{OFF}}$ V</th>
<th>$I_{ON}/I_{OFF}$</th>
<th>$S_{MIN}$ mV/dec</th>
<th>$S_{EFF}$ mV/dec</th>
</tr>
</thead>
<tbody>
<tr>
<td>this work</td>
<td>Al$<em>{0.45}$Ga$</em>{0.55}$Sb</td>
<td>InAs</td>
<td>Al$_2$O$_3$/HfO$_2$</td>
<td>1.6</td>
<td>78</td>
<td>0.5</td>
<td>0.5</td>
<td>1.5</td>
<td>2,000</td>
<td>125</td>
<td>230</td>
</tr>
<tr>
<td>Li [4]</td>
<td>Al$<em>{0.45}$Ga$</em>{0.55}$Sb</td>
<td>InAs</td>
<td>Al$_2$O$_3$</td>
<td>3.9</td>
<td>20</td>
<td>0.3</td>
<td>2</td>
<td>6</td>
<td>210</td>
<td>830</td>
<td>1300</td>
</tr>
<tr>
<td>Mookerjea [5]</td>
<td>In$<em>{0.53}$Ga$</em>{0.47}$As</td>
<td>In$<em>{0.53}$Ga$</em>{0.47}$As</td>
<td>Al$_2$O$_3$</td>
<td>5.6</td>
<td>10</td>
<td>0.5</td>
<td>2.5</td>
<td>2.5</td>
<td>10,000</td>
<td>240</td>
<td>310</td>
</tr>
<tr>
<td>Mohata [6]</td>
<td>In$<em>{0.53}$Ga$</em>{0.47}$As</td>
<td>In$<em>{0.53}$Ga$</em>{0.47}$As</td>
<td>Al$_2$O$_3$</td>
<td>2.8</td>
<td>10</td>
<td>0.5</td>
<td>2</td>
<td>2</td>
<td>400,000</td>
<td>200</td>
<td>180</td>
</tr>
<tr>
<td>Zhao [7]</td>
<td>In$<em>{0.7}$Ga$</em>{0.3}$As</td>
<td>In$<em>{0.7}$Ga$</em>{0.3}$As</td>
<td>HfO$_2$</td>
<td>1.2</td>
<td>20</td>
<td>0.45</td>
<td>1</td>
<td>1.5</td>
<td>30,000</td>
<td>86</td>
<td>170</td>
</tr>
<tr>
<td>Zhao [8]</td>
<td>In$<em>{0.7}$Ga$</em>{0.3}$As</td>
<td>In$<em>{0.7}$Ga$</em>{0.3}$As</td>
<td>HfO$_2$</td>
<td>1.2</td>
<td>40</td>
<td>0.5</td>
<td>2</td>
<td>2</td>
<td>200,000</td>
<td>84</td>
<td>190</td>
</tr>
<tr>
<td>Ford [9]</td>
<td>InAs</td>
<td>InAs</td>
<td>ZrO$_2$</td>
<td>2</td>
<td>0.4</td>
<td>0.5</td>
<td>1</td>
<td>1</td>
<td>200</td>
<td>190</td>
<td>220</td>
</tr>
<tr>
<td>Zhou [10]</td>
<td>InP</td>
<td>In$<em>{0.53}$Ga$</em>{0.47}$,In$_{0.47}$</td>
<td>Al$_2$O$_3$/HfO$_2$</td>
<td>1.3</td>
<td>20</td>
<td>0.5</td>
<td>1</td>
<td>1.75</td>
<td>450,000</td>
<td>93</td>
<td>150</td>
</tr>
</tbody>
</table>

**$S_{MIN}$**: Most common SS which is the inverse of $I_D$-$V_{GS}$ slope at the steepest part

**$S_{EFF}$**: Is the average swing when $V_{TH}=V_{DD}/2$, $V_{OFF}=0$

### Average SS:

$$\text{Average SS: } \frac{V_{TH} - V_{OFF}}{\log(I_D/I_{OFF})}$$

### Effective SS:

$$\text{Effective SS: } \frac{V_{DD}}{2\log(I_D/I_{OFF})}$$
$I_{ON}$ and $I_{OFF}$ of TFETs

[1] G. Zhou et al., pp. 782, vol. 33, no. 6, EDL 2012 (University of Notre Dame)

C. Auth et al., pp.131, VLSI2012 (Intel).
K. Mistry et al., pp.247, IEDM2007 (Intel).
MEMS relay

- Frequency of 1, 5, 25 kHz under operation
- $\frac{I_{ON}}{I_{OFF}}$ of $\sim 10^{10}$
- Ultra-low-power digital logic applications.

