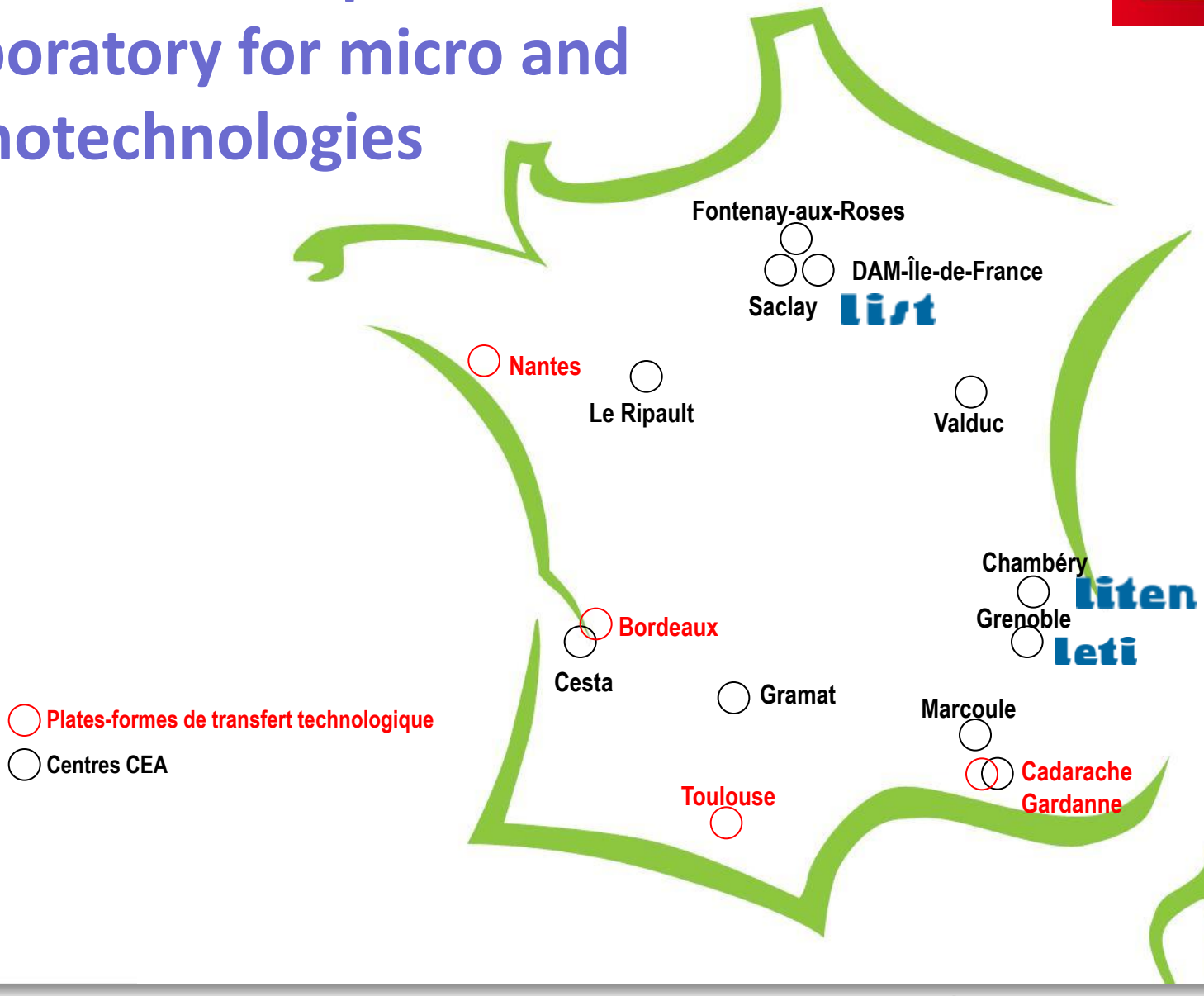
Energy Efficient Nanoelectronics

Hervé Fanet

CEA-Léti

Léti is a french public research Laboratory for micro and nanotechnologies

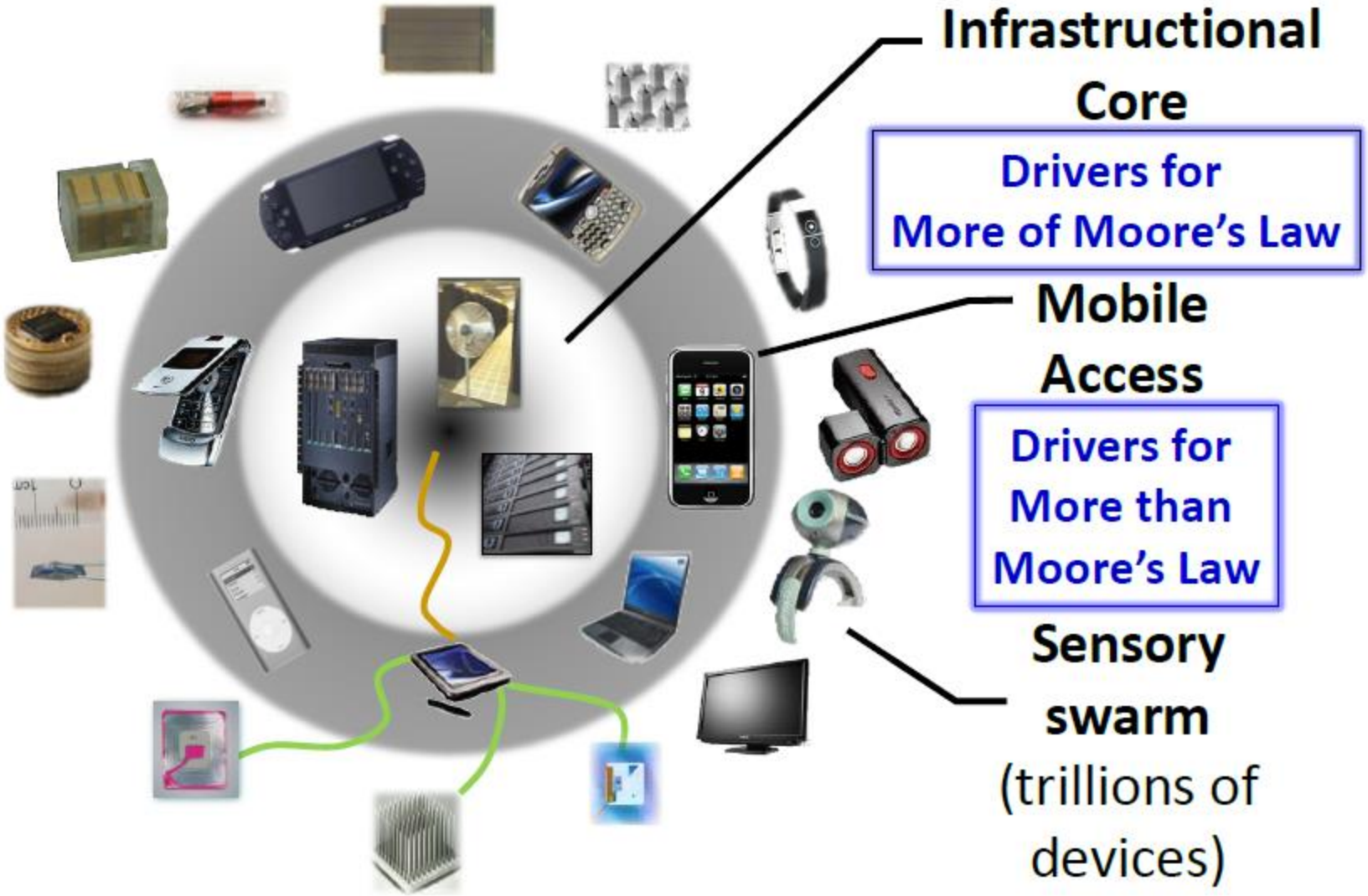


MINATEC on the scientific polygone

Why less power? TODAY

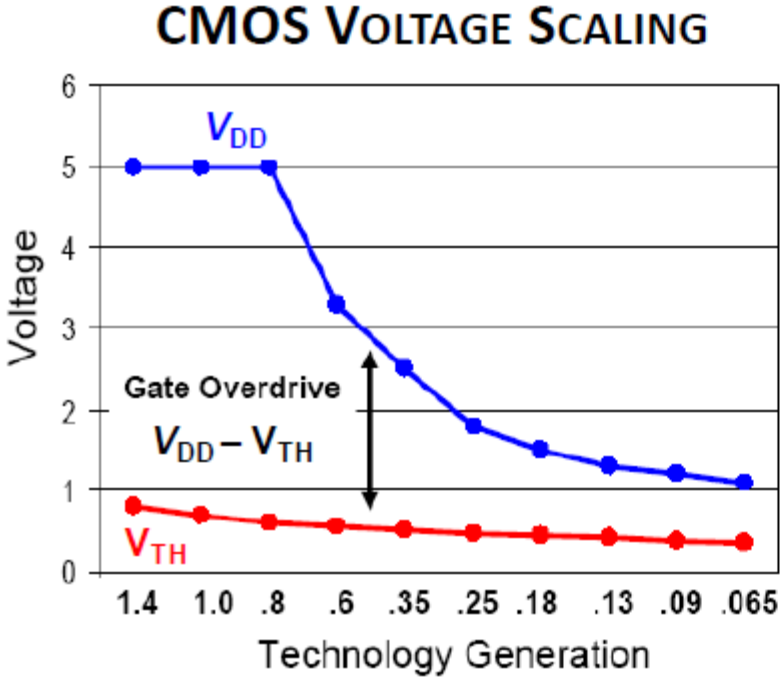
- access to powerful computation at any location
  - restrictions on size, weight and power
  - small, but powerful batteries
- increase in number of portables
  - notebook, laptop
  - games
  - mobile phones
  - mobile DVD players
  - MP3 players
  - GPS

