

Rozwój technologii krzemowej – dokąd zdążamy ?

Marek Idzik

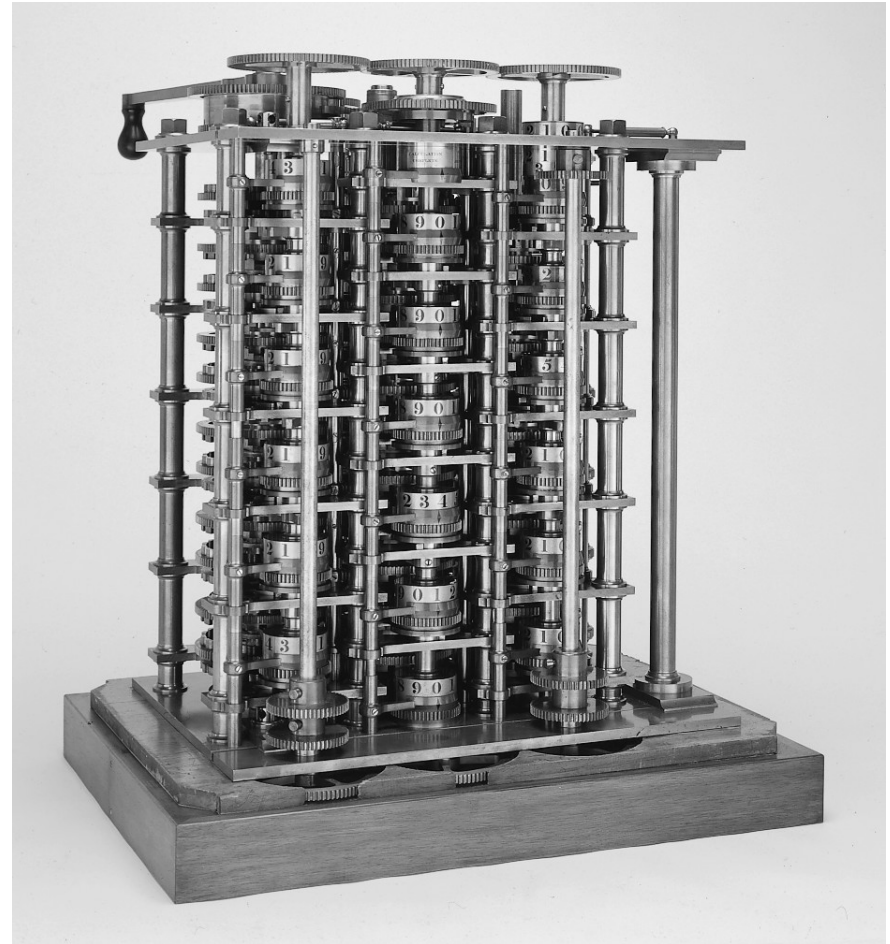
WFiIS AGH

Plan

- Trochę historii
- Gdzie jesteśmy, czyli CMOS i skalowanie
- Dokąd zmierzamy czyli granice obecnego rozwoju
- Co dalej ?

Trochę historii

- ❑ 500BC: **Abacus** liczenie na koralikach, wynaleziony w Chinach
- ❑ 1A.D.: **Antikythera** mechanizm odtwarzający ruchy słońca, księżyca i innych planet
- ❑ 1642: **Blaise Pascal** buduje pierwszą mechaniczną maszynę liczącą
- ❑ 1823: **Charles Babbage** buduje mechaniczny komputer



Babbage Difference Engine
(25000 elementów)

Trochę historii...

- ❑ 1936: Angielski matematyk **Alan Turing** publikuje artykuł “On computable numbers ...”. Pracuje też przy “Colossus” (1943) dla wojska.
- ❑ 1946: **ENIAC**(Electronic Numerical Integrator and Calculator) pierwszy całkowicie elektroniczny komputer na lampach elektronowych (University of Pensylwania)



ENIAC -18000 lamp, 140kW

Troche historii..., przewidywanie przyszłości

“Where a calculator on the ENIAC is equipped with 18,000 vacuum tubes and weighs 30 tons, computers in the future may have only 1,000 vacuum tubes and perhaps weigh 1.5 tons.”

Popular Mechanics, March 1949

“I think there is a world market for maybe five computers.”

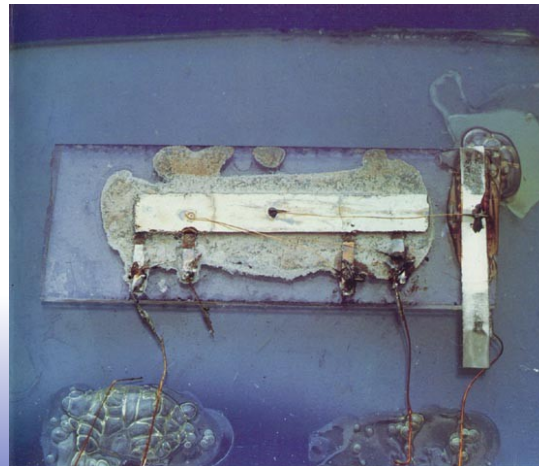
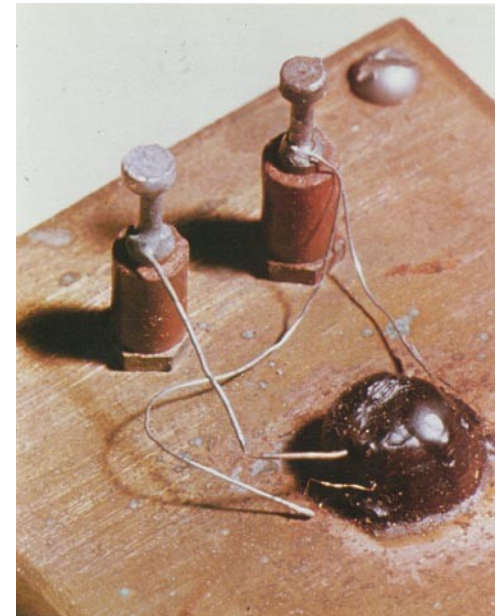
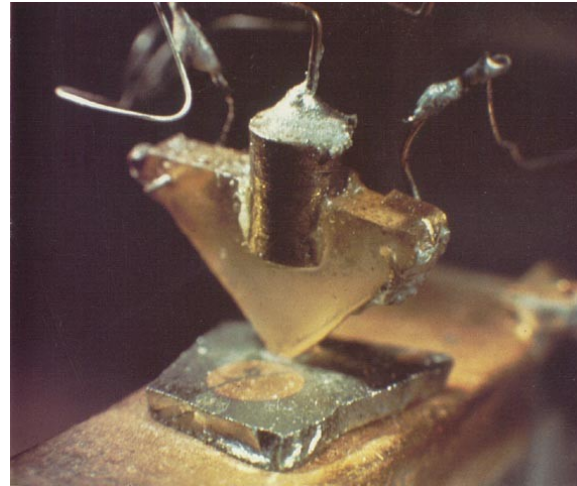
Thomas Watson, Chairman of IBM, 1943

“Prediction is very difficult, especially of the future.”

Niels Bohr

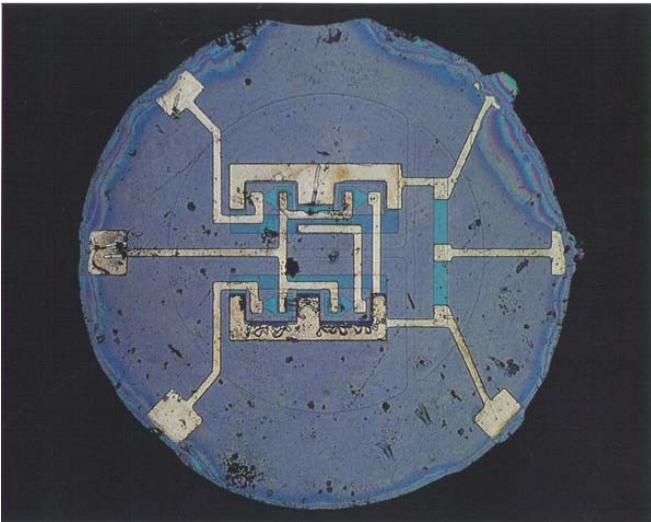
Początek rewolucji

- ❑ 1947: Bardeen and Brattain (Bell Labs) tworzą pierwszy tranzystor ostrzowy (U.S. Patent 2,524,035), nobel 1956
- ❑ 1951: Shockley (Bell Labs) konstruuje pierwsze tranzystory złączowe (U.S. Patent 2,623,105), nobel 1956
- ❑ 1959: Jack Kilby (Texas Instruments) buduje pierwszy układ scalony (“integrated circuit”), nobel 2000