T. K. Liu et al., pp. 43, VLSI2012 (UC Berkeley)
Junctionless Accumulation Mode (JAM) devices have reduced gate control and degraded short-channel characteristics relative to Conventional Inversion Mode (IM).

- JAM devices are not suitable for high-performance logic due to high $I_{on}$ and moderate $I_{off}$.

**Figure Description:**
- IM: Conventional Inversion Mode
- JAM LD: Junctionless Accumulation Mode with low dope
- JAM HD: Junctionless Accumulation Mode with high dope

**Figure Details:**
- **Vt (V):** Voltage at which current starts to flow.
- **DIBL (mV/V):** Drain-Induced Barrier Lowering.
- **Idsat (mA/µm):** Saturation Current Density.
Near-ideal subthreshold slope, close to 60 mV/dec at room temperature, and extremely low leakage currents.

- Silicon nanowire is uniformly doped
- Gate material is opposite polarity polysilicon

\[ \frac{I_{ON}}{I_{OFF}} \approx 1 \times 10^6 \quad (-1 < V_g < 1) \]

\[ I_{OFF} \text{ is smaller than } 10^{-15} \text{ A} \]
**Carbon nanotube and Graphene**

K. Banerjee, UC Santa Barbara, G-COE PICE International Symposium on Silicon Nano Devices in 2030

SWCNT : single wall carbon nanotube  
GNR : graphene nano ribbon

<table>
<thead>
<tr>
<th></th>
<th>Si</th>
<th>SWCNT</th>
<th>GNR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mobility cm²/V/s</td>
<td>1400</td>
<td>&gt;10000</td>
<td>&gt;10000 for large widths</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>200 for W&lt;10nm</td>
</tr>
<tr>
<td>Mean free path (nm) @ room temp</td>
<td>30</td>
<td>&gt;1,000</td>
<td>~1,000</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>~10nm for W&lt;10nm</td>
</tr>
<tr>
<td>Thermal conductivity (W/mK)</td>
<td>150</td>
<td>5800</td>
<td>~5000</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Balandin, et al., Nano Let., 2008</td>
</tr>
</tbody>
</table>

Carbon materials for FET applications
- an ultra-thin body for aggressive channel length scaling
- excellent intrinsic transport properties similar to carbon nanotubes
- pattern the desired device structures
Sub-10nm carbon nanotube transistor

Transistor operation with $L_{ch}$ of 9nm
Graphene Field-effect Transistor

- Ambipolar Characteristics
- Bi-layer graphene and double gates can open the gap
2D material: single layer MoS$_2$

- Mobility of 190 cm$^2$/V·sec
- $I_{on}$ of 1 μA/μm at $V_{DD} = 1$V

Candidate: MoS$_2$, MoSe$_2$, WS$_2$, WSe$_2$, MoTe$_2$, WTe$_2$
Read/write are enabled by using ferromagnetic electrodes and Spin-polarized current
<table>
<thead>
<tr>
<th>Emerging Technology</th>
<th>Advantage</th>
<th>Issues</th>
</tr>
</thead>
<tbody>
<tr>
<td>TFET</td>
<td>Lower $V_{dd}$, Lower $I_{OFF}$</td>
<td>Integration, higher $I_{ON}$</td>
</tr>
<tr>
<td>CNT FET</td>
<td>Higher transport velocity, $L_g$ scaling</td>
<td>High density and alignment, reproducibility, integration</td>
</tr>
<tr>
<td>Graphene FET</td>
<td>RF application, Large area manufacturing</td>
<td>NOT a direct replacement for Silicon logic</td>
</tr>
<tr>
<td>MEMS</td>
<td>Extremely low leakage, Ultra-low digital logic</td>
<td>Endurance, Slow speed, scalability</td>
</tr>
<tr>
<td>Junctionless FET</td>
<td>CMOS process compatibility</td>
<td>Worse gate control in short-channel</td>
</tr>
<tr>
<td>Spin FET</td>
<td>Low power, suitable for memory (nonvolatile info storage)</td>
<td>Low efficiency of spin injection</td>
</tr>
</tbody>
</table>
Conclusions
New device structures (FinFET, Tri-gate) are replacing conventional Planar CMOS. Same performance at lower supply voltage

HKMG: Continuous innovation has enabled EOT scaling to 9 A°, however, new material could be needed for further EOT scaling. La-based high-k material

Recent advances in new channel material shows promising device performances but still far to catch up Si-CMOS. The combination of III-V channel materials with a multi-gate structure appears to be a promising direction. (Higher performance in lower operating voltage)

Device demonstration on emerging technologies (such as Tunnel FET, Junctionless FET, Carbon-based FET..) is increasing, But more time is needed for implementation of these technologies in future generation devices as mature technologies.