Why less power? TOMORROW

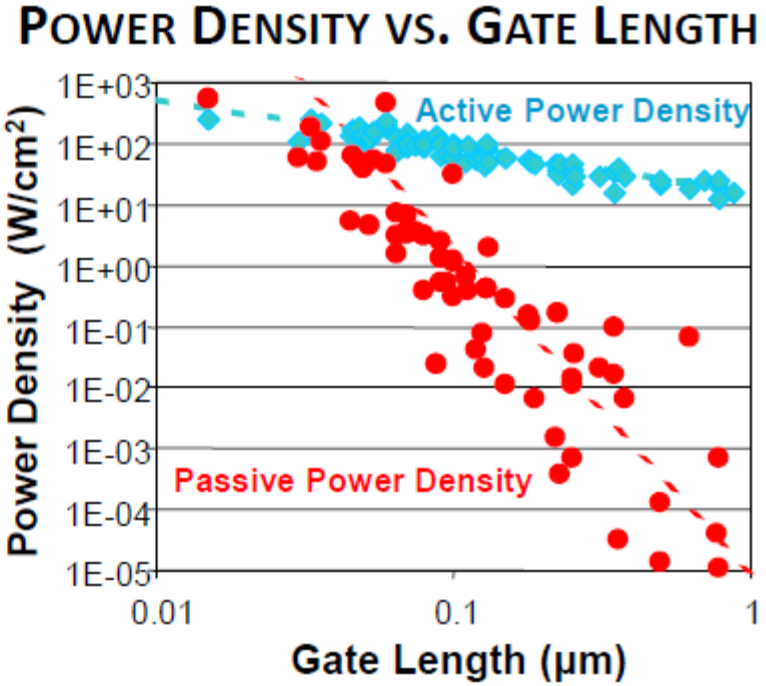


J. Rabaey, *ASPDAC* 2008

Unfortunately the power density is not decreasing with CMOS scaling



Source: P. Packan (Intel), 2007 IEDM Short Course



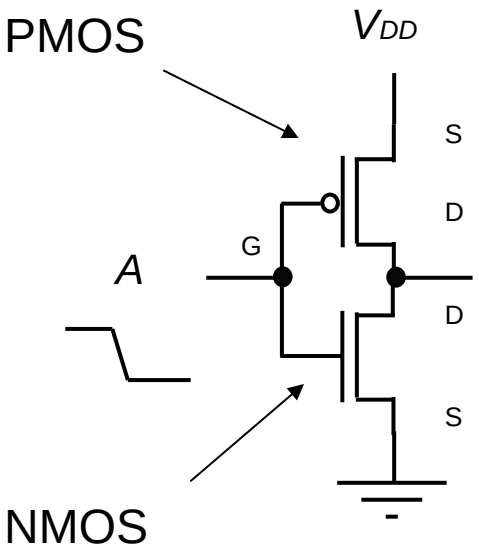
Source: B. Meyerson (IBM) Semico Conf., January 2004

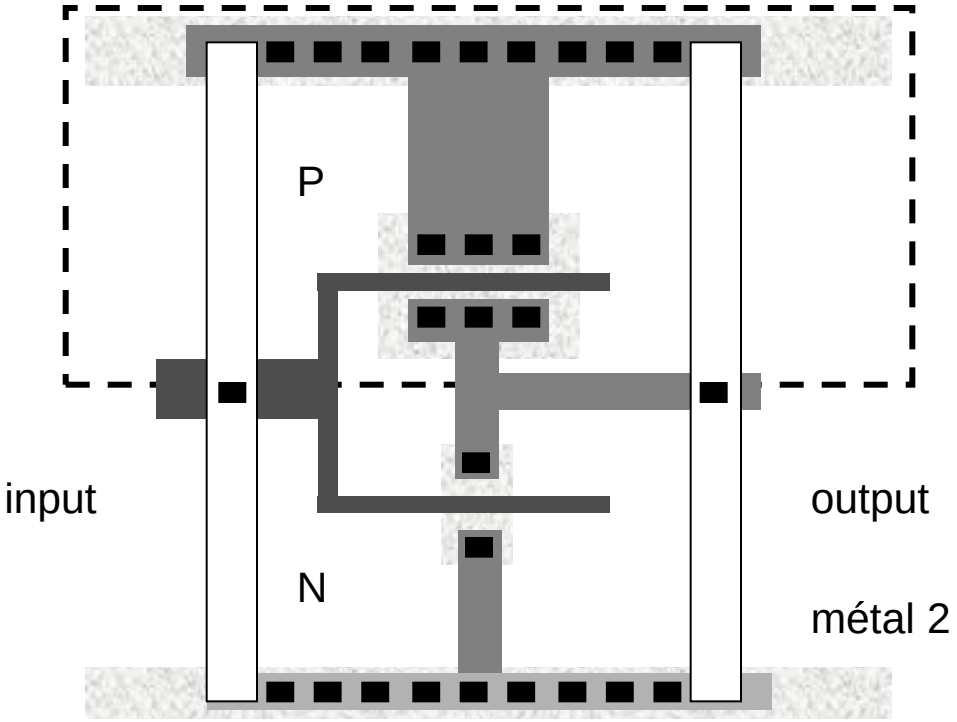
OUTLINE

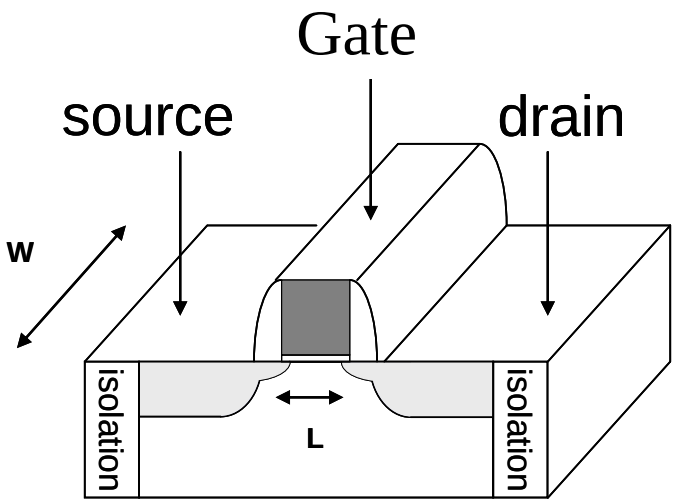
- CMOS: basic elements…
- Power dissipation: static power and dynamic power
- Power optimization at component and circuit level
- Interest of nanotechnologies for energy efficiency
- Adiabatic and reversible computing

OUTLINE

- **CMOS: basic elements…**
- Power dissipation: static power and dynamic power
- Power optimization at component and circuit level
- Interest of nanotechnologies for energy efficiency
- Adiabatic and reversible computing


PMOS
V_{DD}
S
D
G
A
D
S
NMOS


P
input
output
N
métal 2



L around 0,05 micron

W between 0.05 and 0.5 micron (except for analog or buffering)

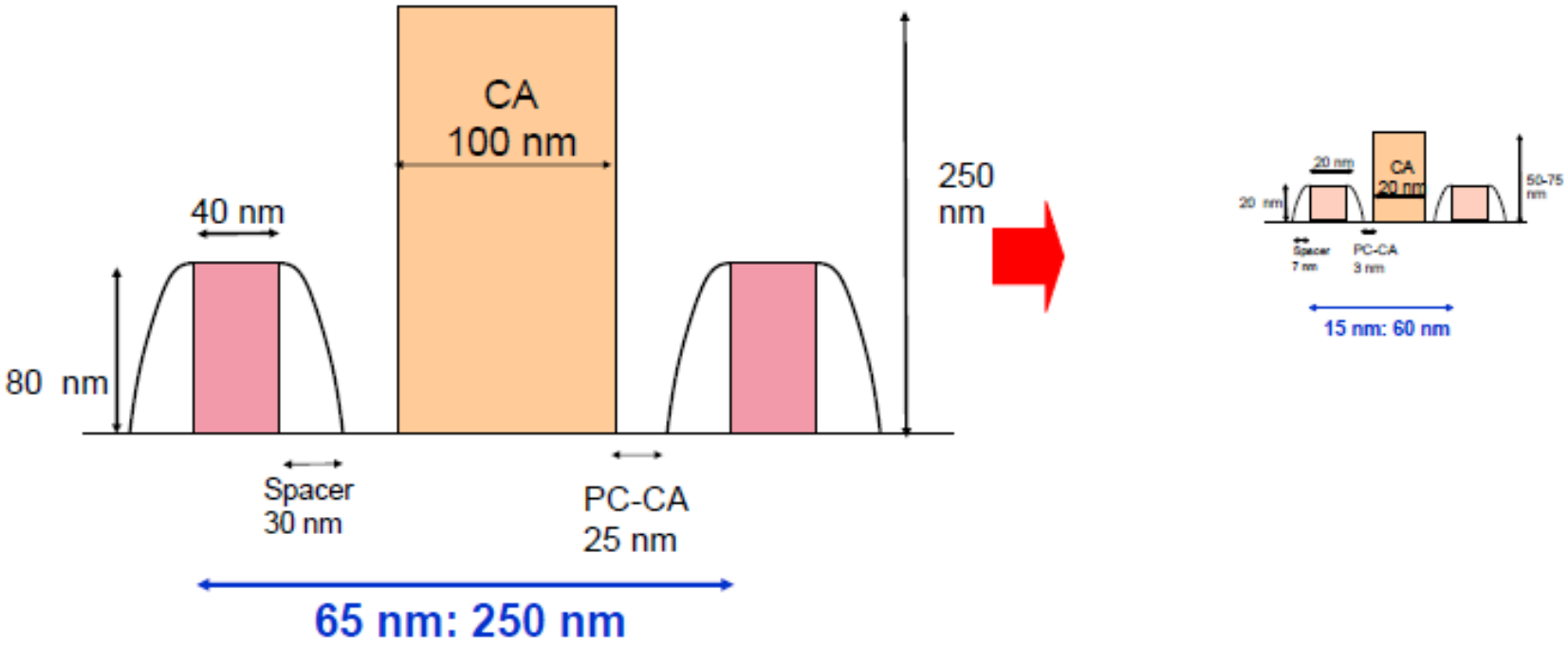
$$I_{on} = WC'_{OX} v_{sat} \left(V_{GS} - V_T \right)$$