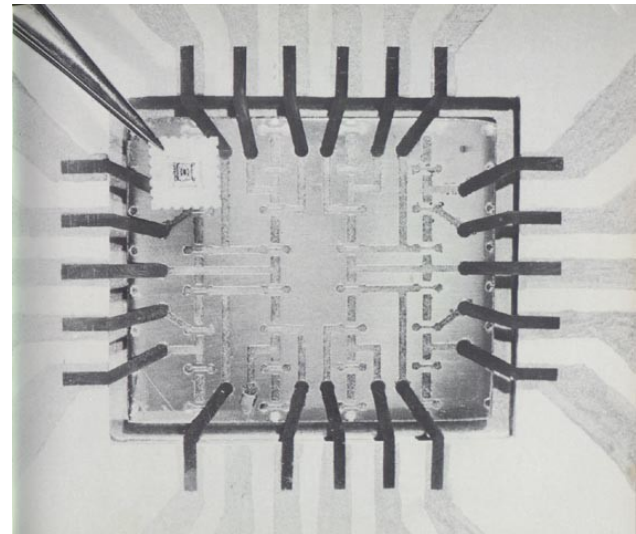


Pierwsze scalone układy logiczne

- ❑ 1961: Texas Instruments i Fairchild wprowadzają pierwsze scalone układy logiczne ICs (~\$50 za sztukę)
- ❑ 1962: RCA wprowadza pierwsze układy na tranzystorach typu MOS



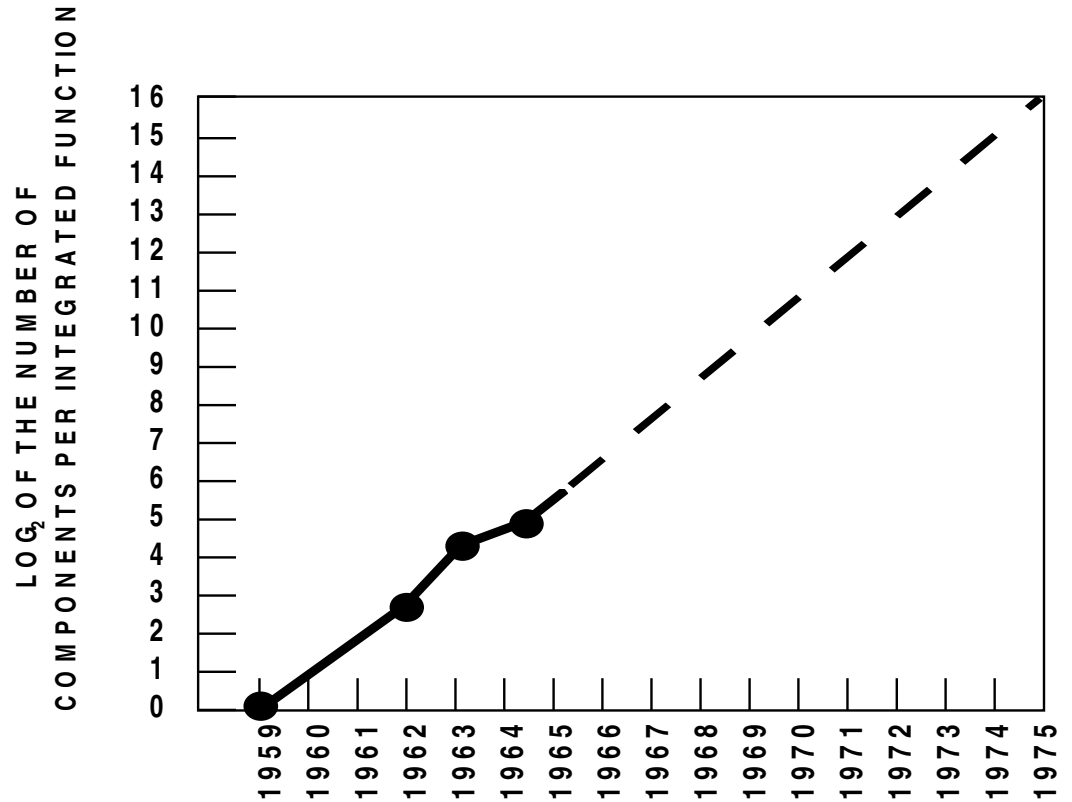
Fairchild bipolar RTL Flip-Flop



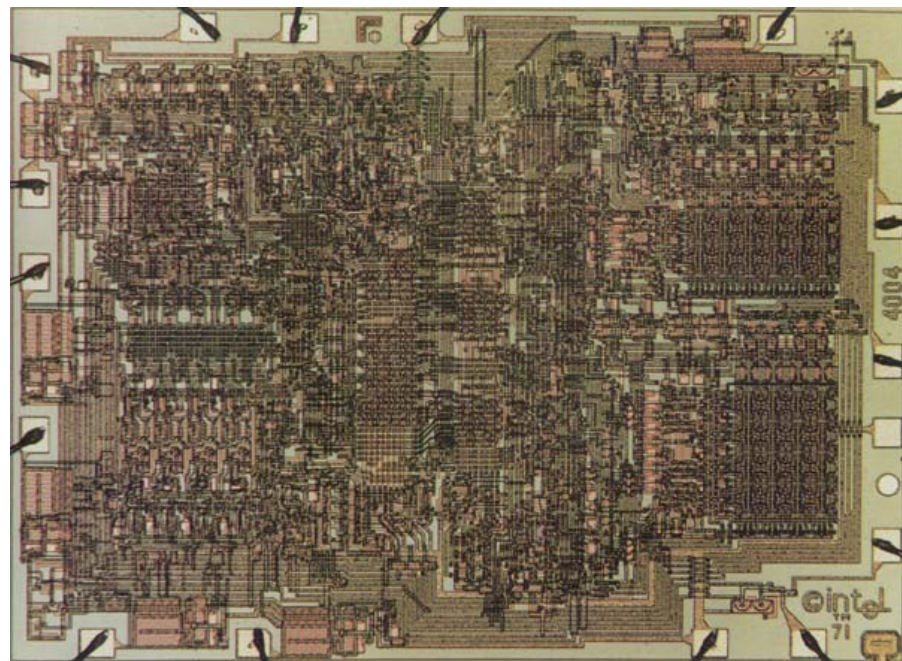
RCA 16-transistor MOSFET IC

“Moore’s Law”

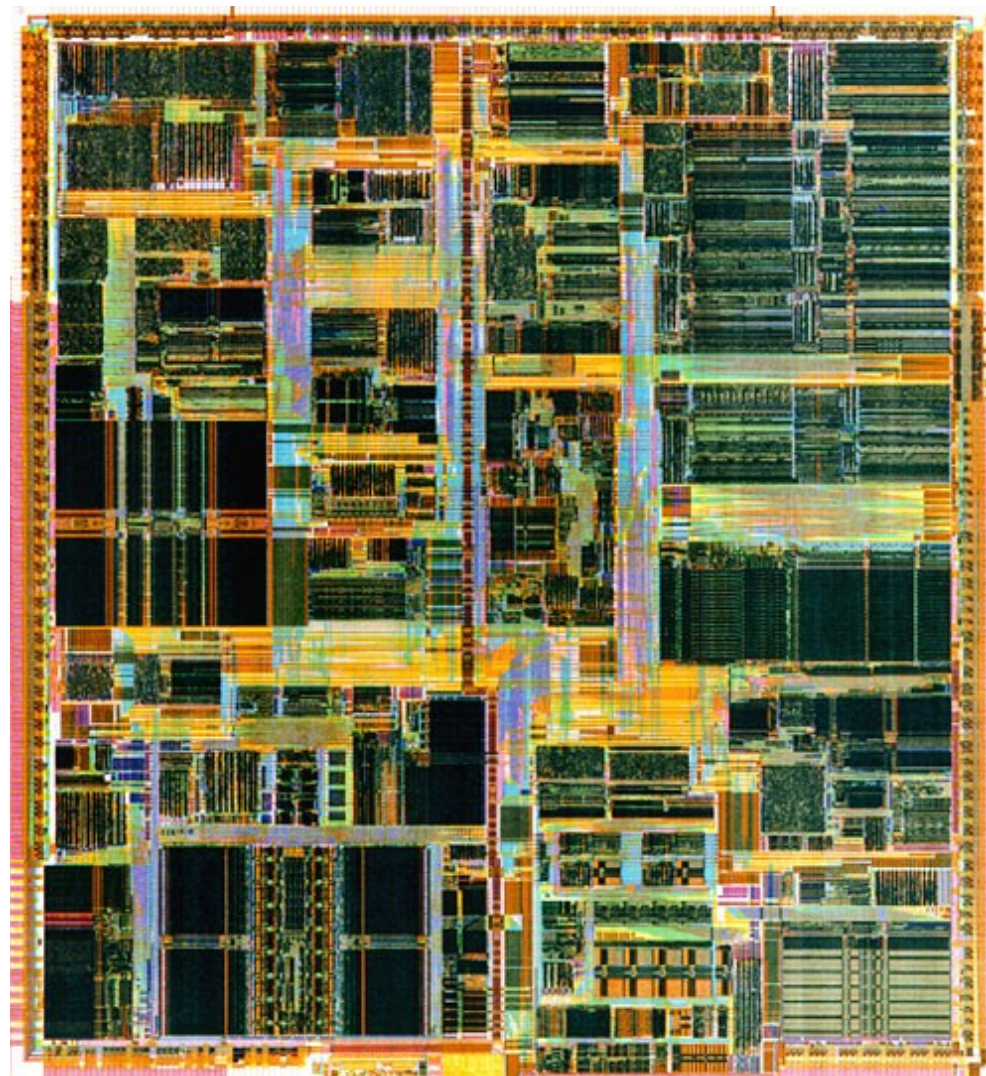
- W 1965 Gordon Moore (Fairchild, w 1968 współzałożyciel Intelu) zauważa, że ilość tranzystorów na “chip-ie” podwaja się co roku
- W 1975 Moore przewiduje, że układy półprzewodnikowe będą podwajać swą moc obliczeniową co 18 miesięcy



Epoka CMOS !

