Lecture 33 - The "Short" Metal-Oxide-Semiconductor Field-Effect Transistor *(cont.)*

April 30, 2007

Contents:

Key questions

- How has MOSFET scaling been taking place?
- Are there fundamental limits to MOSFET scaling?
- How far will MOSFET scaling go?

1. Scaling *(cont.)*

Scaling goal: *extract maximum performance from each generation* (maximize I_{on}), for a given amount of:

- short-channel effects (DIBL), *and*
- off-current

To preserve *electrostatic integrity*, scaling has proceeded in a harmonious way: L ($\downarrow$), W ($\downarrow$), x_{ox} ($\downarrow$), N_A ($\uparrow$), x_j ($\downarrow$), and V_{DD} ($\downarrow$).

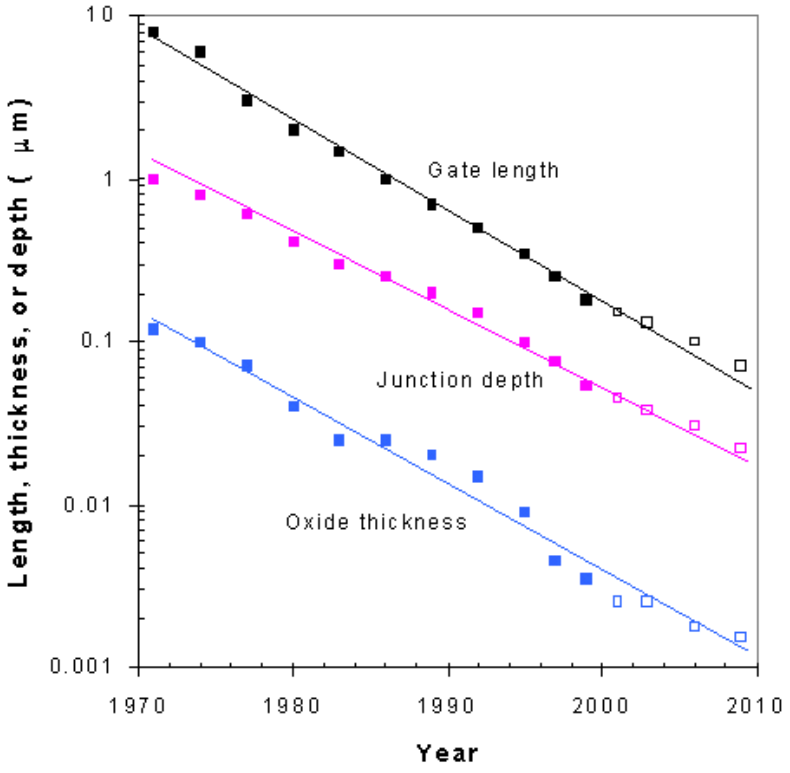
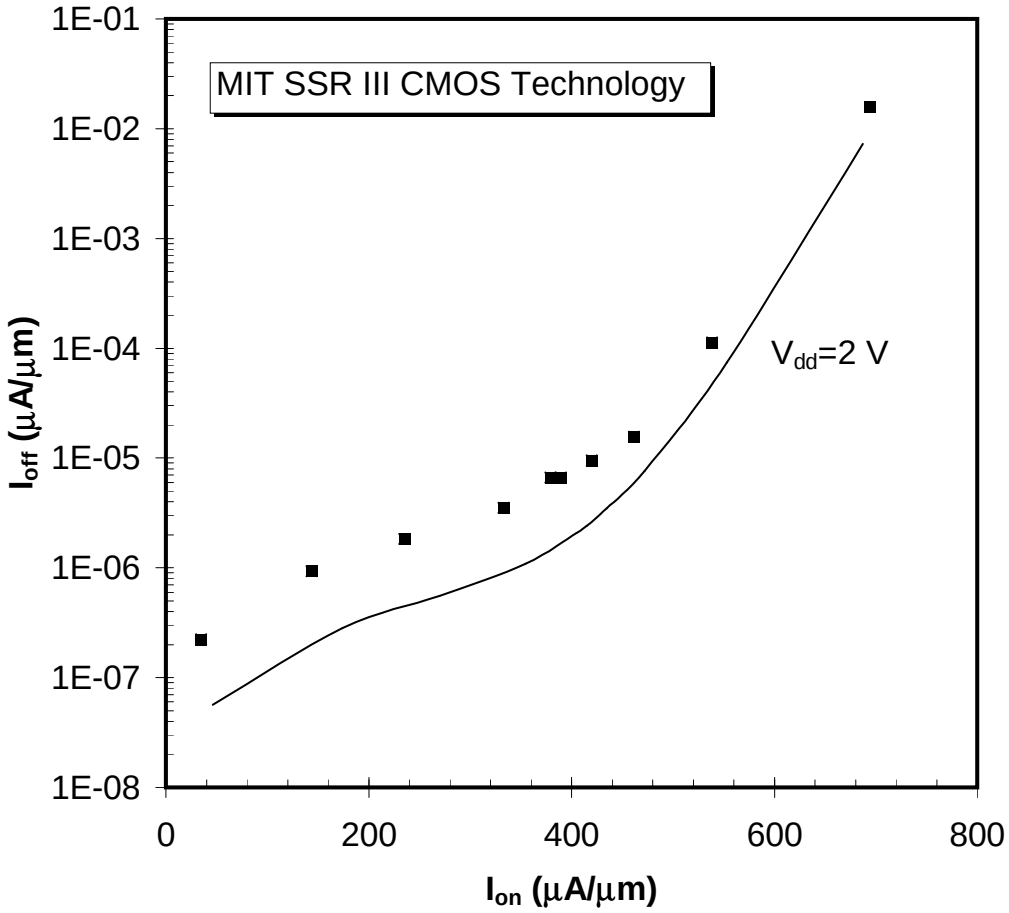
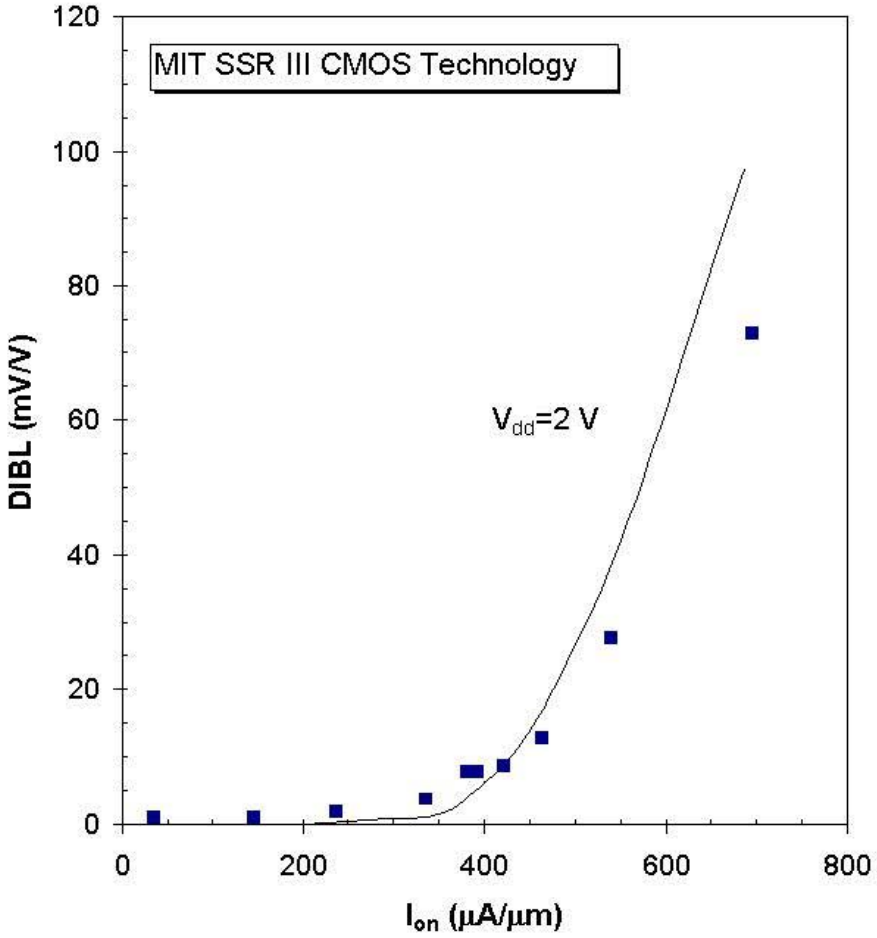