v_{sat} saturation velocity of carriers in the conduction channel

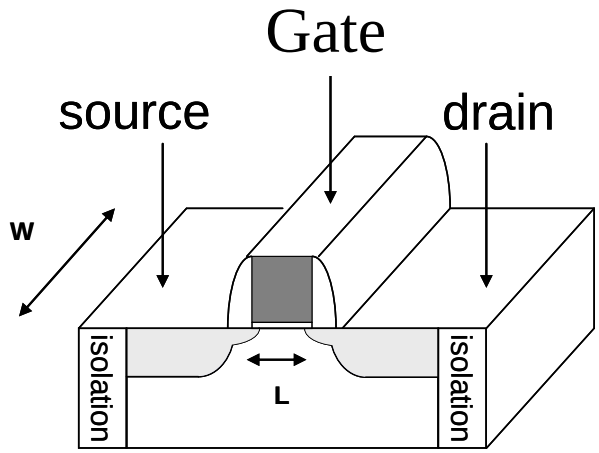
C'_{OX} capacitance per surface unit of the oxide layer

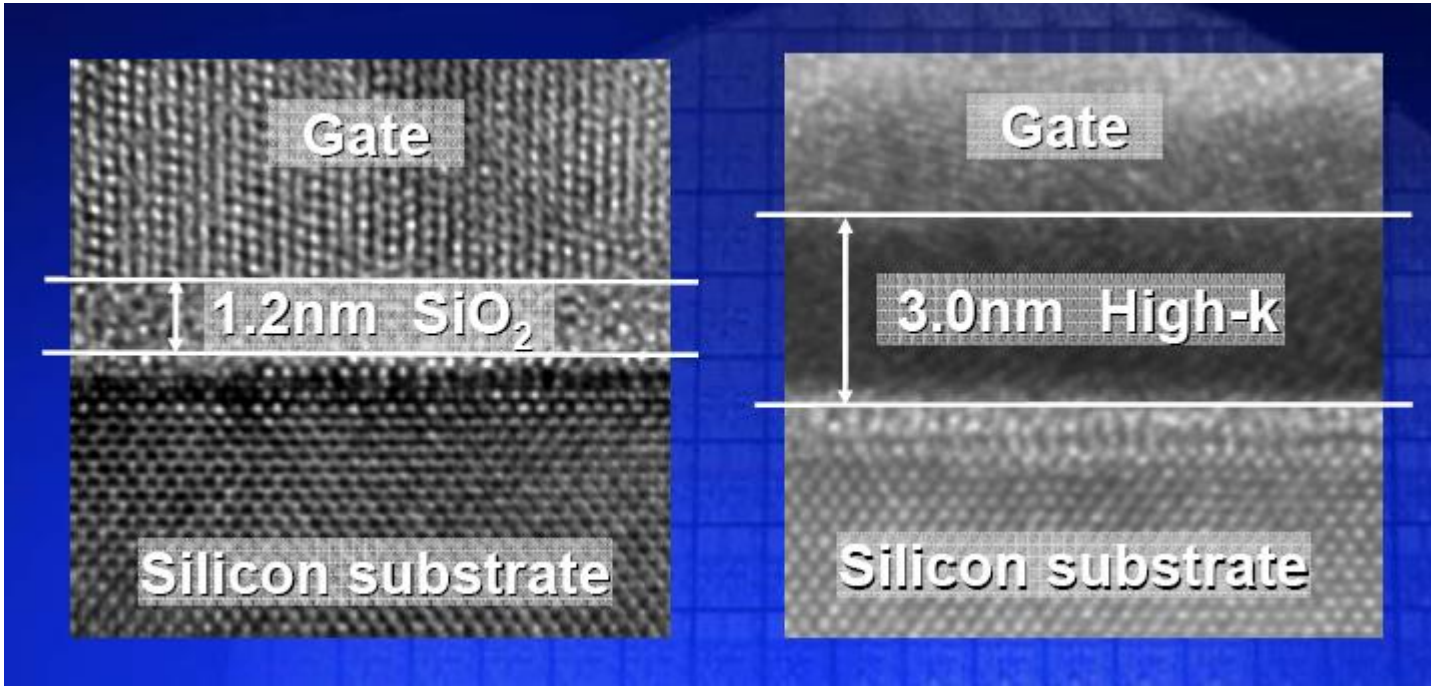
V_T threshold voltage

The microelectronic scaling principle



The MOS transistor is a nanocomponent





Basic FET model

- **The I_{on} current (short channel)**

$$I_{on} = WC'_{OX} v_{sat} \left(V_{GS} - V_T \right)$$

- **The I_{off} current**

$$I_{off} = Wk \exp^{\frac{-V_T}{nV_{th}}} \left(1 - \exp^{-\frac{V_{DS}}{V_{th}}} \right)$$

Strong effect of threshold voltage

Tunnel current in addition

$$V_{th} = \frac{kT}{q}$$

The thermal voltage (25 mV at room temperature)

$$n = 1 + \frac{C_D}{C_n}$$

The body effect factor

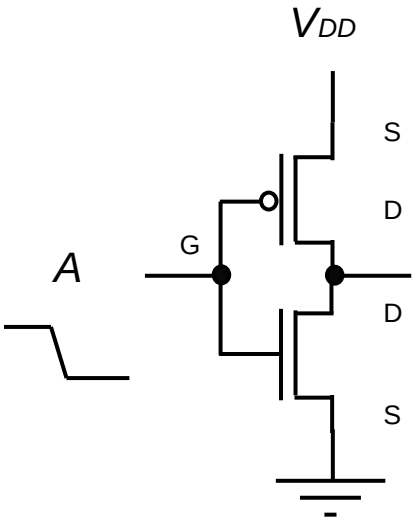
$C_n = WLC'_{OX}$

C_D depletion capacitance

Importance of *Ion* et *Ioff*

- ***Ioff*** **defines leakage power consumption**

- ***Ion*** **defines speed of the circuit**

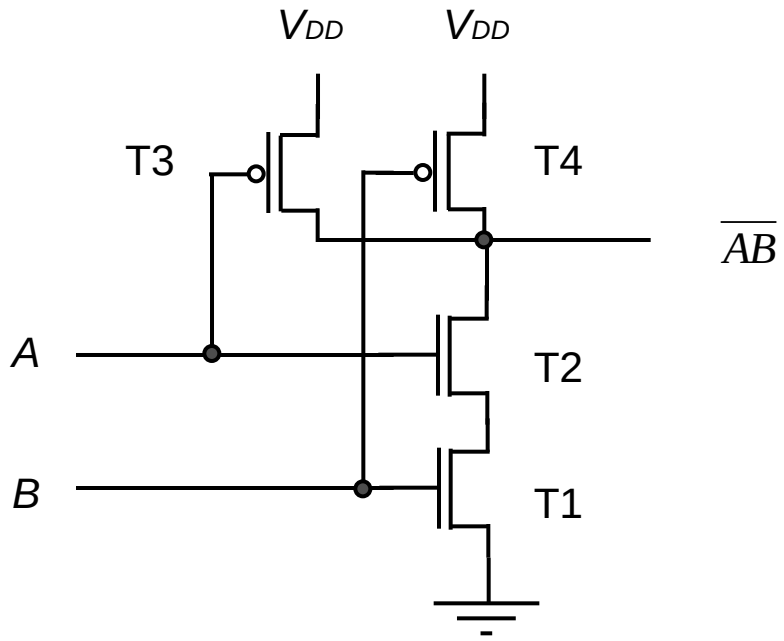


Charging time of the output capacitance

$$i = C\frac{\Delta V}{\Delta t}$$

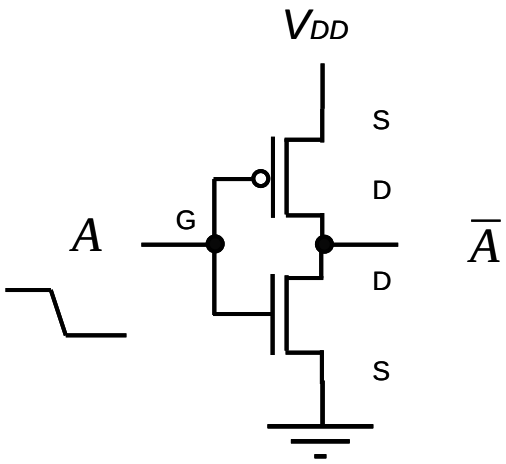
Trade-off between speed and consumption

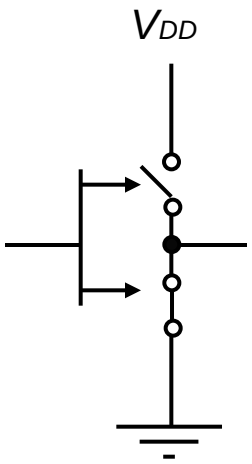
An elementary function (NAND)