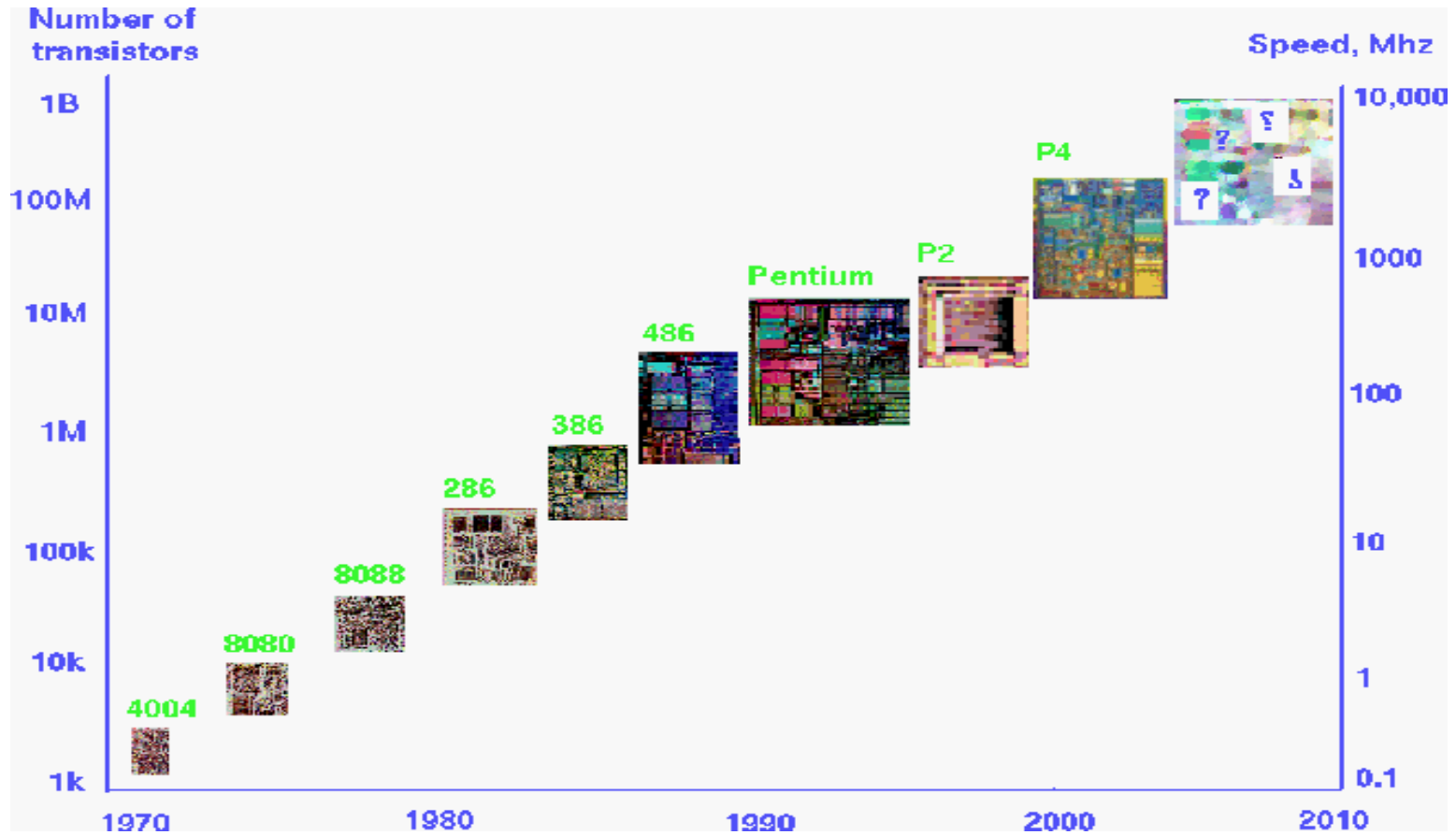


1971-Intel 4004: 0.74 MHz,
2200 tranzystorów, $L_{\min} = 10\mu\text{m}$

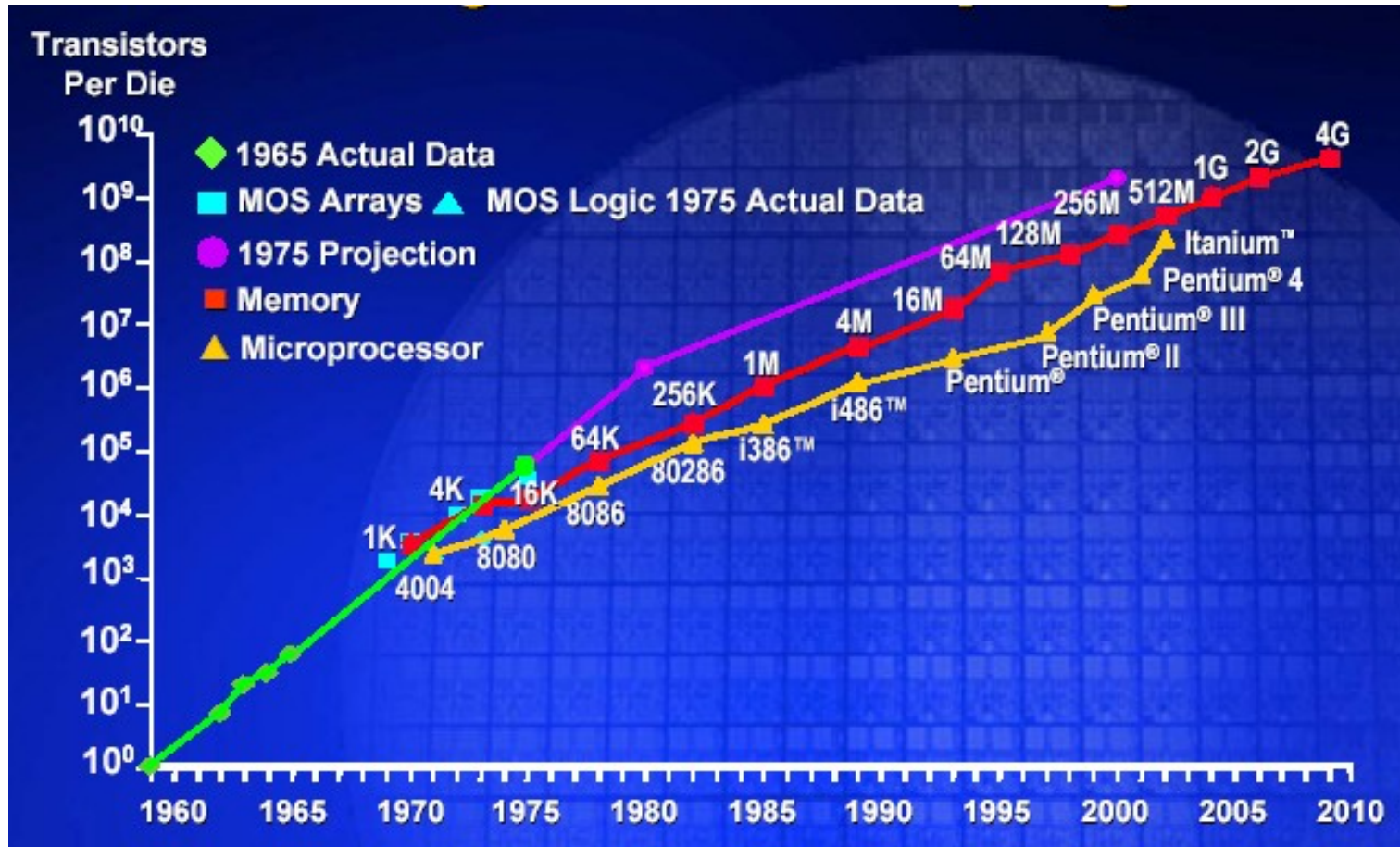


2004-Intel PIV: 3800 MHz,
125 000 000 tranzystorów, $L_{\min} = 0.09\mu\text{m}$

Trendy rozwoju technologii CMOS

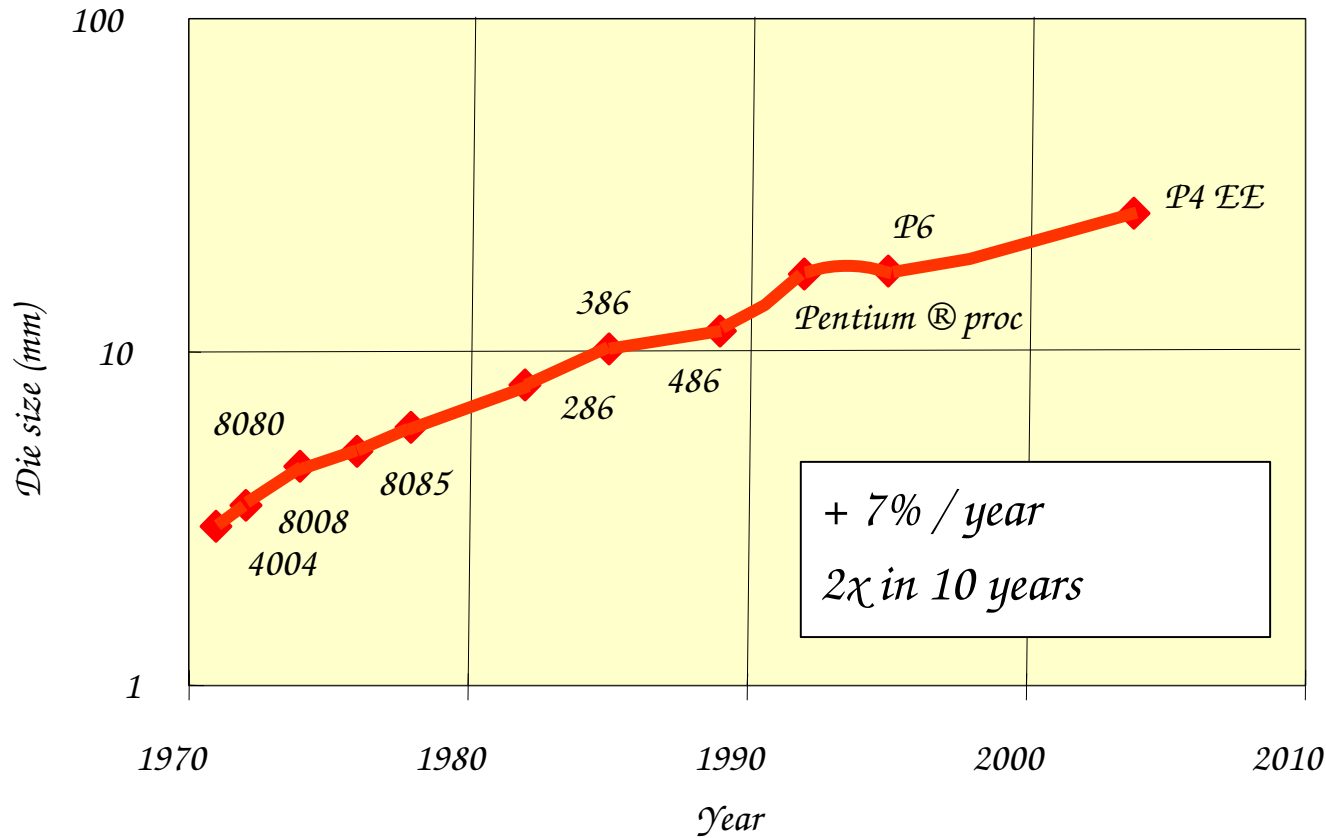