Illustration of key trade-offs:

- I_{on} vs. I_{off}



- I_{on} vs. $DIBL$



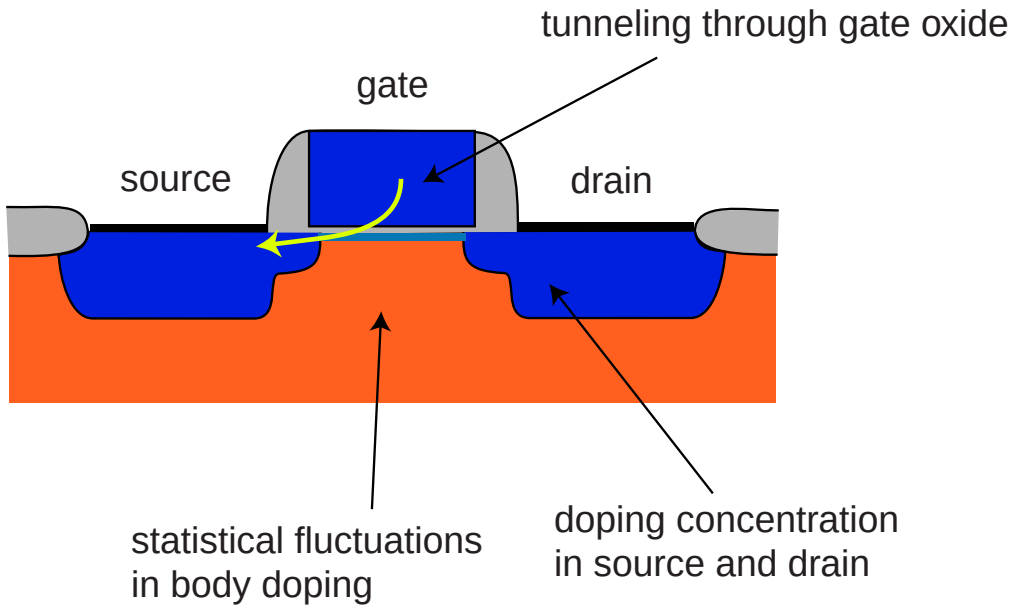
□ Limits to scaling

Text removed due to copyright restrictions.
Markoff, John. "Chip Progress Forecast to Hit a Big Barrier."
The New York Times (October 9, 1999).