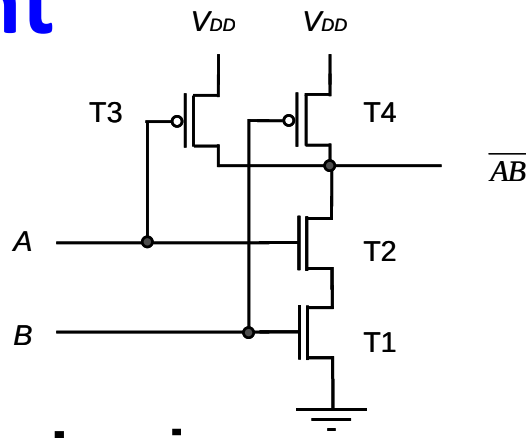


The inverter architecture

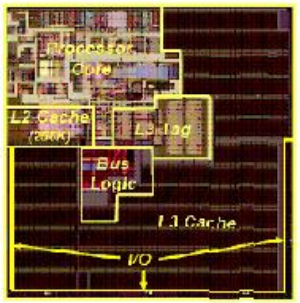
Now systems are switch based systems
Inverter is the basic element

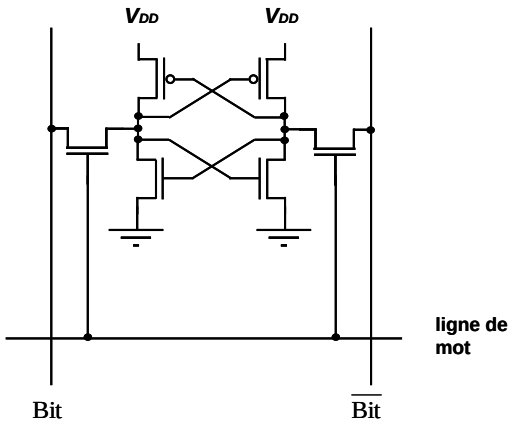






Logic

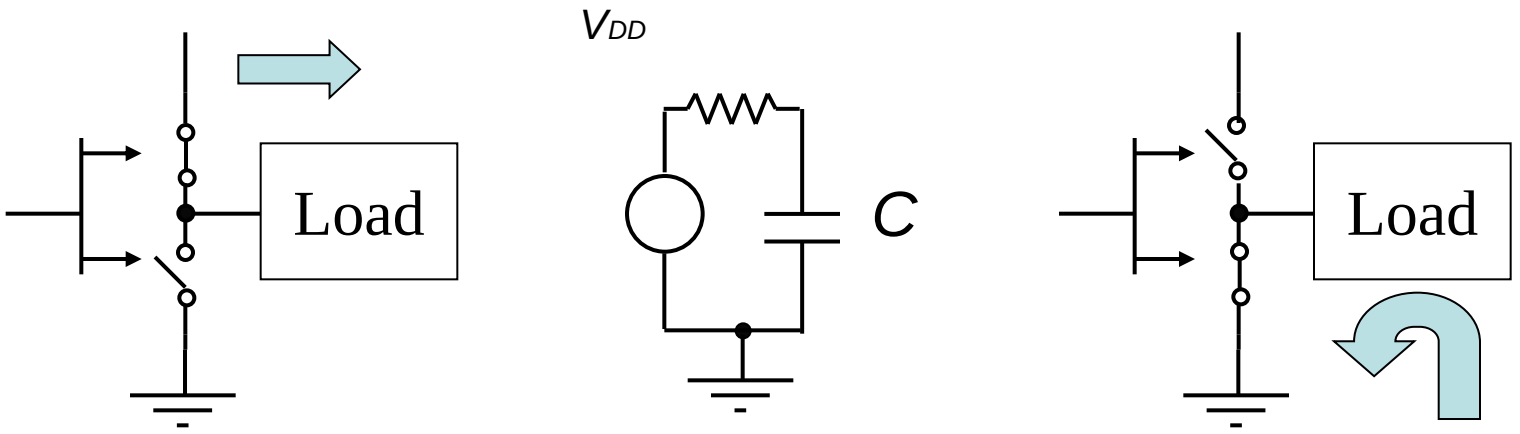




Memory

OUTLINE

- CMOS: basic elements…
- **Power dissipation: static power and dynamic power**
- Power optimization at component and circuit level
- Interest of nanotechnologies for energy efficiency
- Adiabatic and reversible computing



Charge of C

Discharge of C

$$E_{\text{supply}} = V_{DD} \cdot \int_0^{\infty} i(t)dt = V_{DD} \cdot Q = CV_{DD}^{\ 2}$$

$$Q = CV_{DD}$$

$$E_C = \frac{1}{2} CV_{DD}^{\ 2}$$

Dissipated energy

$$E_R = \frac{1}{2} CV_{DD}^{\ 2}$$

Dissipation independent of *R* (switch resistance) and duration of operation

Only true if V_{DD} constant (conventional logic)

Leakage power and dynamic power

$$P_{leak} = I_0 \; V_{DD} \exp^{-\frac{V_T}{nV_{th}}} + V_{DD} \; I_t$$

$$P_{dyn} = \alpha \; f \; C V_{DD}^{\;2}$$

How to reduce leakage power

The subthreshold current

The tunnel current

$$P_{leak} = N \cdot I_0 \; V_{DD} \exp^{-\frac{V_T}{nV_{th}}} + N \cdot V_{DD} \; I_t$$

To use sleep modes

for circuit or blocks

To decrease n

To increase the threshold voltage

To eliminate tunnel effects

How to reduce dynamic power

$$P_{dyn} = \alpha \, N \, f \, C \, V_{DD}^{\,2}$$

To increase parallelism

To reduce interconnect

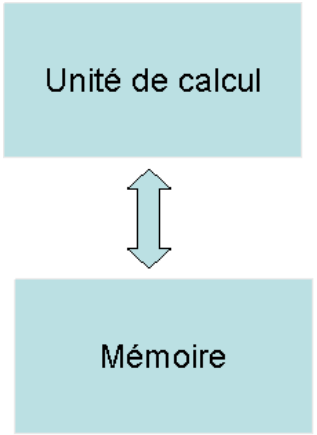
To decrease supply voltage

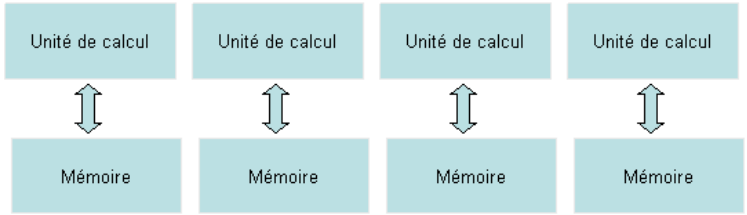
OUTLINE

- CMOS: basic elements…
- Power dissipation: static power and dynamic power
- **Power optimization at component and circuit level**
- Interest of nanotechnologies for energy efficiency
- Adiabatic and reversible computing

Is it possible to reduce the operating frequency ?

Parallelism is an efficient solution



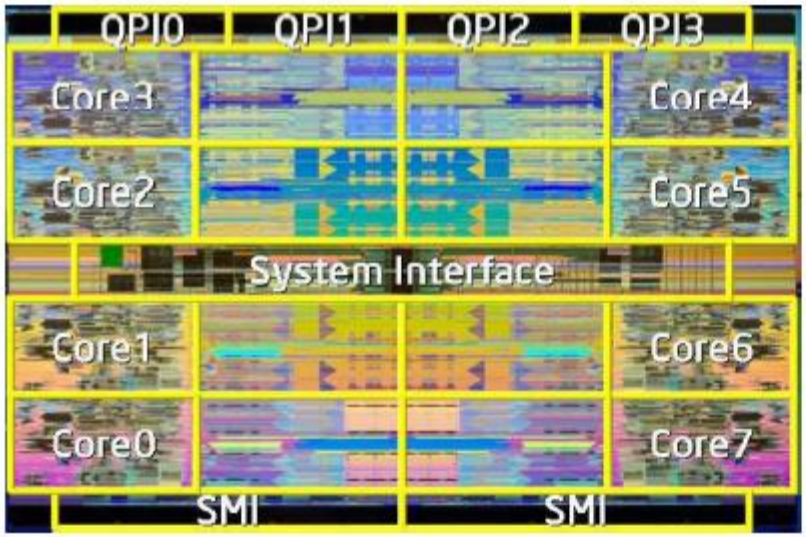


$$P = NfCV^2$$

$$P = 4N\frac{f}{4}\frac{C}{4}\left(\frac{V}{4}\right)^2$$

Today advanced architectures use parallelism but parallelism has limitations

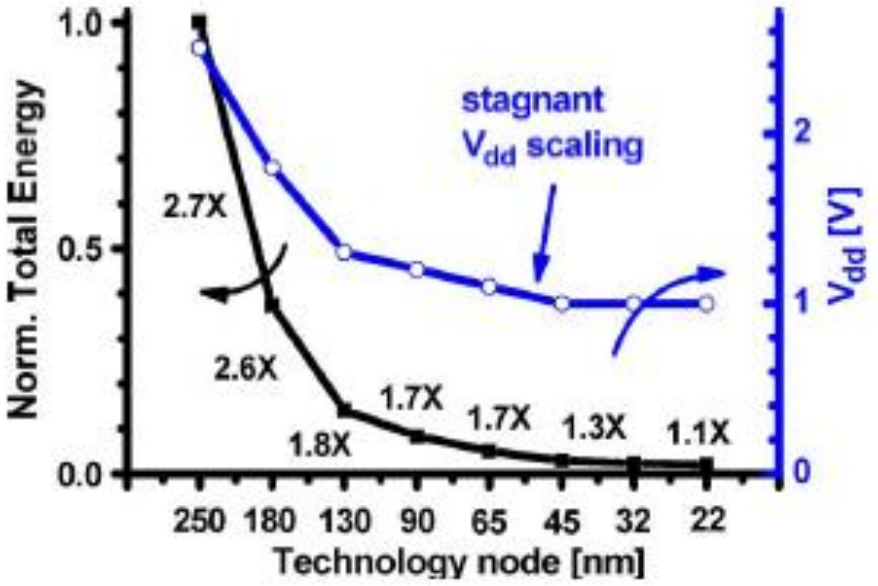
Intel's Nehalem-EX Processor



- 45nm process with Hi-K / Metal Gate
- 8 cores, 16 threads
- 24MB L3 cache
- 4 QPI links
- 2.3B transistors
- 130W TDP