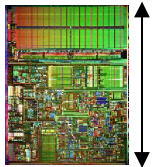


“Moore’s Law” vs rozwój CMOS



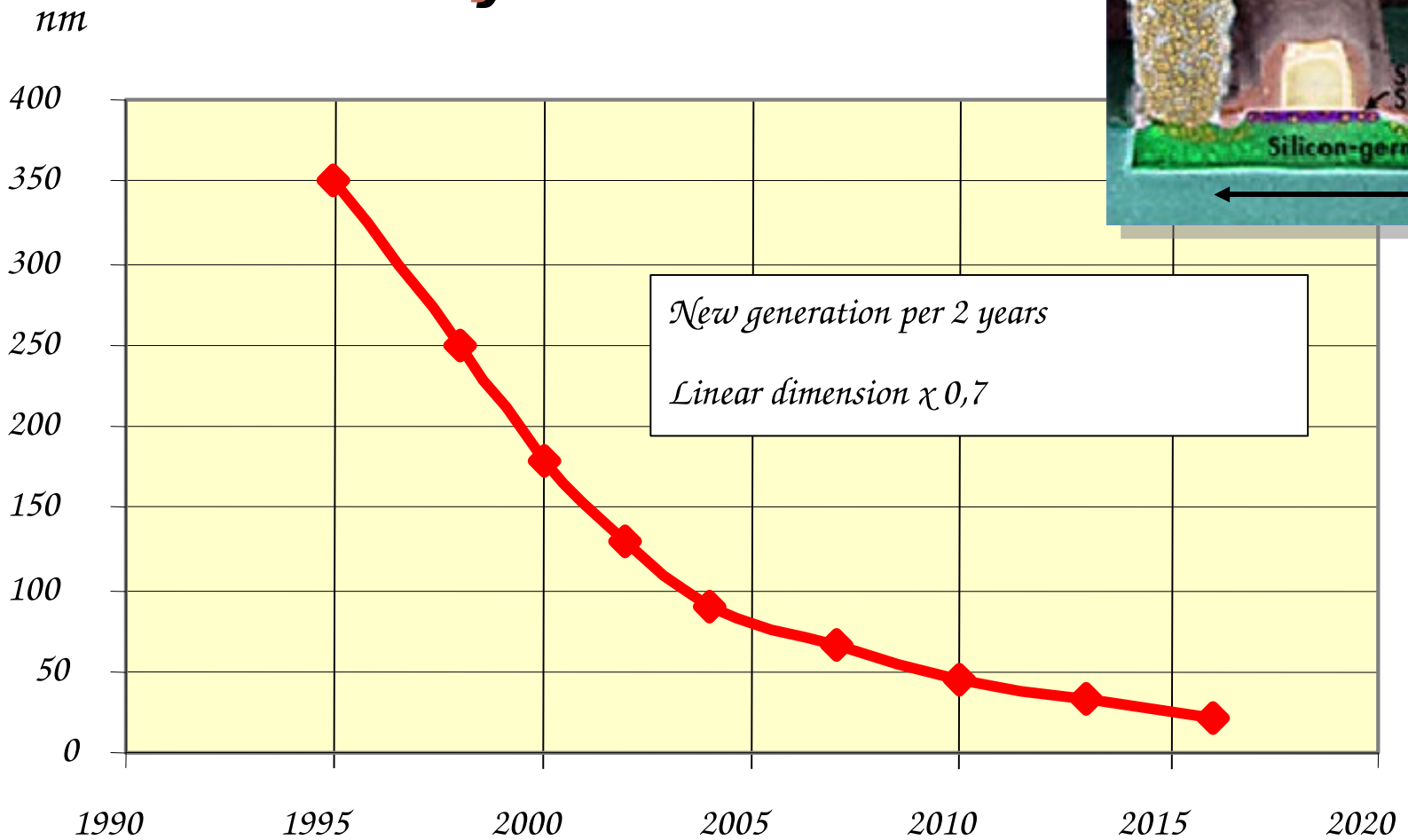
[Source: Intel]

Liniowy wzrost wielkości procesorów



[Source, Intel]

Rozmiar tranzystora



[there are about 5 atoms in 1 nm]

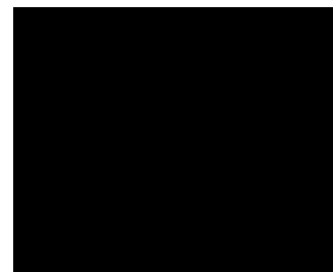
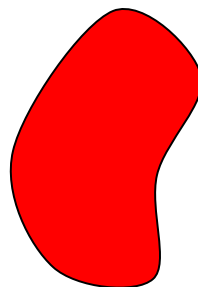
[Source: ITRS 2003 roadmap]