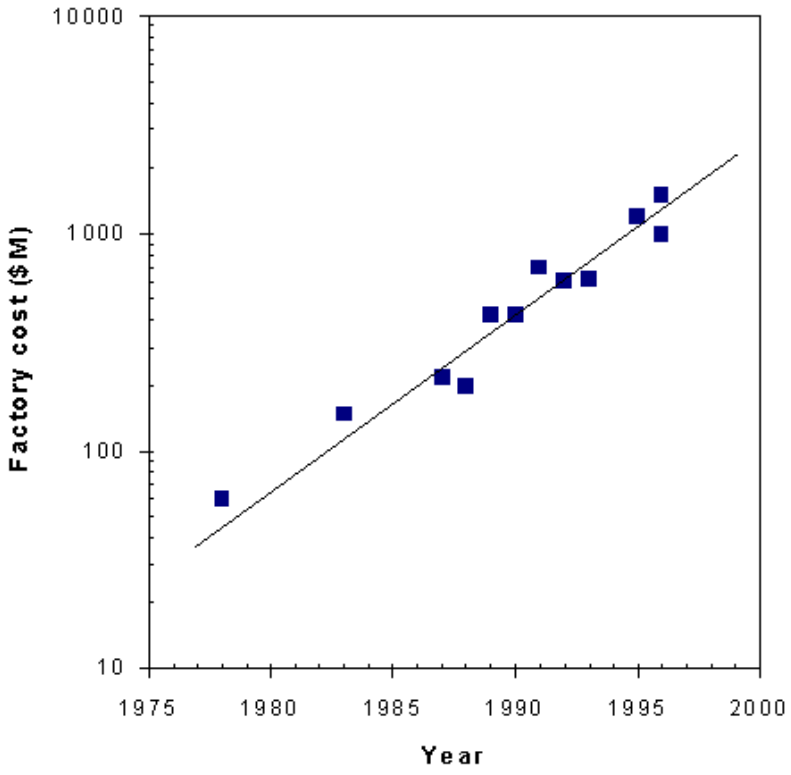
The New York Times (Oct. 9, 1999)

Four kinds of limits:

- Thermodynamics: doping concentration in source and drain
- Physics: tunneling through gate oxide
- Statistics: statistical fluctuation of body doping
- Economics: factory cost



□ Economics: factory cost also follows Moore's law!



New factories cost well in excess of $1B!

□ Physics: tunneling through gate oxide (most severe limit)

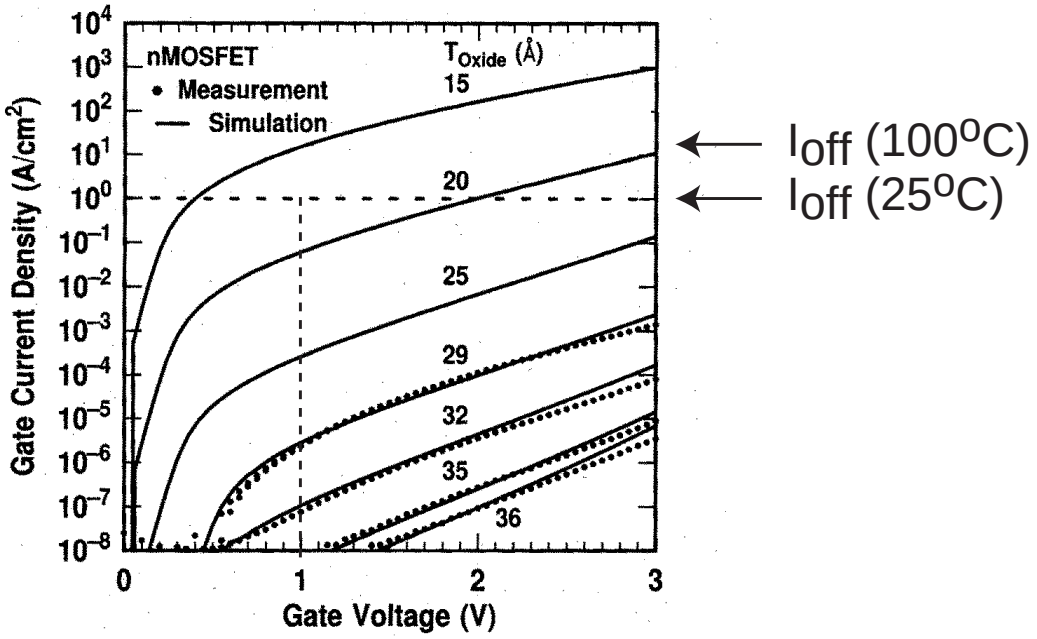


Fig. 2. Measured and simulated I_G–V_G characteristics under inversion conditions of four nMOSFET's. The dotted line indicates the 1 A/cm^2 limit for leakage current as discussed in the text.

Figure 13 on p. 491 in: Taur, Y., et al. "CMOS Scaling into the Nanometer Regime." *Proceedings of the IEEE* 85, no. 4 (1997): 486-504. © 1997 IEEE.

