Implications on the Design

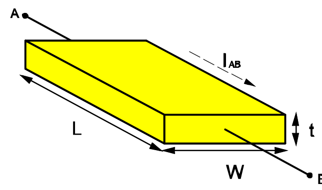
Ramon Canal
CTD – Master CANS

Agenda

- VLSI Basics
  - Resistance:
  - Capacity:
- Energy Consumption
  - Static
  - Dynamic
  - Thermal maps
- Voltage Scaling
- Metrics

Resistance

- In general: $R_{AB} = \rho \frac{L}{tW}$ $\quad I_{AB} = \frac{V_{AB}}{R_{AB}}$

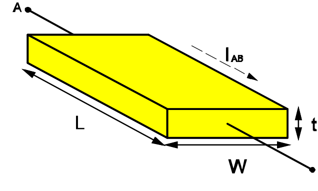


Material	ρ(Ω-m)
Silver (Ag)	1.6×10^{-8}
Cupper (Cu)	1.7×10^{-8}
Gold (Au)	2.2×10^{-8}
Aluminium (Al)	2.7×10^{-8}
Tugnsten (W)	5.5×10^{-8}

- For a given technology
  - *t* is constant
  - Basic unit of measurement: Square resistance (L=W=2λ): ❑, R_s

Resistance

- In general: $R_{AB} = \rho \frac{L}{tW}$ $\quad I_{AB} = \frac{V_{AB}}{R_{AB}}$

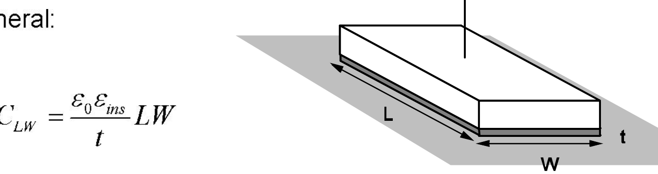


Material	Sheet Res. (Ω/❑)
n, p well diffusion	1000 to 1500
n+, p+ diffusion	50 to 150
n+, p+ diffusion with silicide	3 to 5
polysilicon	150 to 200
polysilicon with silicide	4 to 5
Aluminum	0.05 to 0.1

- For a given technology $\qquad R_s = \frac{\rho}{t}$
  - *t* is constant
  - Basic unit of measurement: Square resistance (L=W=2λ): ❑, R_s

Capacity

- Directly depending on the ...
- In general:



$$C_{LW} = \frac{\varepsilon_0 \varepsilon_{ins}}{t} LW$$

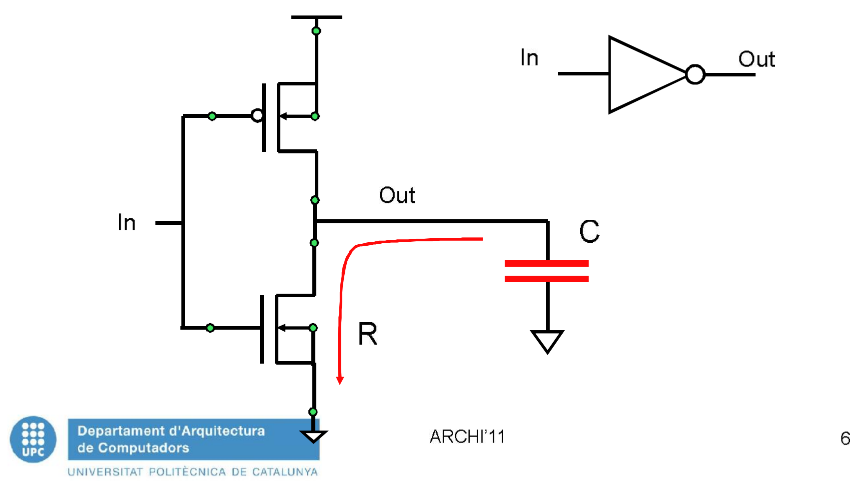
- For a given technology:
  - The permitivity ε_{ins} is constant
  - t is constant
  - Basic unit of measurement: (L=W= 2λ): Cg

$$C_g = \frac{\varepsilon_0 \varepsilon_{ins}}{t}(2\lambda \times 2\lambda) \rightarrow C_{LW} = \frac{LW}{4\lambda^2} C_g$$