Rozmiar charakterystyczny

Komórka roślinna

100 μm

Komórka zwierzęca



*1971
10 μm*

Bakteria



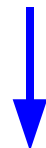
*1989
1 μm*

virus



*2004
0.09 μm*

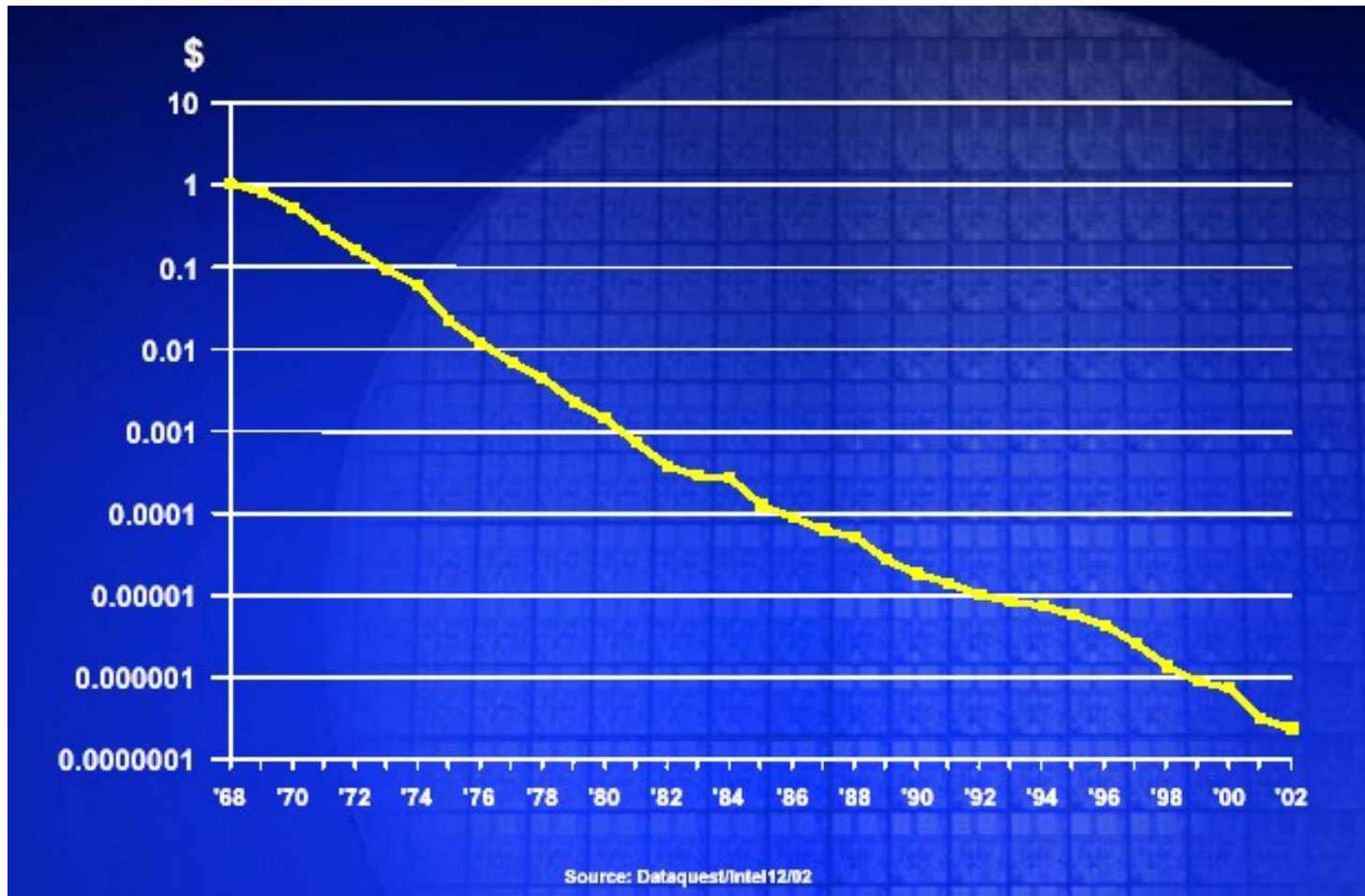
*Dużo do
zrobienia!*



atom

0.0002 μm

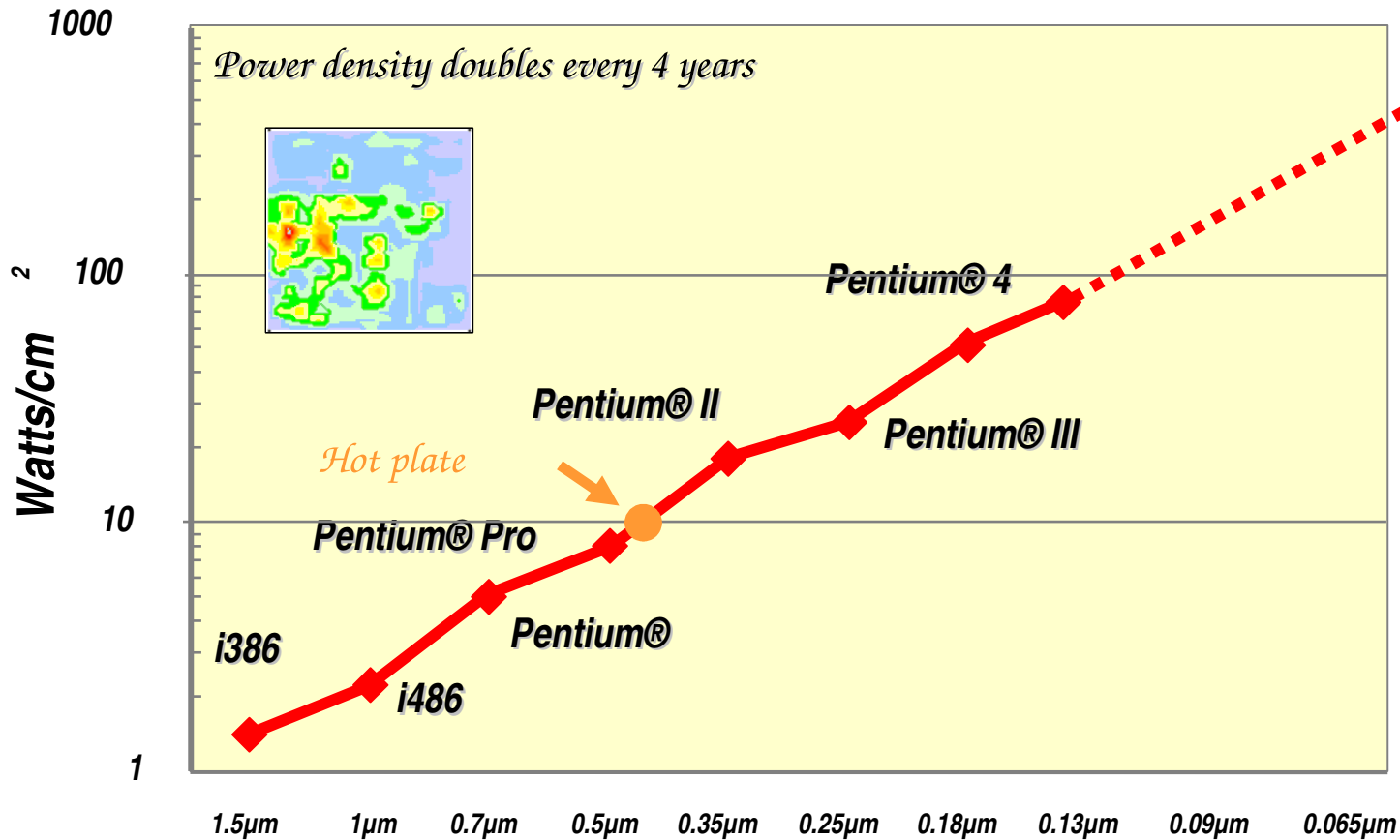
Koszt tranzystora



100 nanodollar or the price of one grain of salt

Tranzystory są za darmo !

Gęstość wydzielanej mocy

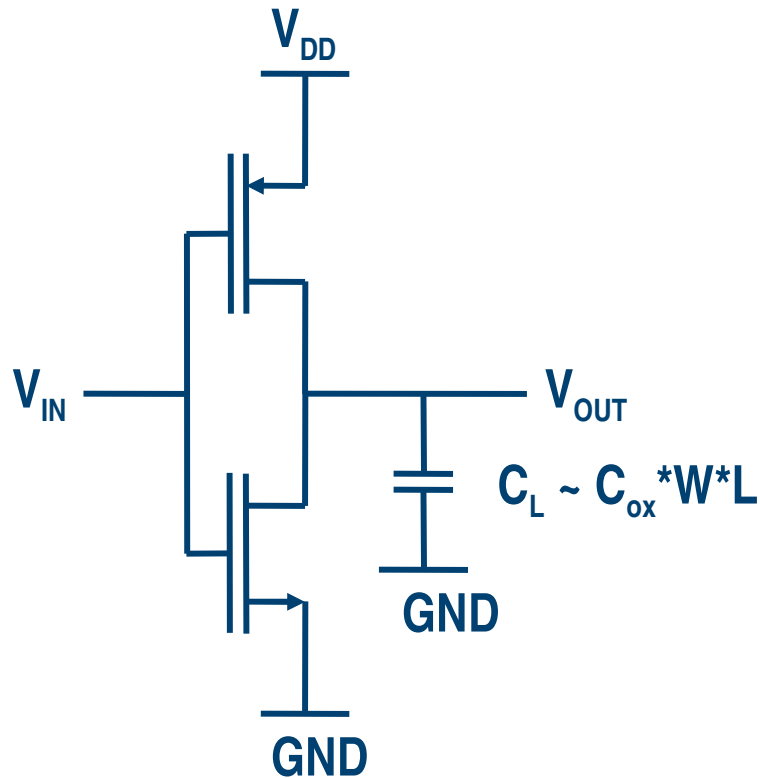


[Source: Intel]

Pojawia się problem...