- Oxide's thickness limit when:

$$I_{gate} \simeq I_{off} \ @ \ V_{DD} \simeq 1 \ V, \ T_{oper} (\simeq 100^o C)$$

- Translates to limiting gate current:

$$I_{gate}(25^o C) \simeq 100 \ pA$$

- Limiting gate current density:

$$A \simeq 0.1 \ \mu m \times 0.1 \ \mu m = 10^{-10} \ cm^2 \ \Rightarrow \ J_{gate}(25^o C) \simeq 1 \ A/cm^2$$

- Limiting $x_{ox} \simeq 1.6 \ nm \ \Rightarrow \ L \sim 35 - 50 \ nm$
- Solution: *high-dielectric constant gate insulator*

Current wisdom for limiting bulk CMOS (with nitrided gate oxides):

$x_{ox} \simeq 1.2\ nm \ \Rightarrow \ L_{eff} \sim 25 - 35\ nm$

but... unclear if industry will do it (there are better options).

□ What does this mean?

Arno Penzias [1997]: *"We can look forward to a million-fold increase in the power of microelectronics".*

10X transistor size reduction ⇒

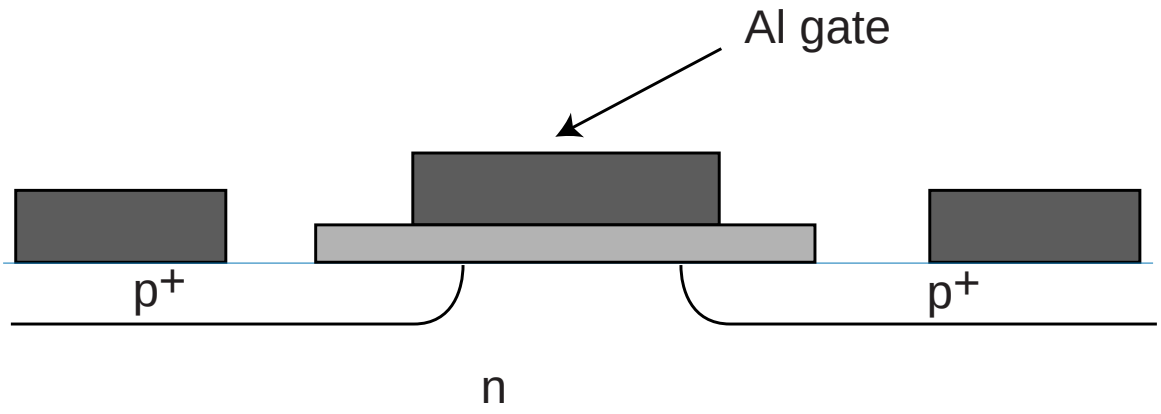
⇒ 100X device density
⇒ 100X circuit speed
⇒ 100X surprise
⇒ 10^6X TOTAL

□ To go beyond this, need:

- new materials that squeeze more performance out of existing device architecture
  - new channel materials: strained Si, Si/SiGe heterostructores
  - new gate insulators: high-K dielectric, such as HfO
  - new gate conductors: metal gate, such fully silicided gate
- new device architecture (SOI, double gate, trigate) to improve electrostatic integrity

2. Evolution of MOSFET design

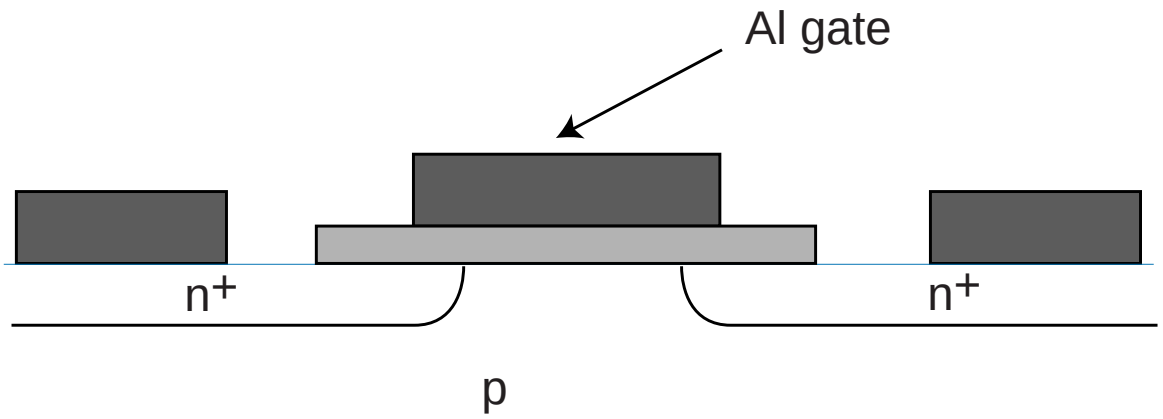
• PMOS with metal gate:



circa∼early 70's
$L \sim 20\ \mu m$
$x_{ox} \sim 1000\ \mathring{A}$
$x_j \sim 3\ \mu m$
$V_{DD} = 12\ V$

Main point: Na^+ contamination made NMOS devices to have too negative a threshold voltage