Departament d'Arquitectura de Computadors
UNIVERSITAT POLITÈCNICA DE CATALUNYA

MOS gate delay

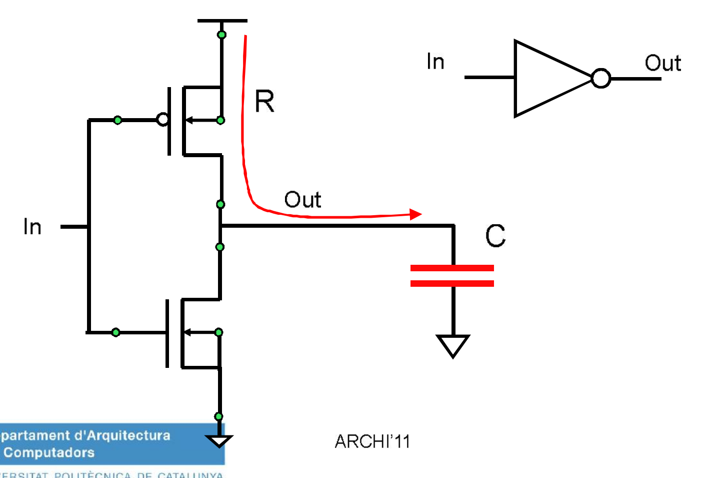
- Delay depends on RC relationship:



Departament d'Arquitectura de Computadors
UNIVERSITAT POLITÈCNICA DE CATALUNYA

MOS gate delay

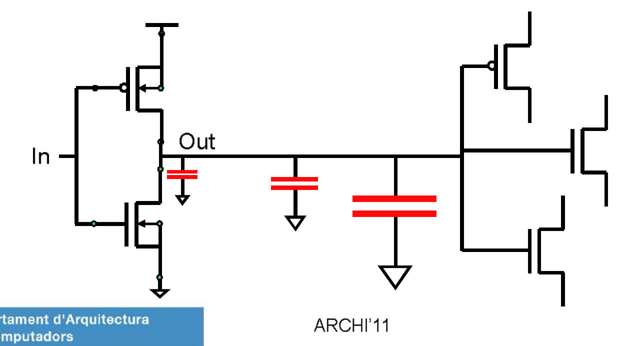
- Delay depends on the RC relationship:



Departament d'Arquitectura de Computadors
UNIVERSITAT POLITÈCNICA DE CATALUNYA

MOS gate delay

- The capacity C depends on:
  - Capacity of the difussions of the output nodes
  - Capacity of the connections
  - Capacity of the nodes connected to the output



Departament d'Arquitectura de Computadors
UNIVERSITAT POLITÈCNICA DE CATALUNYA

MOS gate delay

INV device model

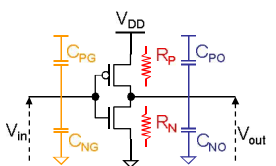
- R_N, R_P : ON-resistance of nMOS / pMOS
  - → Proportional to L_g / W
- C_{NG}, C_{PG} : Gate capacitance of nMOS / pMOS
  - → Proportional to $L_g \cdot W$
- C_{NO}, C_{PO} : Drain capacitance of nMOS / pMOS
  - → Proportional to $L_d \cdot W$ and $L_d + W$
- ➤ *In reality, R_N, R_P, C_{NO}, C_{PO} are not constant, but dependent on the voltages at the gate and drain*

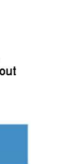
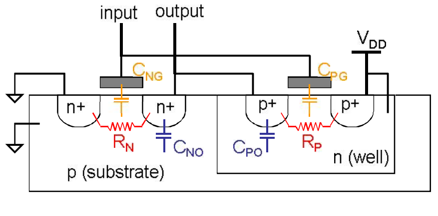
MOS layout

L_g

W s g d

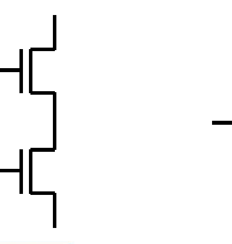
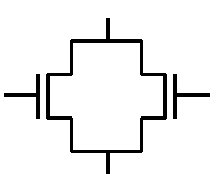
L_d