Moc w technologii CMOS

Przykład: CMOS inverter



$$P_{\text{static}} = I_{\text{leakage}} \cdot V_{\text{DD}} \cong 0 !!!$$

$$P_{\text{dynamic}} = C_L \cdot V_{\text{DD}}^2 \cdot f$$

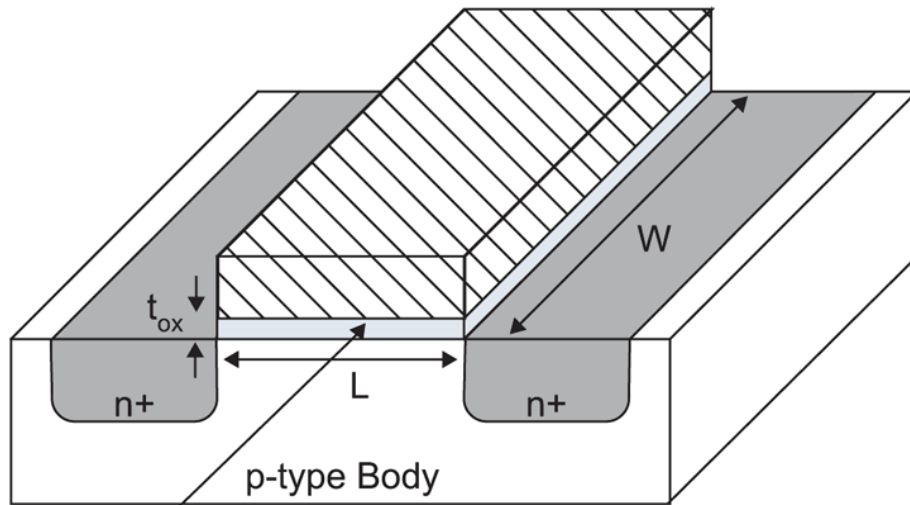
$$\text{PDP} = C_L \cdot V_{\text{DD}}^2$$

Power-delay product

Kontrprzykład TTL(low power): $2\text{mW} \cdot 25\text{M}$ bramek = 50 kW !!!

Skalowanie

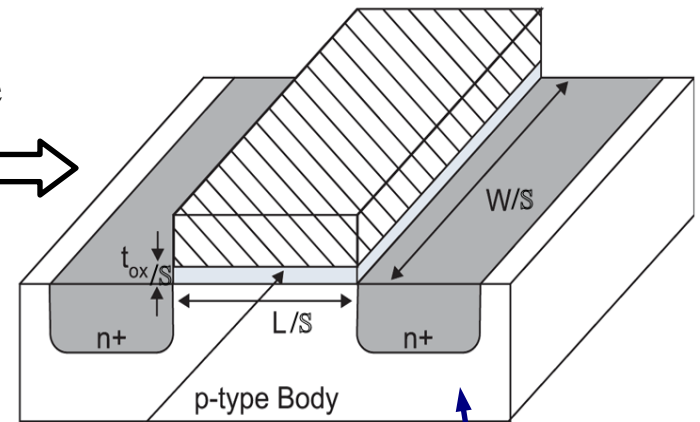
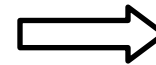
(pierwotny tranzystor NMOS)



SiO₂ Gate Oxide
(Good insulator, $\epsilon_{ox} = 3.9\epsilon_0$)

(przeskalowany NMOS)

scale



SiO₂ Gate Oxide

t_{ox} ↓ V_{DD} ↓ C_L ↓

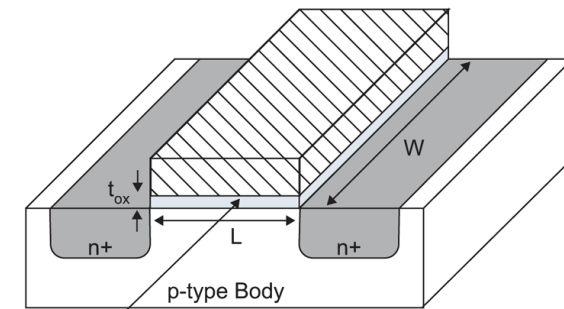
domieszkowanie
zwiększone S razy

Skalowanie pozwala zwiększyć gęstość, szybkość oraz zmniejszyć pobór mocy układów cyfrowych

$$W_D = \sqrt{\frac{2 \epsilon_{si} (\psi_{bi} + V_{dd})}{q N_a}}$$

zmniejsza się S razy

Skalowanie



SiO₂ Gate Oxide
(Good insulator, $\epsilon_{ox} = 3.9\epsilon_0$)

Parameter	Constant-field scaling	Generalized field scaling
Physical dimensions: L, W, Tox, wire pitch	1/S	1/S
Body doping concentration	S	E/S
Voltage	1/S	E/S
Circuit density	1/S ²	1/S ²
Capacitance per circuit	1/S	1/S
Circuit speed	S	S (goal)
Circuit power	1/S ²	E ² /S ²
Power density	1	E ²
Power-delay product	1/S ³	E ² /S ³

S – czynnik skalujący
E=V/S - znormalizowane
pole elektryczne

Granice skalowania

V_{DD}/L rośnie zatem i E
rośnie, maleje ruchliwość

$$\frac{V_T = kT/q}{V_{DD}}$$

Dyfuzja termiczna
elektronów powoduje
wzrost prądu upływu
źródło-dren

$$\frac{V_T}{V_{th}}$$

Tunelowanie przez SiO_2
powoduje wzrost prądu
upływu bramki

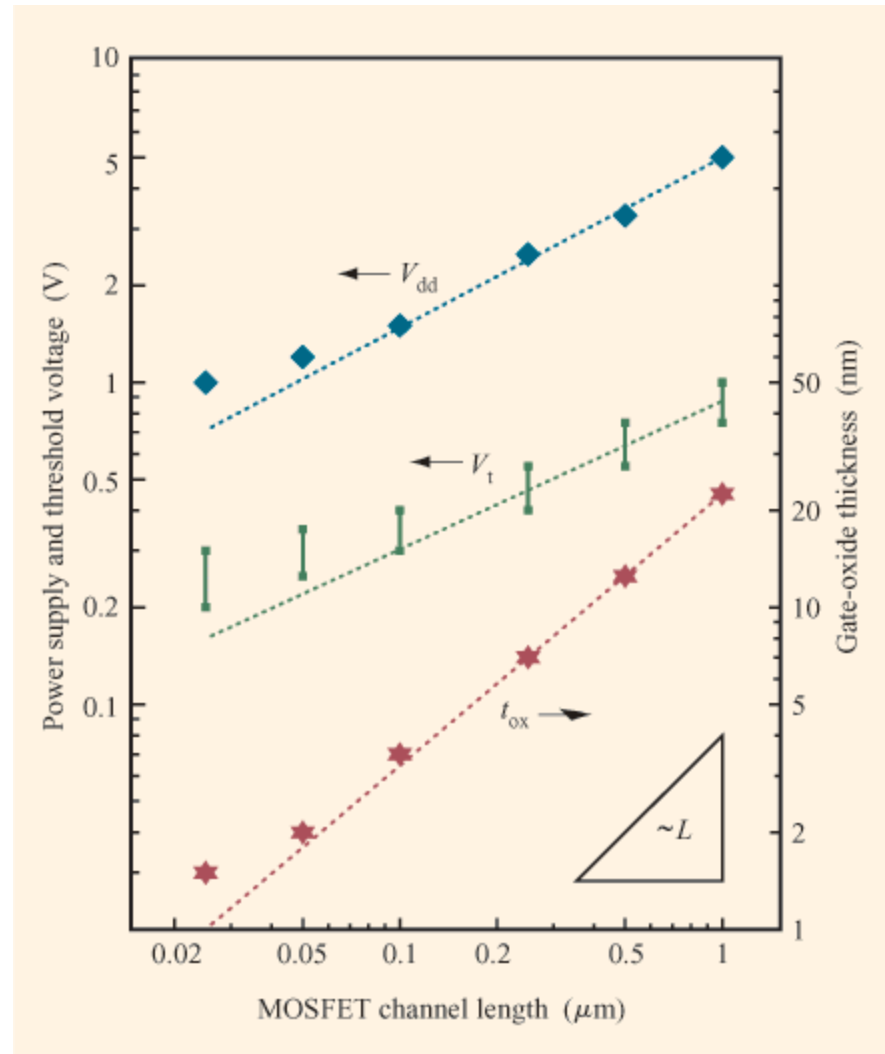


Figure 1

History and trends of power-supply voltage (V_{dd}), threshold voltage (V_t), and gate-oxide thickness (t_{ox}) vs. channel length for CMOS logic technologies.

Prąd upływu I_{DS}

Dominującą składową prądu upływu jest prąd źródło-dren. Niestety poniżej napięcia progowego nie spada on do zera. Tranzystor wchodzi wtedy w zakres podprogowy z eksponencjalną zależnością prądu od napięcia bramki

$$I_D \propto \frac{W}{L} \exp\left(\frac{q(V_{GS} - V_T)}{nkT}\right)$$

Przy 100 milionach tranzystorów
akceptowalny prąd upływu $\sim 10^{-8}$ A
co daje traconą moc ~ 1 W