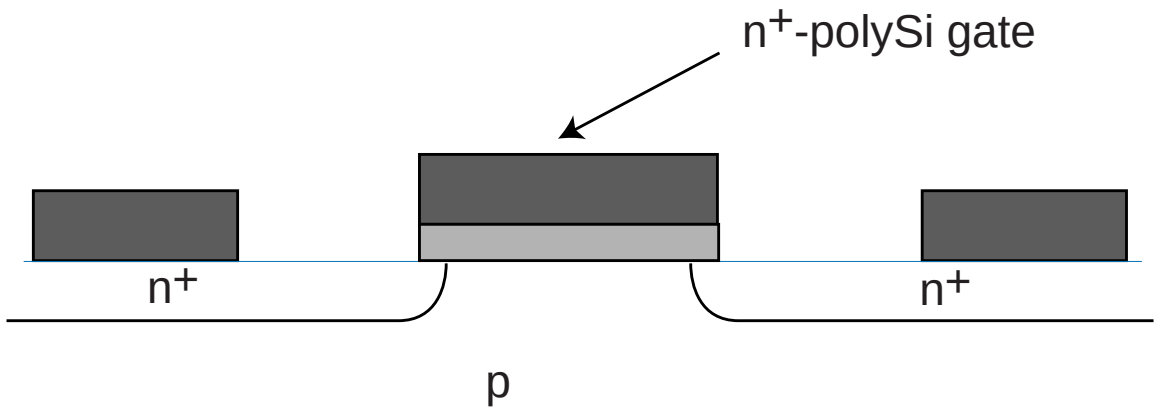
- NMOS with metal gate:



circa~1975
$L \sim 15\ \mu m$
$x_{ox} \sim 600\ \mathring{A}$
$x_j \sim 2\ \mu m$
$V_{DD} = 12\ V$

Main point: with Na^+ contamination under control, NMOS devices became possible (higher performance).

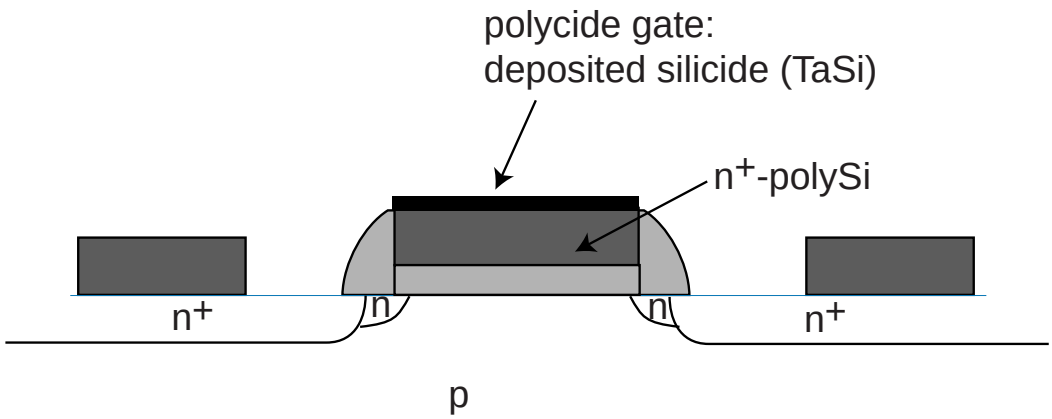
• CMOS with self-aligned polySi gate:



circa∼1980
$L \sim 2\ \mu m$
$x_{ox} \sim 400\ \mathring{A}$
$x_j \sim 1\ \mu m$
$V_{DD} = 5\ V$

Main point: self-aligned process allows tighter overlap between gate and n^+ regions and results in lower parasitic capacitance.

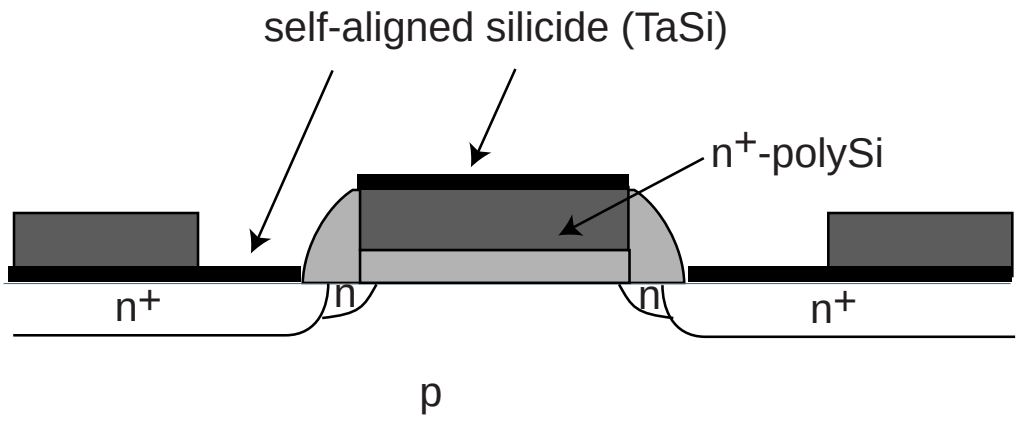
• Lightly-doped drain MOSFET (LDD-MOSFET):



circa∼1985
$L \sim 0.75\ \mu m$
$x_{ox} \sim 200\ \mathring{A}$
$x_j \sim 0.2\ \mu m$
$V_{DD} = 5\ V$

Main point: lightly-doped n-region on drain side reduces electric field there and allows a high V_{DD} to be used.