- Power reduction techniques:
  - Operate at the lowest possible voltage (0.85V cores, 0.9V cache)
  - Clock and power gate inactive cores and cache slices
  - Multiple voltage and frequency domains, on-die power management unit
  - Extensive use of long channel devices on paths with timing slack

Is it possible to reduce the supply voltage ?



It is not easy to reduce supply voltage…

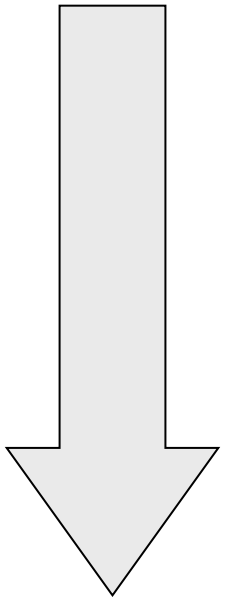
How to reduce supply voltage, the limits

1 Trade-off speed-dissipation

2 Variability and design constraints

3 Cascadability of gates

4 Thermodynamic limit



4 The fundamental limits

- Boltzmann + Shannon : Minimum switching energy

$$E \succ k_B T \ln 2$$

- Heisenberg : Minimum channel length

$$L \succ \frac{\hbar}{\sqrt{2mE}}$$

- With tunnel effect (Zhirnov-Cavin)

$$E \succ k_B T \ln 2 + \frac{\hbar^2 (\ln 2)^2}{8mL^2}$$

What is the minimum value of VDD- summary

- Trade-off Speed-Power

  200-700 mV (less if low frequency operation)

- Variability and design constraints

  100-300 mV

- Cascability of devices

  20-50 mV

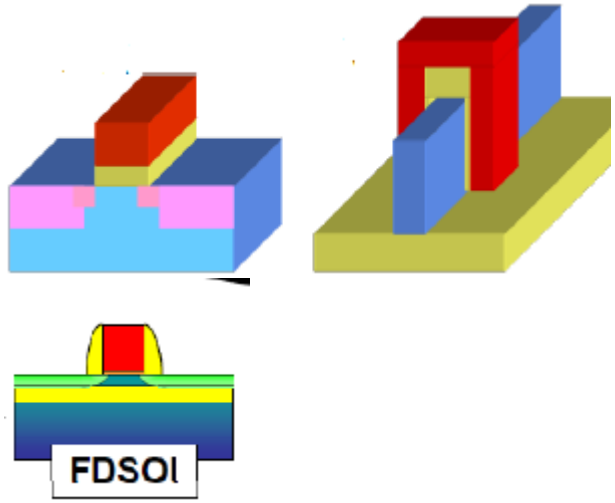
- Fundamental limits

  Below 20 mV towards 2 mV

OUTLINE

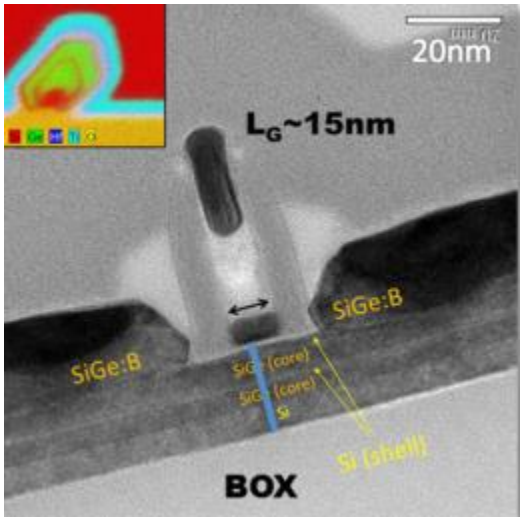
- CMOS: basic elements…
- Power dissipation: static power and dynamic power
- Power optimization at component and circuit level
- **Interest of nanotechnologies for energy efficiency**
- Adiabatic and reversible computing