MOS gate delay

- The resistance R depends on:
  - Dimensions of the transistors that charge/discharge the output capacity (~ L/W).
  - Transistors in a row add resistance: L_1/W_1 L_2/W_2
    - $R \sim L_1/W_1 + L_2/W_2$
  - Transistors in parallel reduce resistance: L_1/W_1 L_2/W_2
    - $R \sim 1 / \left(1/(L_1/W_1) + 1/(L_2/W_2)\right)$

Modular Characterization

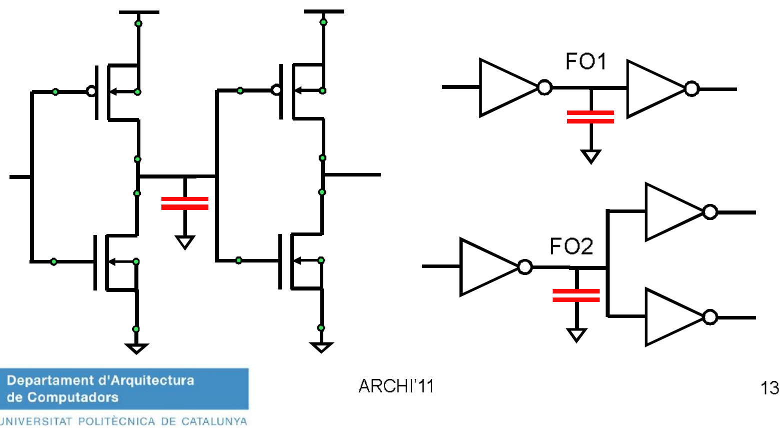
- It is impossible to analyze a million transistor circuit looking at every single transistor
- Modular approach makes it feasible
- Grouping transistors in gates, gates in blocks and so on.
- Keys to the analysis of MOS circuits:
  - The behavior depends on the input values.
  - Interface parameters: capacity of the inputs/outputs and impedance of the outputs
  - Transmission gates are an exception.

Modular Characterization

- Elements that define a module:
  - Logic function
  - Capacity of the inputs (~ gate capacity of the transistors)
  - Capacity of the outputs (~ drain capacity of the output transistors)
  - Impedance of the outputs
  - Internal propagation time
- Capacity and impedance are fixed parameters
- Internal propagation time depends on the inputs:
  - Characterize the highest/lowest times
  - Characterize depending on functionality (e.g. ALU), statistical or sampling (e.g. FU inputs), etc.

MOS gate delay

- FO1 (Fan Out of 1) is defined as the delay of an inverter connected to another inverter of equal sizes
  - FO2 is equivalent to connecting two inverters, etc.

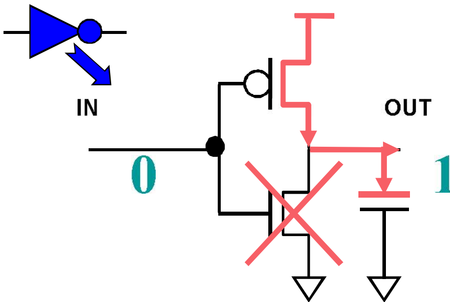


Power

- Two types:
  - Dynamic: proportional to the activity
  - Static: “just for being there”

Dynamic Power

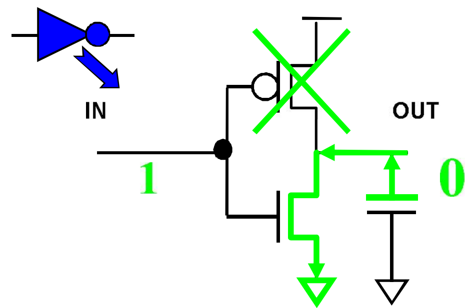
- Charging and discharging capacitors



- The capacitors are the transistors/buses connected to the output

Dynamic Power

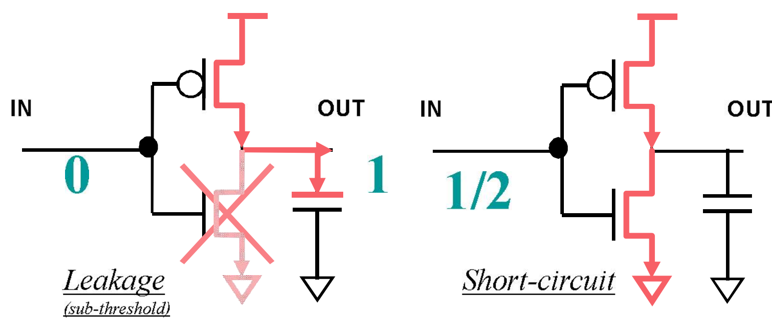
- Charging and discharging capacitors



- The capacitors are the transistors/buses connected to the output

Static Power

- Short-circuit and leakage currents of the transistor



- Leakage is growing dramatically
  - ~7% (0'25µm), ~20% in (130nm) process technology, ~50% in 65nm