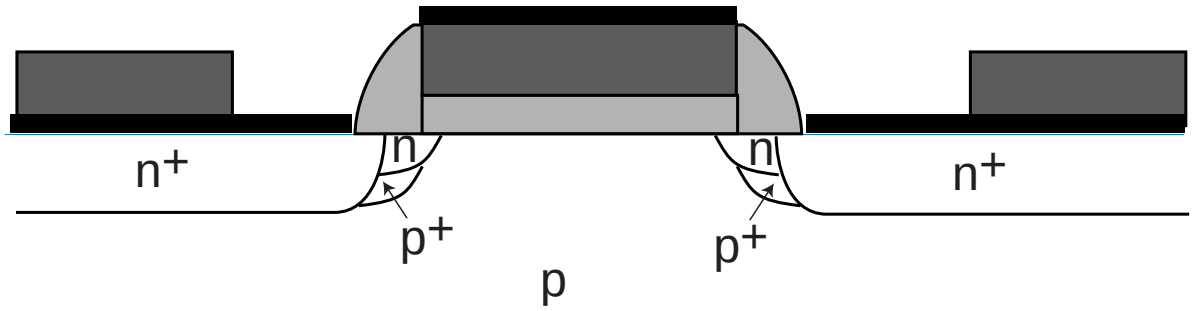
• Salicide (self-aligned silicide) MOSFET:



circa∼1989
$L \sim 0.4\ \mu m$
$x_{ox} \sim 125\ \mathring{A}$
$x_j \sim 0.15\ \mu m$
$V_{DD} = 3.3\ V$

Main point: salicided gate, source and drain reduces all parasitic resistances.

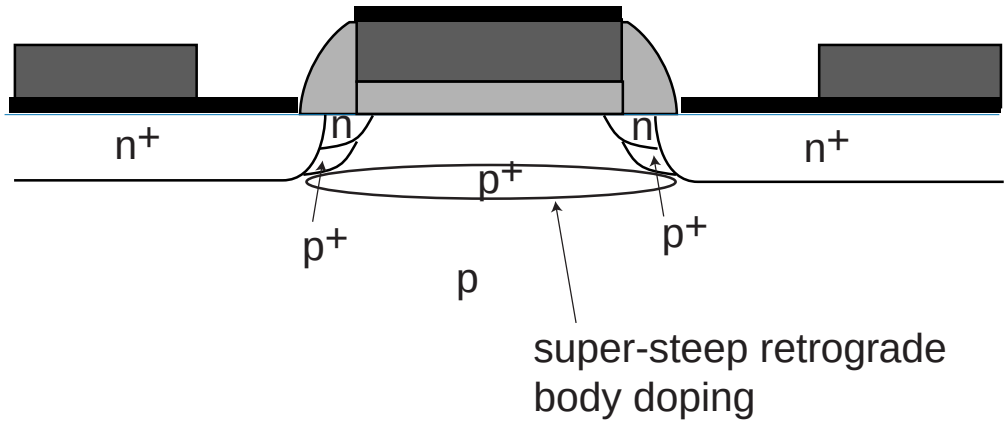
• MOSFET with p-pocket or halo implants:



circa∼1994
$L \sim 0.15\ \mu m$
$x_{ox} \sim 60\ \mathring{A}$
$x_j \sim 0.08\ \mu m$
$V_{DD} = 2.5\ V$

Main point: p^+ pockets control short-channel effects.

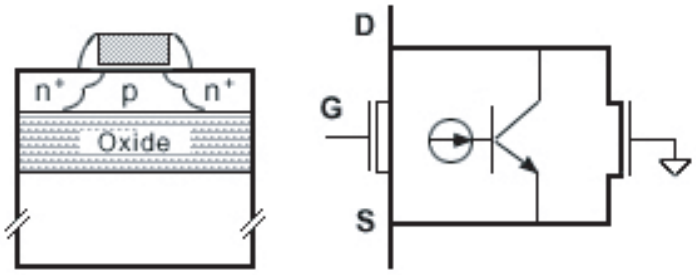
• Sub-0.1 μm MOSFET:



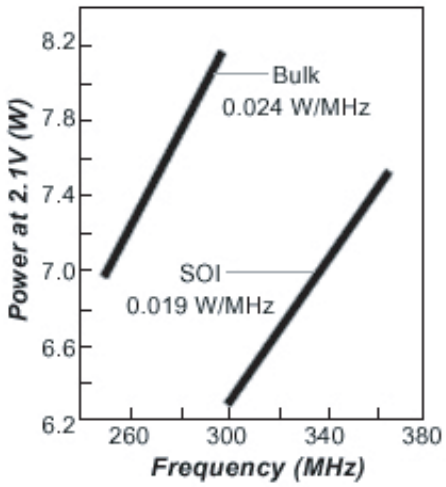
circa~late 90's (manufacturing in early 00's)
$L < 0.1\ \mu m$
$x_{ox} \sim 30\ \mathring{A}$
$x_j \sim 0.06\ \mu m$
$V_{DD} = 0.8 - 1.5\ V$

Main point: p^+-super-steep retrograde body doping controls short-channel effects while preserving high mobility.