Planar FDSOI and FinFET are solutions for today



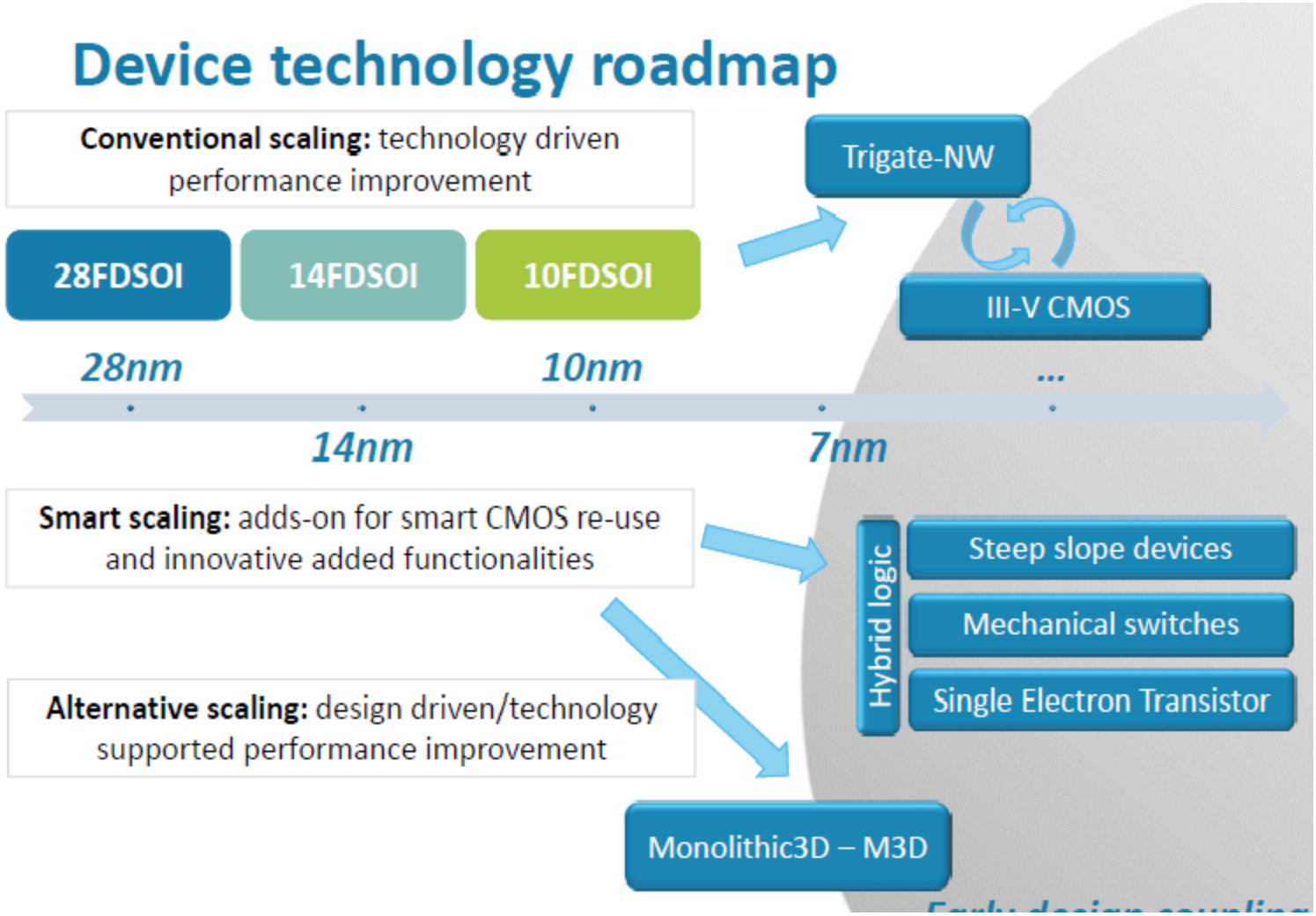
Planar

FinFET

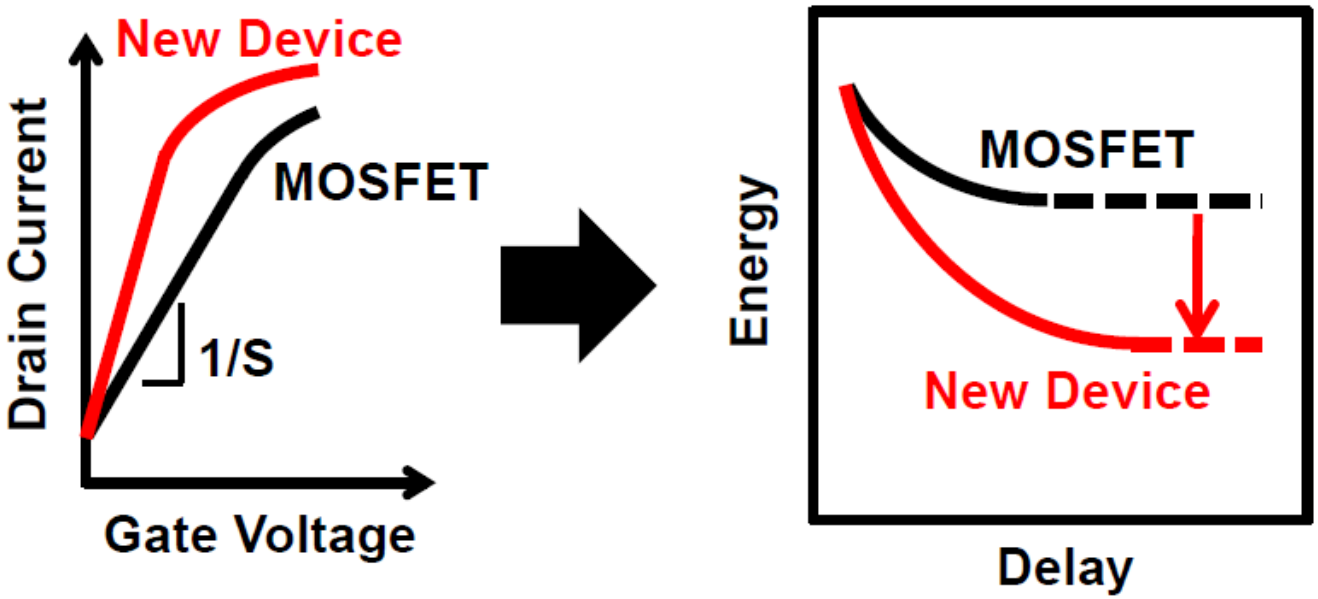


FD SOI transistor

After FinFET and FDSOI

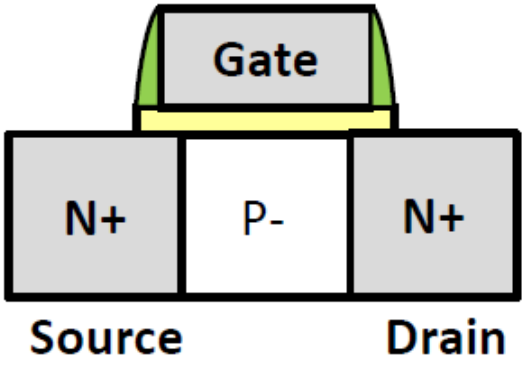


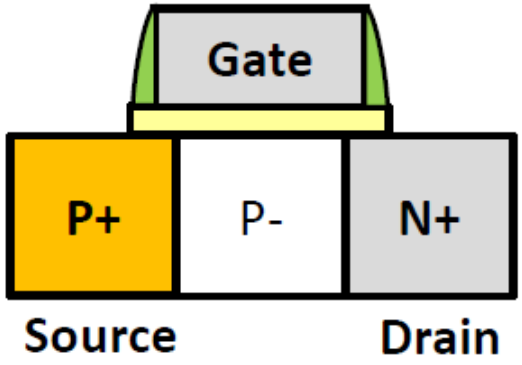
Interest of steep slope devices ?

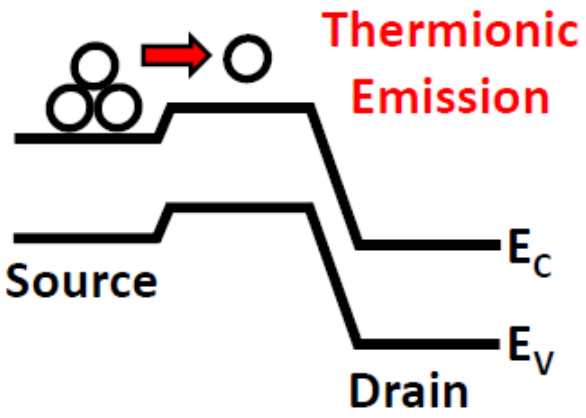


Prof. Tsu-Jae King Liu
Workshop on Zero power technologies for Autonomous Smart Systems

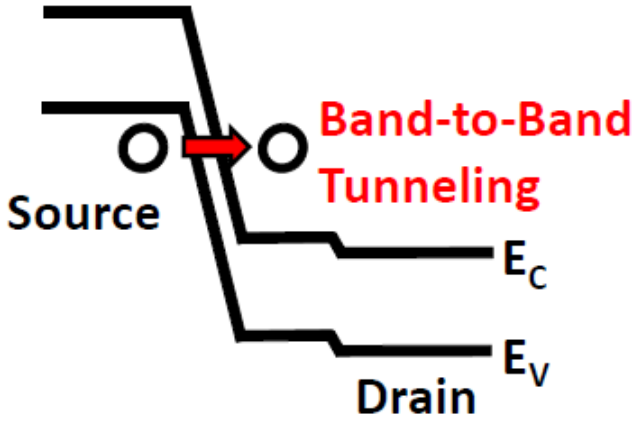
Interest of steep slope devices/ the Tunnel FET







$I_D \propto \exp(qV_{GS}/nkT)$

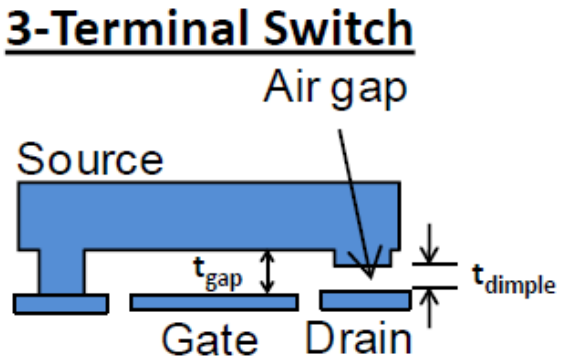


$I_D = AE_S\exp(-B/E_S)$

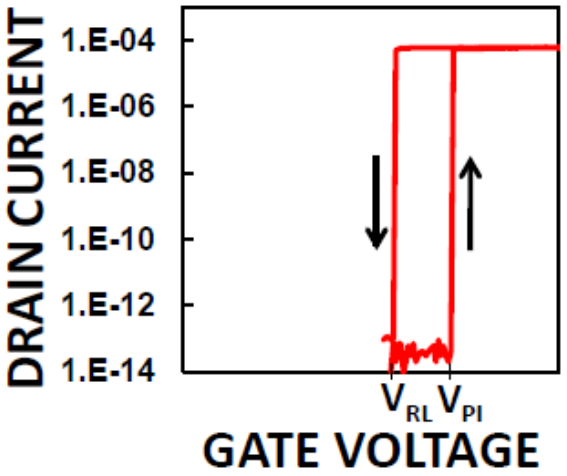
(E_S = electric field)

Interest of steep slope devices/ the nanomechanical relay

- **Zero off-state leakage**
  → zero leakage energy

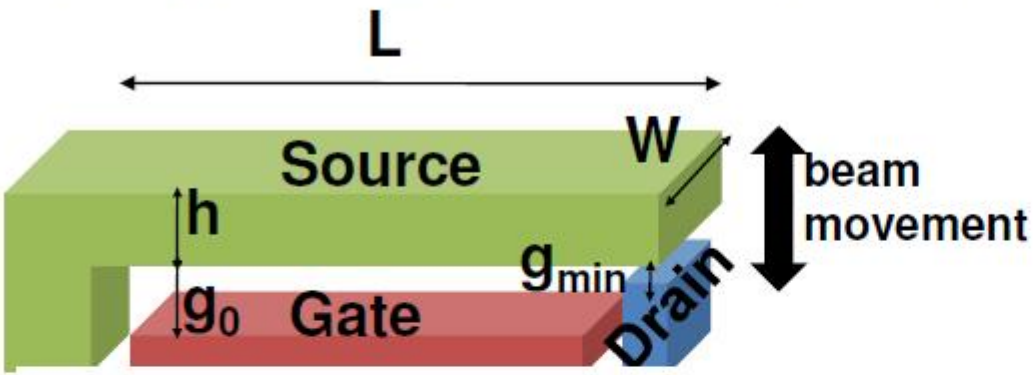


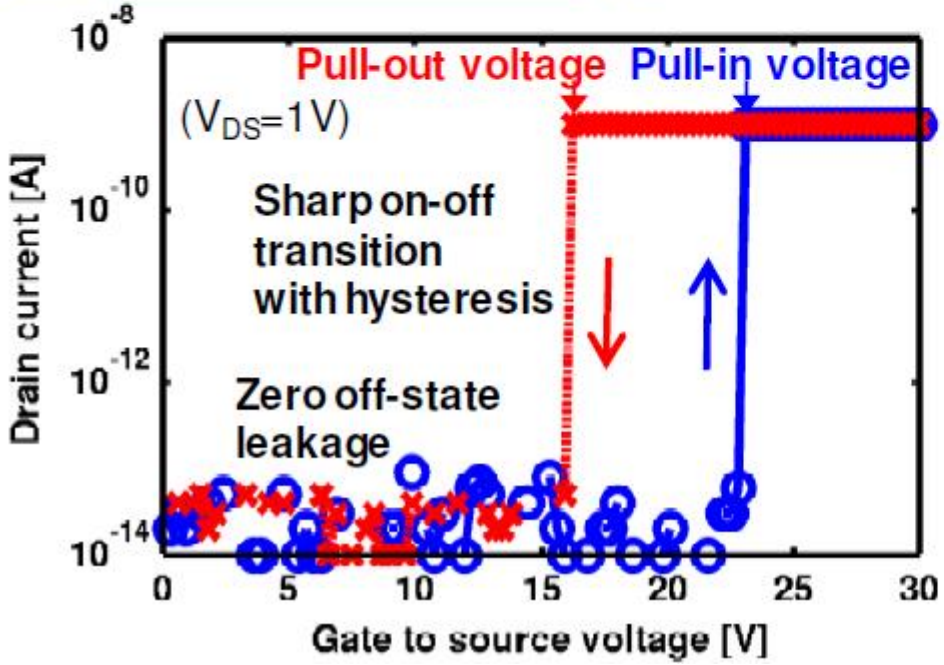
- **Abrupt switching behavior**
  → allows for aggressive V_{DD} scaling (ultra-low dynamic energy)