Power and Energy

Energy

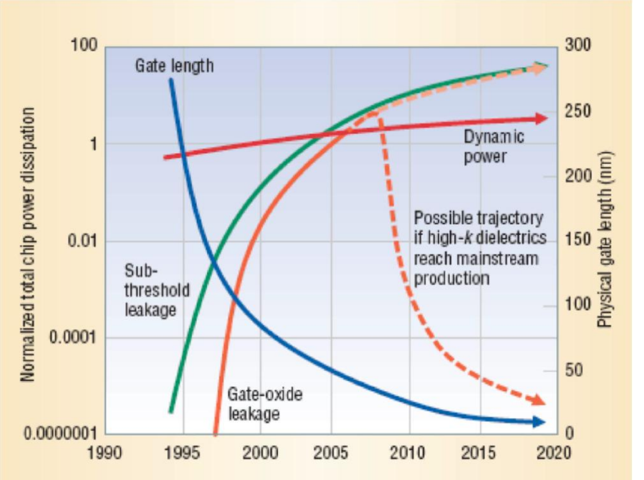
- *"the potential for causing changes"*
  - Mesured in Joules
- Important for
  - Battery duration – less need of energy ➔ longer life
  - Electricity bill – less need of energy ➔ saving $$

Power and Energy

Energy

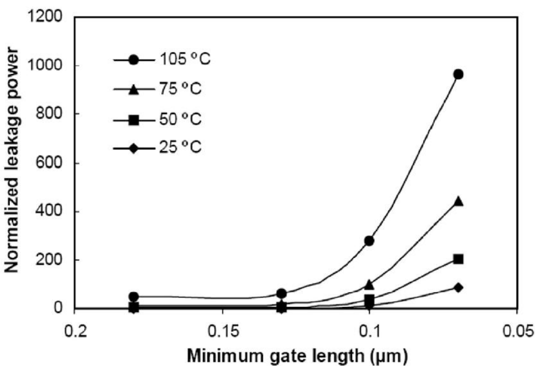
- Dynamic energy (CMOS):
  - Proportional to the activity, the capacitance and the square of the supply voltage ($E= \alpha CV^2$)
- Static energy (CMOS):
  - Short-circuit: $E_{sc}=t_{sc}\, V_{DD}\, I_{peak}\, P_{0\rightarrow 1}$
  - Leakage: $E_{leakage}=V_{DD}\, I_{leakage}$

Importance of leakage



Kim et al. "Leakage Current: Moore's Law Meets Static Power", IEEE Computer, 2003.

Thermal effects on leakage



The circuit simulation parameters including threshold voltage were obtained from the Berkeley Predictive Spice Models. The leakage power numbers were obtained by HSPICE simulations.

Flautner et al. "Drowsy Caches: Simple Techniques for Reducing Leakage Power", ISCA 2002

Power and Energy

Power

- *"amount of work done per unit of time"*
  - Measured in Watts
- Important for:
  - Higher power ➔ Higher currents (I)
    - Processors are limited (power delivery constraints)
  - Higher power ➔ Higher temperature
    - Processors are limited again (power envelop)

Power and Energy

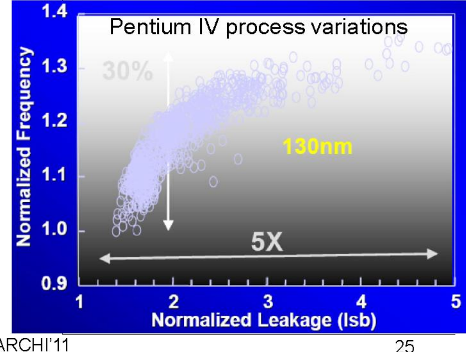
- CMOS
  - $P_{dynamic} = \alpha C V^2 f$
    (α: activity factor, C: capacitance, V: voltatge, f: frequency)
  - $P_{static} = nIV$
    (n: number of transistors *off*, I: leakage current, V: voltatge)