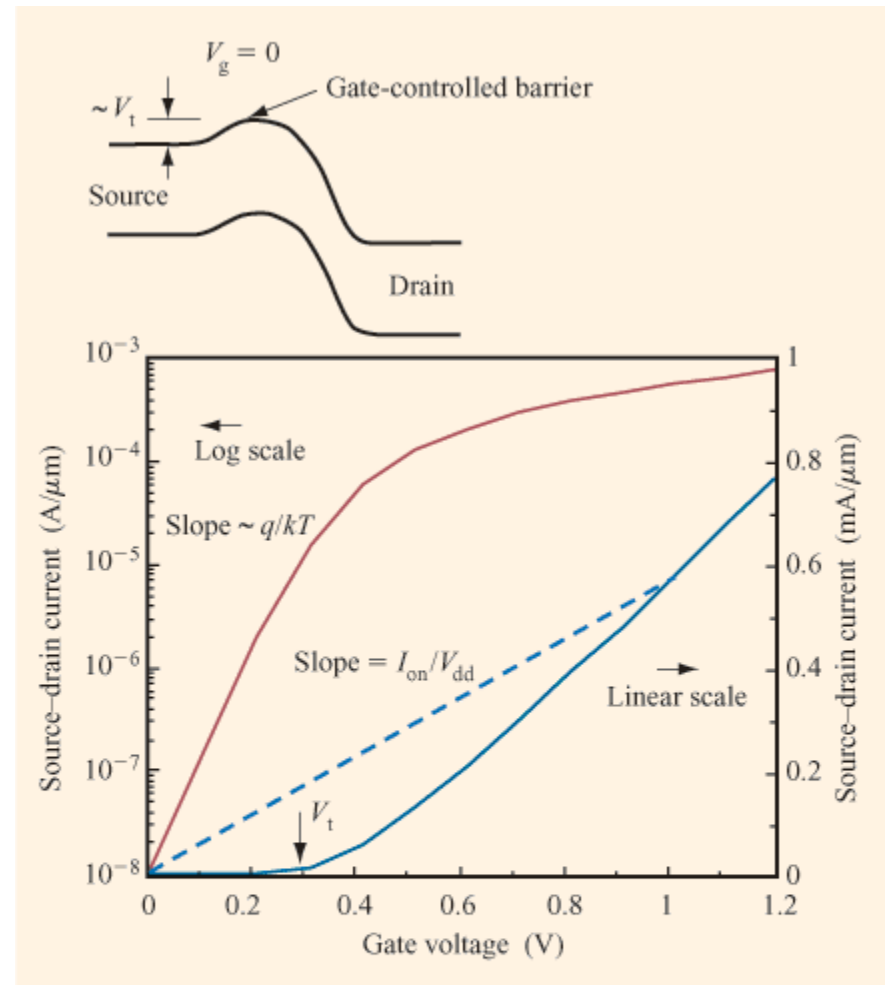


Figure 2

MOSFET current in both logarithmic (left) and linear (right) scales vs. gate voltage. The slope of the dotted line represents the large-signal transconductance for a digital circuit. Inset shows the band diagram of an n-MOSFET. The barrier height at $V_g = 0$ is proportional to V_t .

Prąd upływu bramki

Prąd tunelowy bramki zaczyna istotnie rosnąć gdy grubość t_{ox} jest rzędu 5-10 warst atomowych.

Zadziwiające jest, że przy tak cienkiej SiO_2 nie ma problemów z defektami !

Przyjmując powierzchnię bramki $0.1\mu\text{m}^2$, maksymalny akceptowalny prąd tunelowy wynosi $\sim 10\text{ A/cm}^2$ (moc $\sim 1\text{W}$). Wyznacza to minimalne $t_{\text{ox}} \sim 1.5\text{nm}$

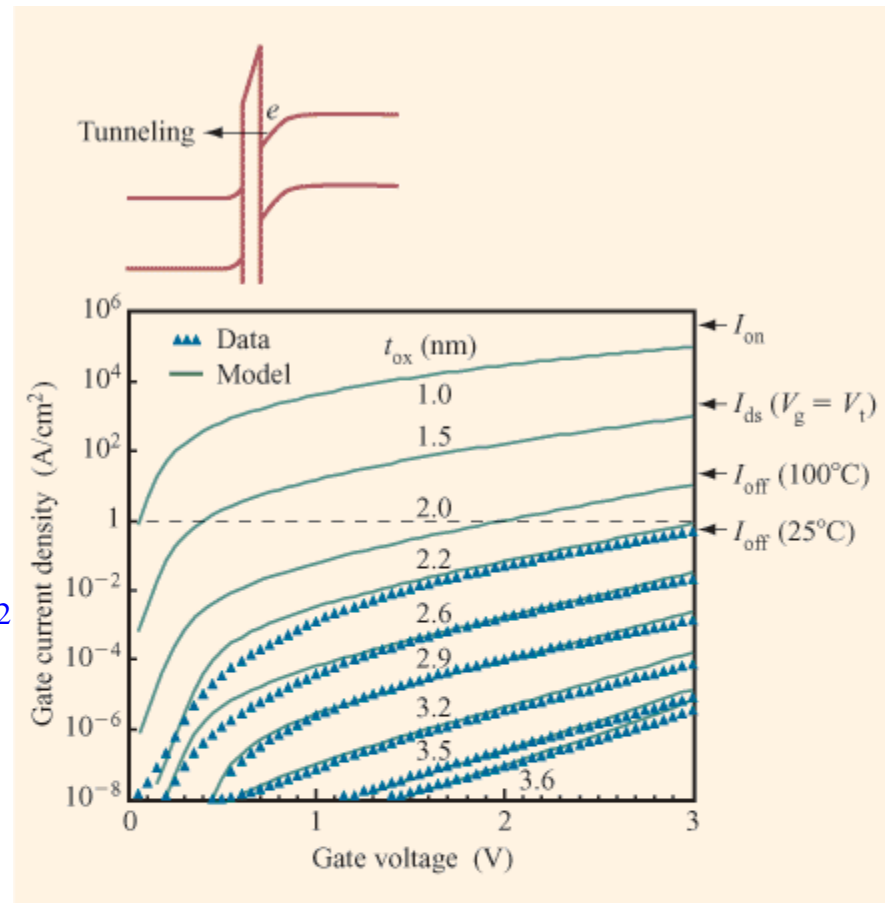


Figure 3

Measured and calculated oxide tunneling currents vs. gate voltage for different oxide thicknesses. Labels on the right, from the bottom up, mark the order of magnitude of off-currents at room and worst-case temperatures, source-to-drain current at $V_g = V_t$ ($V_{ds} = V_{dd}$), and on-current at $V_g = V_{ds} = V_{dd}$. The inset shows the band diagram for tunneling in a turned-on n-MOSFET.

Moc i szybkość

Przy długości bramki poniżej 100nm moc spoczynkowa stanowi ograniczenie. Dlatego nie można dalej stosować “tradycyjnego” skalowania

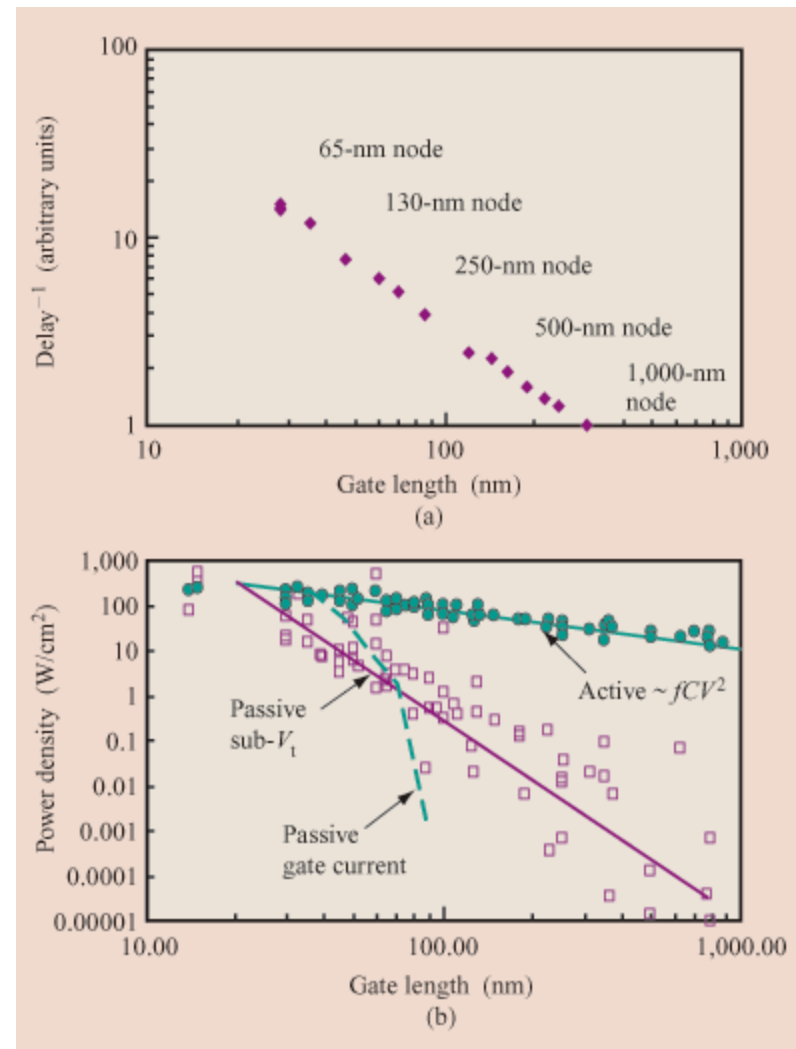
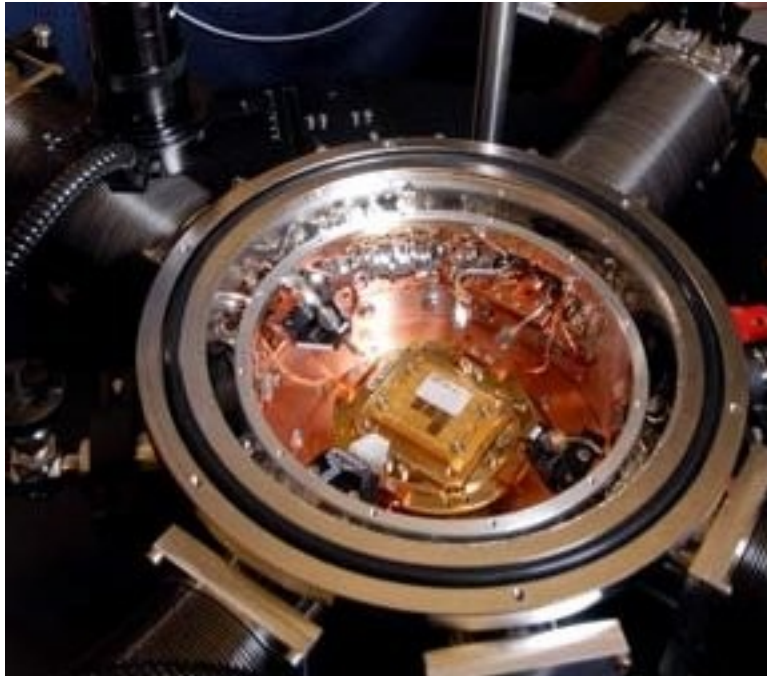