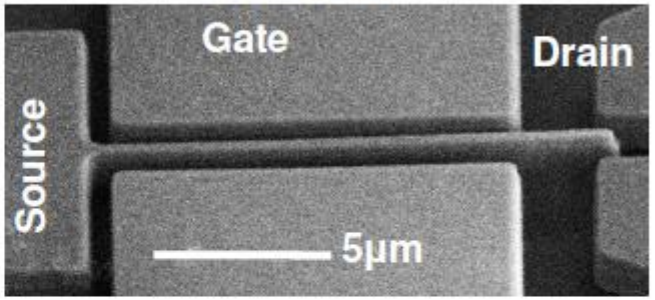
Prof. Tsu-Jae King Liu
Workshop on Zero power technologies for Autonomous Smart Systems

Nanoelectromechanical (NEM) Relay

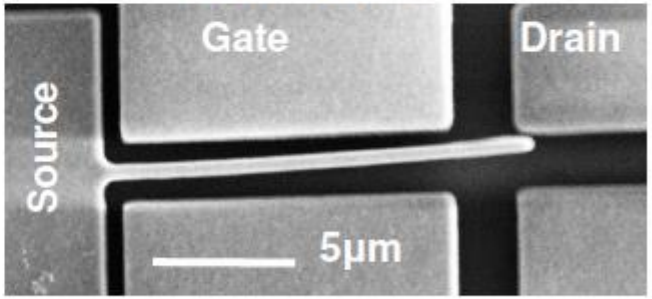




Off-state



On-state



S. Chong …H.-S. P. Wong, "Nanoelectromechanical (NEM) Relays Integrated with CMOS SRAM for Improved Stability and Low Leakage," *ICCAD* 2009.

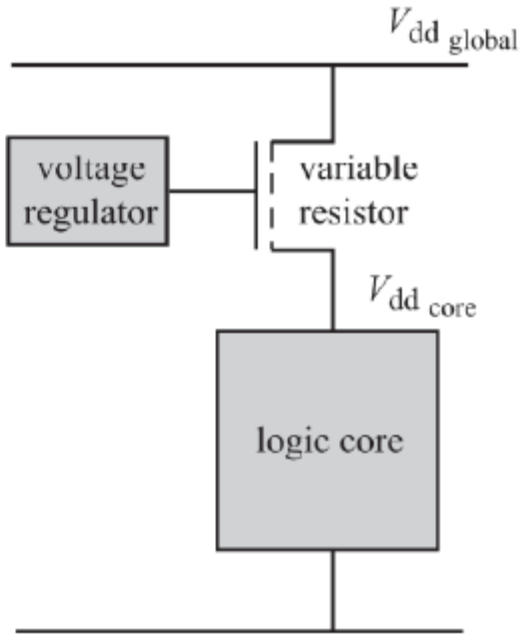
- **NEM switches have several potential advantages compared with nanoscale CMOS transistors:**
  - low power supply (100 mV)
  - high on current
  - 'zero' leakage
  - 'infinite' subthreshold slope
  - high temperature operation
  - radiation-hard operation
  - compatible with other substrates … glass, plastics

- **… and "just a few" problems:**
  - reliability (stiction, contact degradation, wear, …)
  - slow speed

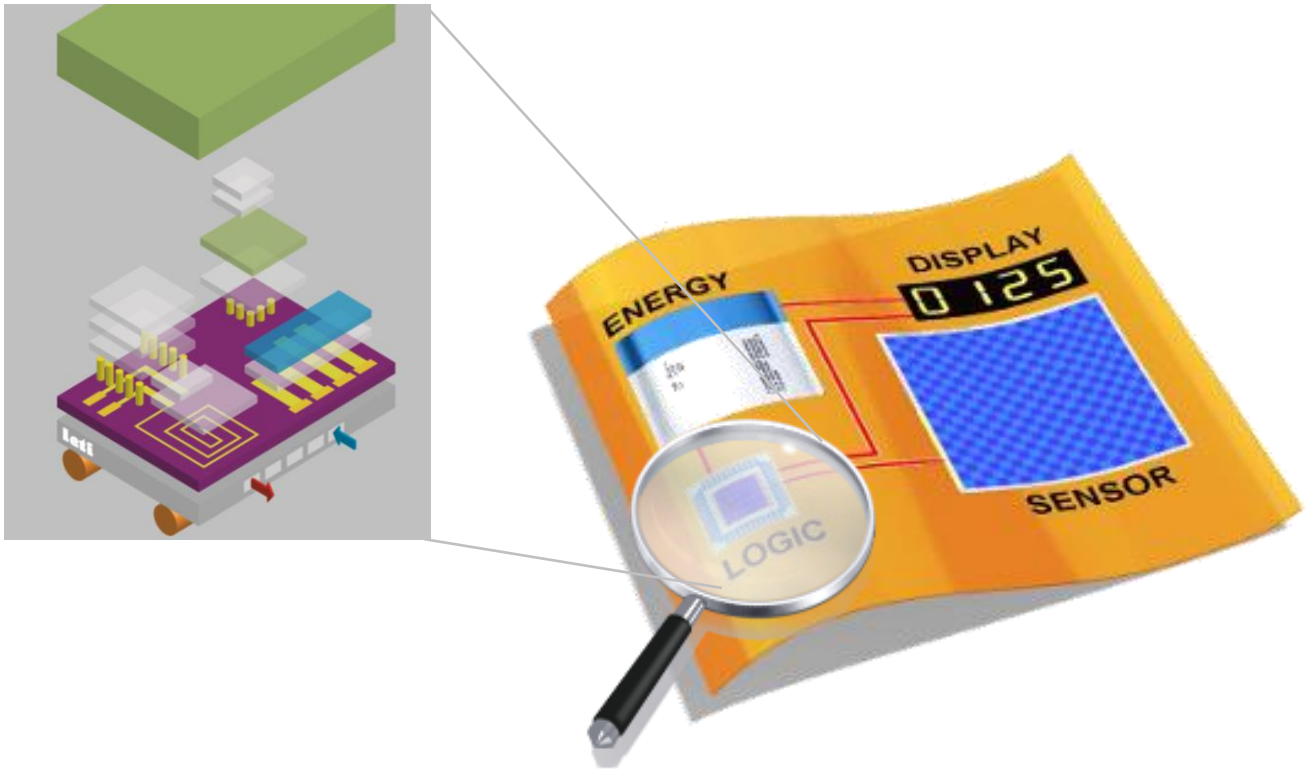
Optimization at system level

dynamic voltage and frequency scaling (DVFS)

- dynamically adjusting the voltage to the performance needs
- requires on- or off-chip voltage regulator
- varying the resistance of a big transistor between the core and the supply voltage
- but power reduces only linearly with the V_{ddcore} instead of quadratically, since part of the power saving in the core is now consumed in the variable resistor.

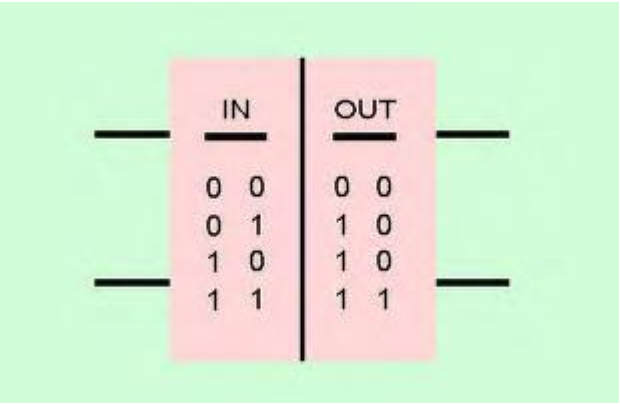


Interest of 3D and heterogeneous technologies



- **CMOS: what is necessary to know…**
- Power dissipation: static power and dynamic power
- Power optimization at component and circuit level
- Power dissipation at architecture level
- **Adiabatic and reversible computing**

Adiabatic logic	Reversible logic
How to charge and discharge a capacitance with minimum energy dissipation To work at constant current	How to take into account the Landauer principle in order to minimize energy dissipation To use only reversible logic