New device architecture: Silicon-on-Insulator (SOI)



Schematic of nFET on SOI and equivalent devices.
*Adapted from Shahidi et al., **Proc. ISSCC**, 1999 (426).*



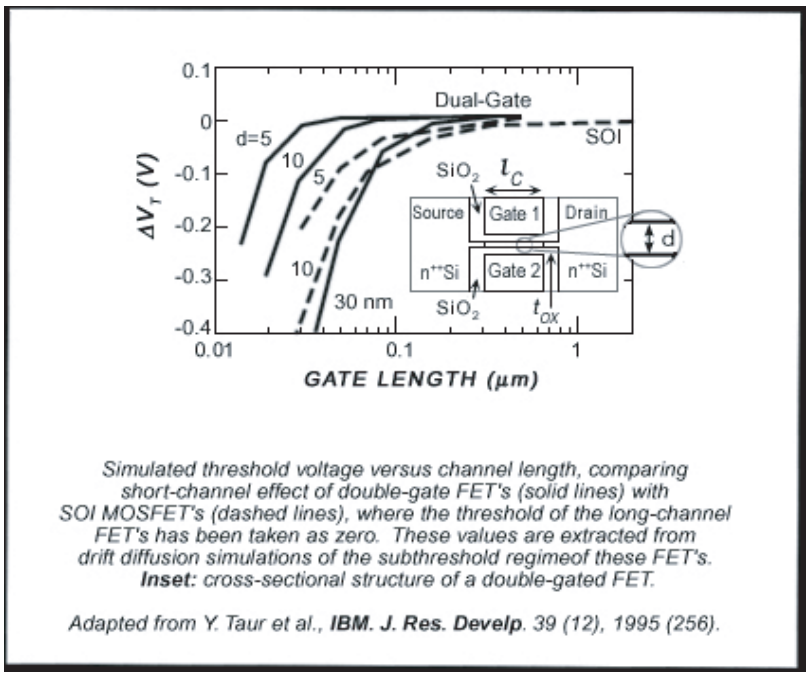
Power vs. Frequency
*Adapted from Shahidi et al., **Proc. ISSCC**, 1999 (426).*

Figure 25.1.1 in: Shahidi, G.G., et al. "Partially-depleted SOI Technology for Digital Logic." *International Solid-State Circuits Conference, San Francisco, CA, Feb. 15-17, 1999. Digest of Technical Papers*. New York, NY: Institute of Electrical and Electronics Engineers, 1999, pp. 426-427. ISBN: 9780780351264. © 1999 IEEE.

A number of issues associated with existence of buried oxide:

- reduced junction capacitance
- floating body: kink effect, extra drive ($V_{BS} > 0$ during switching)
- increased thermal resistance

New device architecture: Dual-gate MOSFET



Simulated threshold voltage versus channel length, comparing short-channel effect of double-gate FET's (solid lines) with SOI MOSFET's (dashed lines), where the threshold of the long-channel FET's has been taken as zero. These values are extracted from drift diffusion simulations of the subthreshold regimeof these FET's. ***Inset:*** *cross-sectional structure of a double-gated FET.*

Adapted from Y. Taur et al., ***IBM. J. Res. Develp****. 39 (12), 1995 (256).*

Figure 26 in Taur, Y., et al. "CMOS Scaling into the Nanometer Regime." *Proceedings of the IEEE* 85, no. 4 (1997): 486-504. © 1997 IEEE.