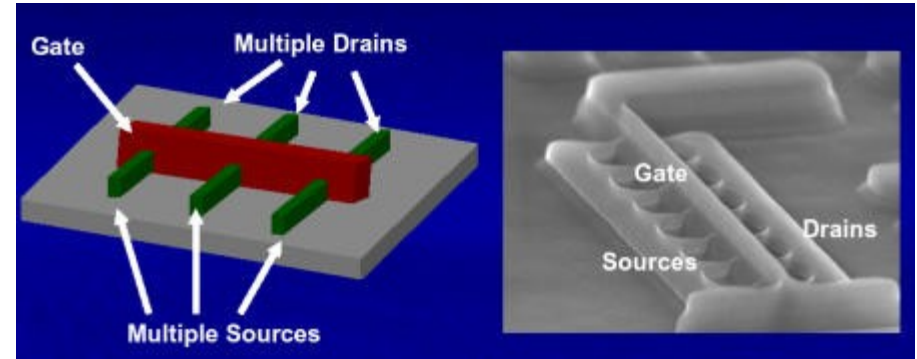
Figure 1

(a) MOSFET performance vs. gate length; normalized MOSFET intrinsic device delay (CV/I_{eff}) vs. gate length. (b) Power density vs. gate length; data collected from literature for active power density and passive power density. Lines are intended to show trend. ($fCV^2 = \text{frequency} \times \text{capacitance} \times \text{voltage}^2$.)

Przykładowe kierunki rozwoju

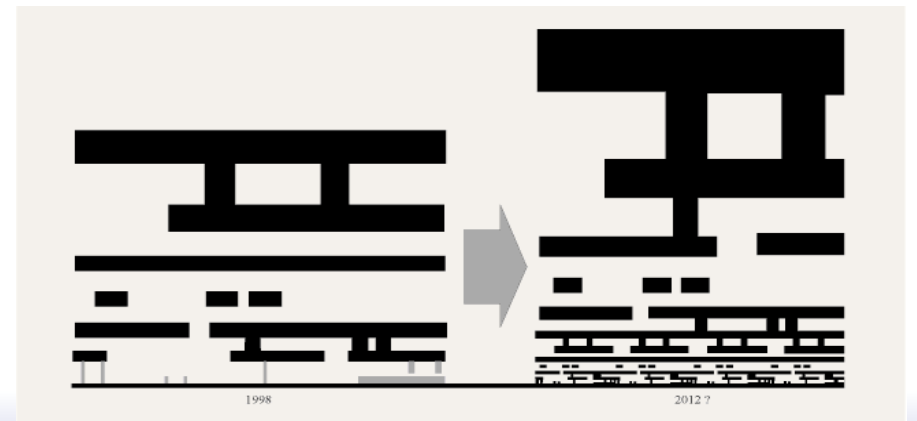


Tranzystory Si-Ge, $T = -268$ °C
(2006 VI - “mikrochip” IBM
pracuje przy $f = 500$ GHz)



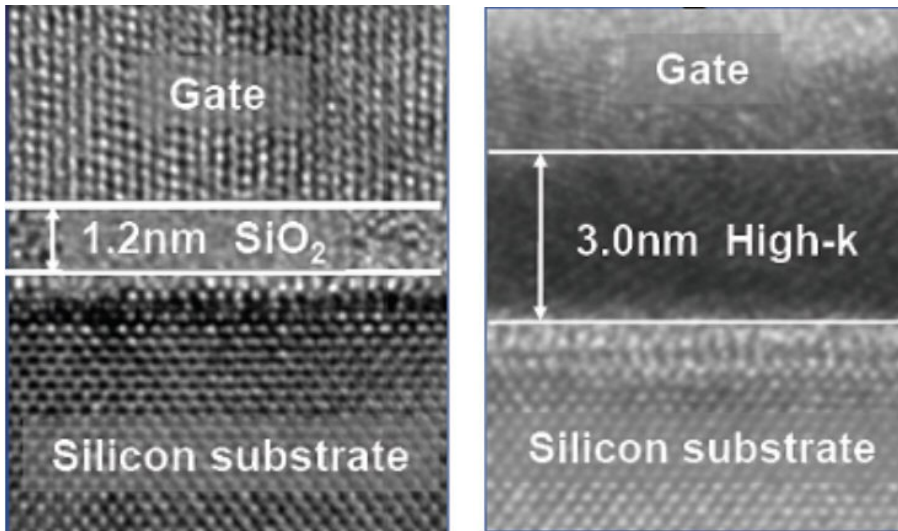
Tranzystory 3D

(moc, prądy upływu, ruchliwość)



3D “wiring” (moc, częstotliwość, RC)

Bariera 1 nm - dlaczego tlenki high-K ...



1.2nm(5 warstw atom.) SiO₂ stosowany w techn. 90nm

- Aby zapobiec zwiększeniu prądów upływu – grubszy izolator
- Aby nie zmniejszyć pojemności – większa stała dielektryczna

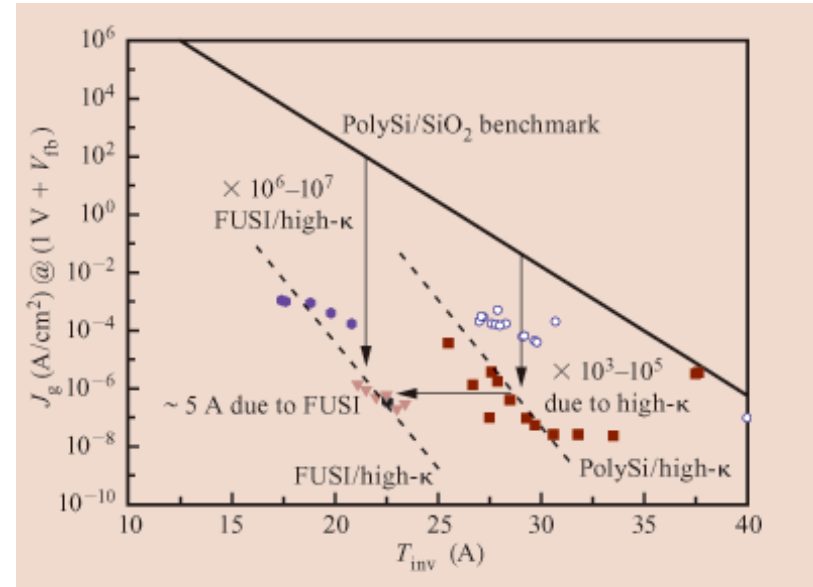


Figure 16

Gate leakage current density as a function of inversion thickness for polySi-gated devices with SiO₂ and high- κ dielectrics and fully silicided (FUSI) devices with high- κ dielectrics. From [131], reproduced with permission; ©2004 IEEE.

Granice skalowania CMOS

Dotychczasowe przewidywania sceptyków nie sprawdziły się zupełnie, “Moore's Law” jest ciągle aktualne. Ale:

*No exponential is forever...
but we can delay “forever”*

Gordon Moore, ISSCC 2003

- ❑ Obecne szacunki określają technologicznie możliwą granicę na $L \sim 7\text{nm}$ (od 2.5nm do 10nm), w czasie ~ 10 lat. Oczekuje się przy tym ~ 10 krotnego wzrostu wydajności
- ❑ Rezultat na pewno zależeć będzie od kosztów udoskonalania technologii i opłacalności tej inwestycji

Co po ~2020 roku ?

Carbon Nanotubes

V-Groove Transistors

Organic Transistors

Quantum Computing

Molecular Devices

The collage features several images: a 3D model of carbon nanotubes, a cross-section of a V-groove transistor with a 100 Å scale bar, a scanning electron micrograph of a transistor with labels for Source (n⁺⁺ Si), Gate (SiO₂), and Drain (n⁺⁺ Si) and a 20 nm scale bar, a diagram of a pentacene structure on a flexible organic transistor, a 3D molecular model of a quantum computing structure with Fe and ¹⁹F atoms, a 3D surface plot, and a scanning electron micrograph of a molecular device array.

A może kolejna epoka krzemu ?

“Prediction is very difficult, especially of the future”