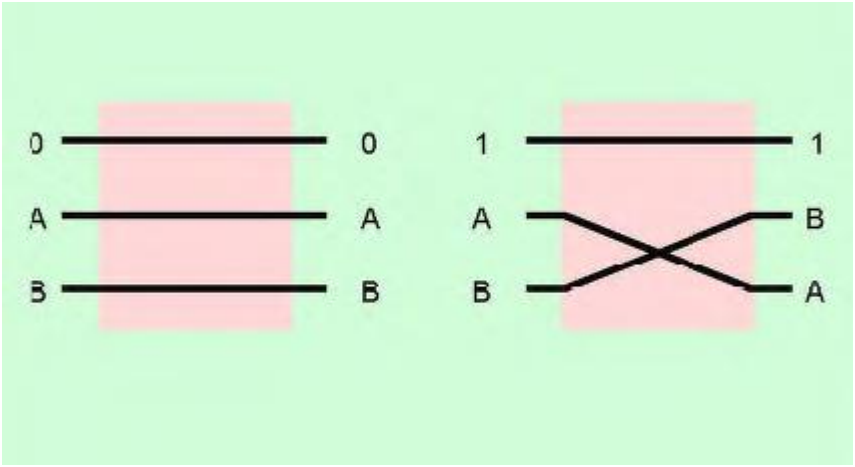


$$H_{in} = -\sum_{states} p_{state} \log_2 p_{state} = -4 \times \frac{1}{4} \log_2 \frac{1}{4} = 2$$

$$H_{out} = -2 \times \frac{1}{4} \log_2 \frac{1}{4} - \frac{1}{2} \log_2 \frac{1}{2} = \frac{3}{2}$$

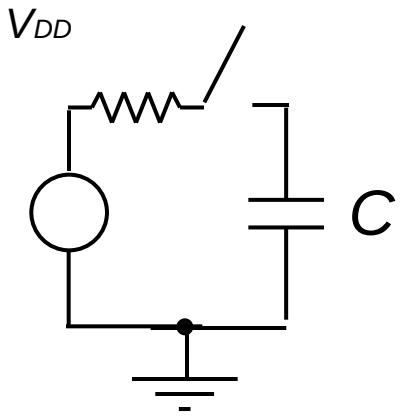
An example of non reversible gate

Impossible to guess inputs from outputs



An example of a reversible gate: the Fredkin Gate

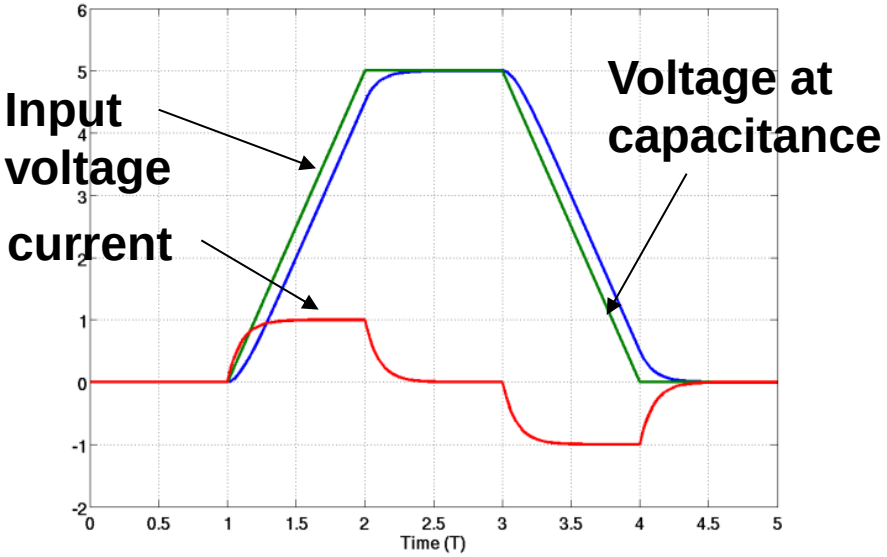
The adiabatic charge of a capacitance



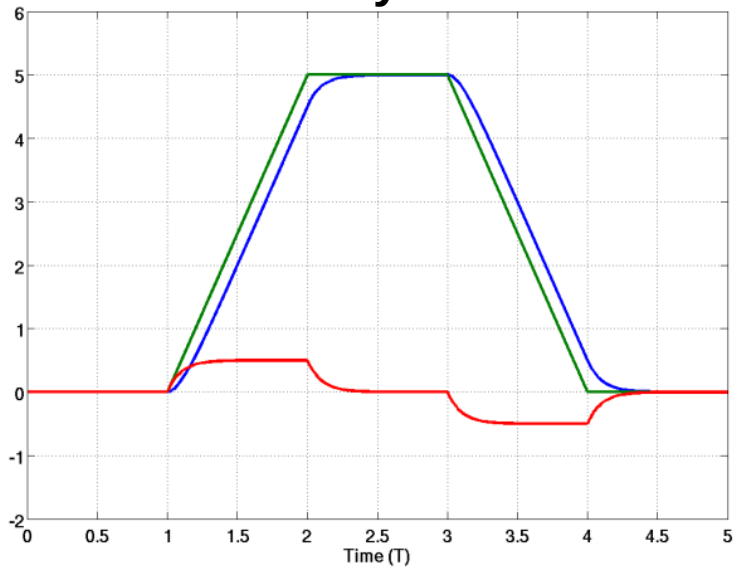
Not an abrupt variation but a linear variation is applied to the capacitor

If *RC* is small compared to the duration *T* of the ramp, charge-discharge currents and Joule dissipation are reduced by RC/T

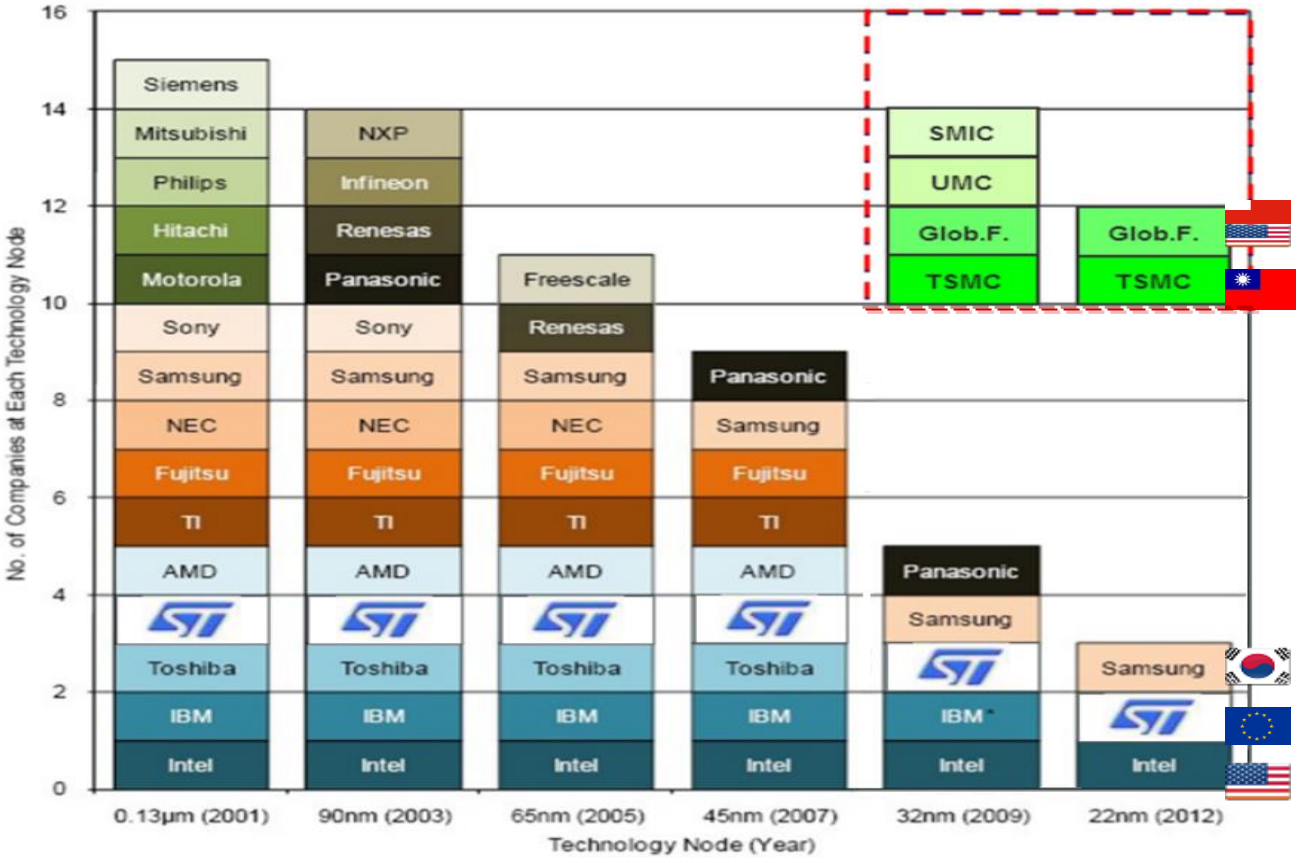
***RC/T* small**



***RC/T* very small**



Nanoelectronics is not only a technical issue



Foundries

Integrated Device Manufacturers

*Source IBS – IDMs - ** IBS fait la supposition que la partie manufacturing d'IBM sera absorbée par Global Foundries*

Conclusions

- Energy efficiency is the main nanoelecronic driver…
- Many improvements at circuit level
- New switches (Tunnel FET and NEMS) have to be confirmed for future ultra low power electronics but FDSOI and FinFETs are solutions for today
- Long term solutions are not defined
- Adiabatic logic in association with new switches has to be investigated

Bibliography

- **Low power electronics design**

  Christian Piguet

  CRC PRES

- **Proceedings of the IEEE**

February 2010 vol 98 Number 2

Circuit technology for ultra low power