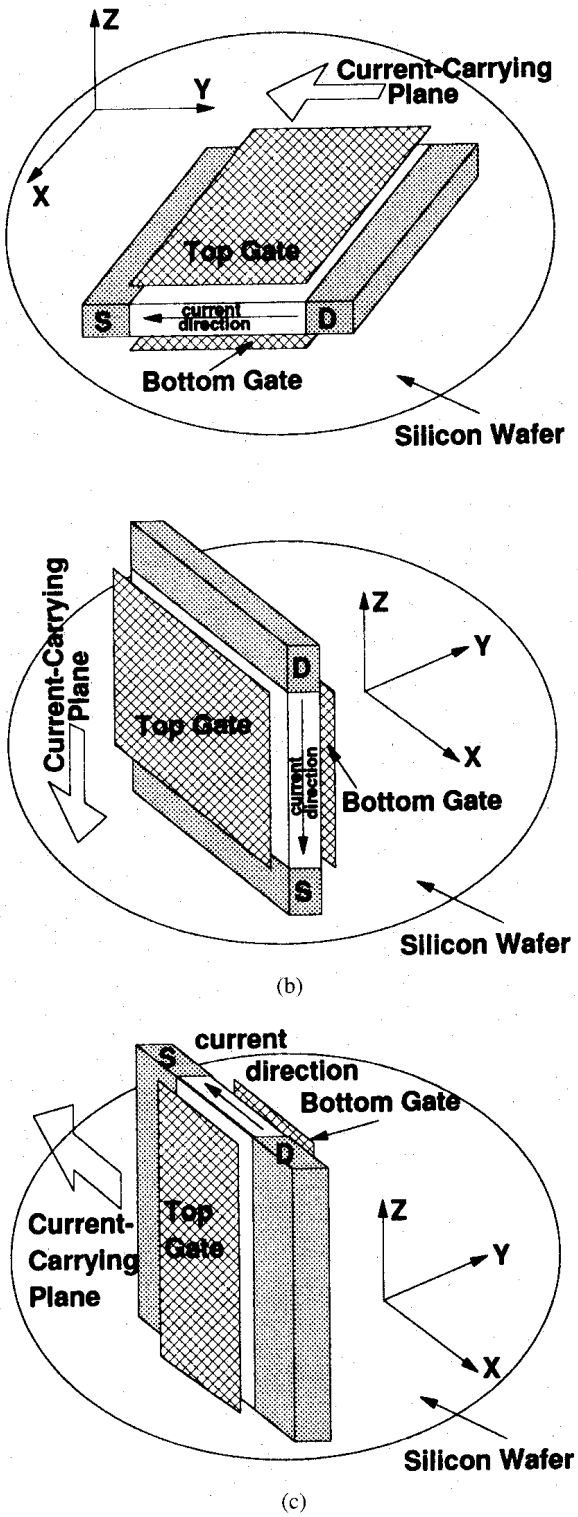
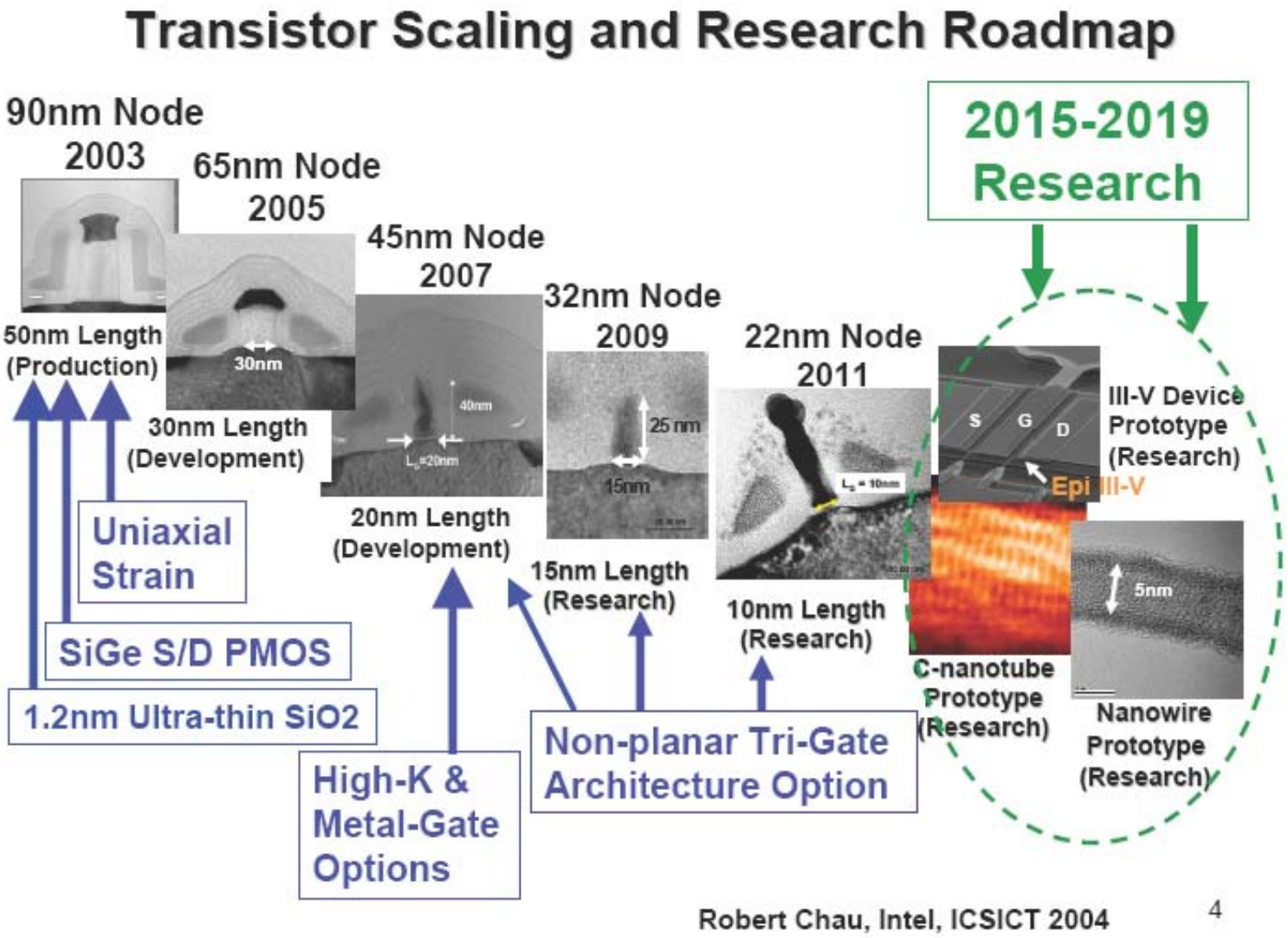


Fig. 29. Three possible orientations of a double-gate MOSFET on a silicon wafer. Examples of devices fabricated are: (a) Colinge *et al.* [54], Tanaka *et al.* [55], (b) Takato *et al.* [56], and (c) Hisamoto *et al.* [57].

Figure 29 in Taur, Y., et al. "CMOS Scaling into the Nanometer Regime." *Proceedings of the IEEE* 85, no. 4 (1997): 486-504. © 1997 IEEE.

□ Intel's current (public) view of MOSFET scaling...



Chau, R., et.al. "Advanced CMOS Transistors in the Nanotechnology Era for High-Performance, Low-Power Logic Applications." In *Proceedings of the 7th International Conference on Solid-State and Integrated Circuit Technology*. Beijing, China: IEEE Press, 2004, pp. 26-30. Copyright 2004, IEEE. Used with permission.

Key conclusions

- MOSFET scaling has taken place in a harmonious way with all dimensions and voltage scaling down.
- *The end of conventional MOSFET scaling is close!* Biggest barrier to MOSFET scaling is gate oxide leakage: need new gate dielectric with higher dielectric constant.
- To improve electrostatic integrity with limited oxide scaling: SOI, double gate designs, triple gate designs.
- To improve performance: use strained Si or strained-Si/SiGe heterostructures.
- Also, use metal gate.