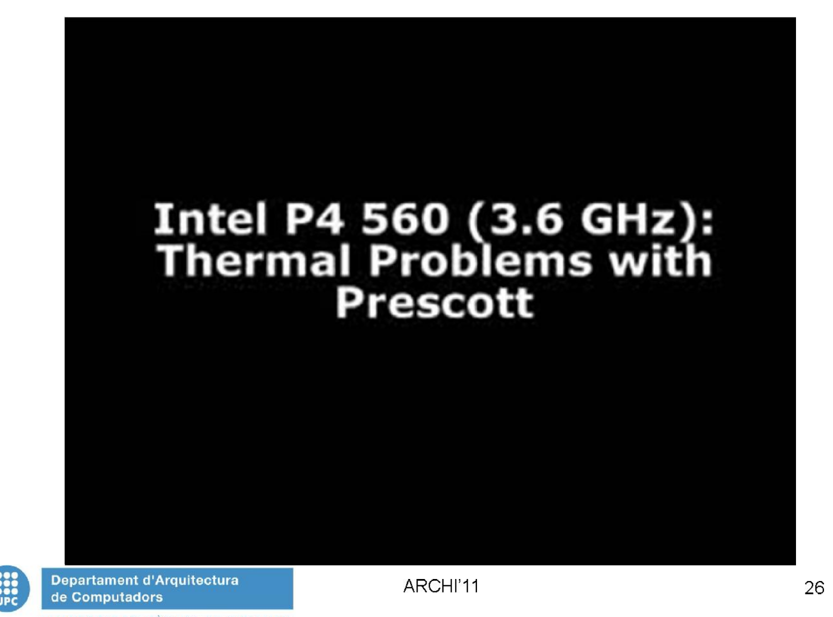
Power Envelop / Thermal Design Power

- **CPU "power envelop"**
  - Maximum power the cooling mechanisms can dissipate.
- **Limitted by**
  - System power budget
  - Processor power budget
  - Usually it is around:
    Server <140W, Desktop 50-100W, Laptop 20-35W, PDA <10W
- **Larger systems can afford large cooling mechanisms such as**
  - *Heat sinks*
  - *Heat pipes*
  - Better TIM (*Thermal Interface Materials*)
- **Power distribution is important:**
  - The most uniform the distribution is the better the dissipation capabilities

Voltage Scaling

- **Given a voltage range, the higher the voltage the higher the frequency the circuit can opperate**
- **Good approach to trade-off power and frequency.**
  - Statically, when manufacturing the circuit
  - Dynamically, while the circuit is opperating
    - Intel's SpeedStep® Technology
    - Transmeta LongRun
    - AMD PowerNow, Cool'n quiet
- **The voltage/frequency range depends on the technology used.**





Voltage Scaling (cont.)

- **Impact on power consumption:**
  Frequency 20% ↓➔ ↓ 20% voltage
  ➔ 35% energy reduction ($\alpha CV^2 = \alpha C(0.8V)^2 = \alpha CV^2 \cdot 0.64$)
  ➔ 50% power reduction ($\alpha CV^2 f = \alpha CV^2 f 0.8^3 = \alpha CV^2 f \cdot 0.51$)
- **Minimal impact on performance:**
  Frequency 20% ↓ ➔ performance 10%-15% ↓*
- ***Voltage scaling* leverages energy consumption and performance**

* Depending on the cache size and the ratio between the chip and off-chip buses.

Leveraging power and performance

Voltage scaling **may be not enough to reduce power**

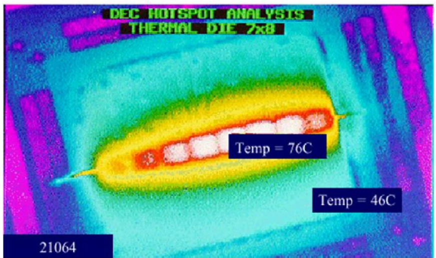
We need better:

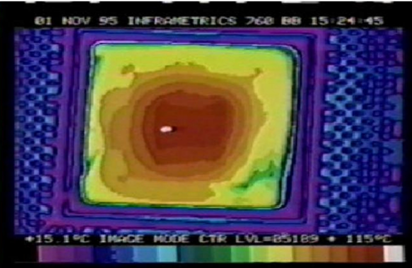
1. **Designs (Computer Architecture)**
   - **Power-Aware designs**
     - **Energy*Delay (EDP) metrics**
2. **Technology (physics)**

Alpha hot spots

21064 Thermal Plot

21164 Thermal Plot

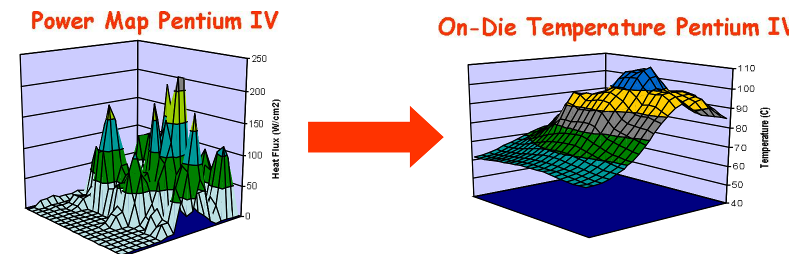




	Power (Watts)	Freq. (MHz.)	Die Size (mm^2)	Vdd
Alpha 21064	30	200	234	3.3
Alpha 21164	50	300	299	3.3

Source - CoolChips-99

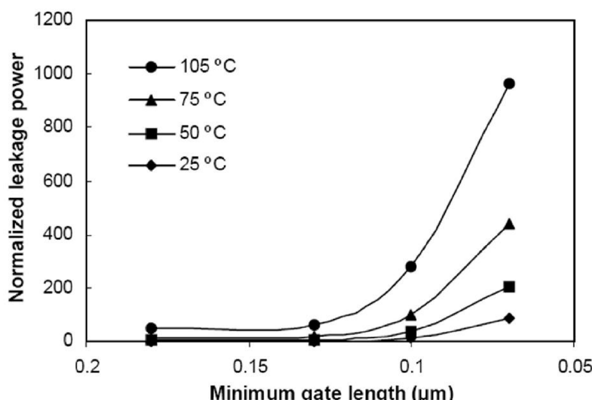
Hot Spots and thermal problems



- Silicon is not a good heat conductor
- Temperature ↑ ➔ *leakage* ↑ ➔ power ↑➔ temperature ↑
- A big circuit does not help. *Hot* spots must be reduced.
- With a good layout, hotspots can be put apart and thus reducing the heat dissipation needs –*power envelope*.

* "New Microarchitecture Challenges in the Coming Generations of CMOS Process Technologies" – Fred Pollack, Intel Corp. Micro32 conference key note - 1999.

Thermal effects on leakage



Flautner et al. "Drowsy Caches: Simple Techniques for Reducing Leakage Power", ISCA 2002

Temperature

- Typical temperatures of Pentium 4 processor
  - P-N junction temperature = up to 73 C
  - Die temperature = 100 C
- Current (& thus speed) decreases exponentially
  - Diodes: $I = I_s(e^{V/nVt} - 1)$, Vt = kT/q, T=temperature
  - Resistors: $R \propto T$, I = V/R
  - Rule of thumb: Speed ↓ 0.15% per °C
- But leakage current increases with temperature
  - Thermal runaway
  - ↑ temp → ↑ leakage → ↑ self-heating → ↑ temp
  - Motivates multiple Vt (see 0.1 µm paper by Taur)
- Reliability decreases exponentially
  - T ↑ 10-15°C → Chip lifetime ↓ 50%!
- Physical warping & cracking
  - Different CTE at package material interfaces

Effects of High Temperature

- Typical temperatures of Pentium 4 processor
  - P-N junction temperature = up to 73 C
  - Die temperature = 100 C
- Current (& thus speed) decreases exponentially
  - Diodes: $I = I_s(e^{V/nVt} - 1)$, $Vt = kT/q$, T=temperature
  - Resistors: $R \propto T$, $I = V/R$
  - Rule of thumb: Speed ↓ 0.15% per °C
- But leakage current increases with temperature
  - Thermal runaway
  - ↑ temp → ↑ leakage → ↑ self-heating → ↑ temp
  - Motivates multiple Vt (see 0.1 μm paper by Taur)
- Reliability decreases exponentially
  - T ↑ 10-15°C → Chip lifetime ↓ 50%!
- Physical warping & cracking
  - Different CTE at package material interfaces

Effects of High Current

- Voltage drop
  - IR drop due to high current flowing through power grid
- di/dt current transients
  - Clock edges
  - Powering up/down large logic blocks
  - Add decoupling capacitors on chip and package to mitigate
- Electromigration
  - Metal atoms of thin wires physically move over time
  - Eventually, short circuits & open circuits which break the chip

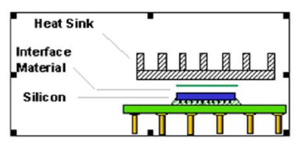
Temperature During Test

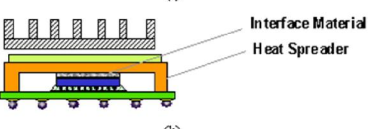
- Burn-in testing
  - Test at high Vcc and high temperature
  - Goal: Induce "infant mortality" failures without breaking good, reliable specimens
- Test 1,000 chips at once
  - Burn-in "oven" is now a refrigerator!
  - Need to dissipate up to 50 kW!
- Emerging tester technology: Hydraulic cooling

"Thermal Challenges during Microprocessor Testing", Intel Technology Journal, Q3 2000

Thermal Solutions

- Heat sink
  - Mounted on processor package
- Passive cooling
  - Remote system fan
- Active cooling
  - Fan mounted on sink
- Heat spreaders
  - Increase surface area
  - Example: Metal plate under laptop keyboard





a. Heat sink mounting for low-power chip
b. Package design for high-power chips

"Thermal Challenges during Microprocessor Testing", Intel Technology Journal, Q3 2000

Power Management on Pentium 4 Processor

- Over 400 power-saving features!
  - 20% of features = 75% of saved power
- Clock throttling
  - Thermal diode temperature sensor
  - Stop clock for a few microseconds
  - Output pin can be used by system to trigger other responses
- SpeedStep technology for mobile processors
  - Switch to lower frequency and voltage
  - Depends on whether power source is battery or AC
  - Can be manually overridden by Windows control panel

"Managing the Impact of Increasing Microprocessor Power Consumption", Intel Technology Journal, Q1 2001

Pentium 4 Multi-level Powerdown

- Level 0 = Normal operation (includes thermal throttle)
- Level 1 = Halt instructions (less processor activity)
- Level 2 = Stop Clock (internal clocks turn off)
- Level 3 = Deep sleep (remove chip input clock)
- Level 4 = Deeper sleep (lower Vdd by 66%)
  - For "extended periods of processor inactivity"
  - QuickStart technology – resume normal operation from Deeper Sleep
- Note: We haven't even talked about system powerdown modes, like removing power from processor, stopping hard disks, dimming or turning off the display…

"Managing the Impact of Increasing Microprocessor Power Consumption", Intel Technology Journal, Q1 2001

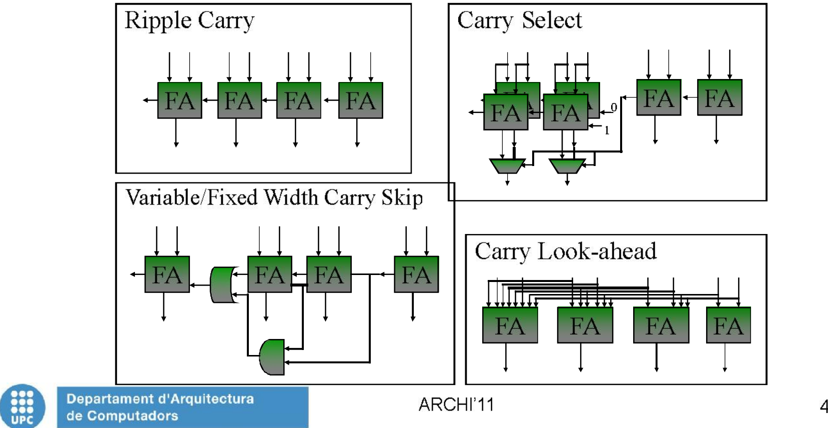
Intel
Pentium 4/2000
Socket 423
Intel 850 Chipset

Power - Performance Metrics

- Power α C V^2 f
- Metric: suppose we introduce a technique in the design that it has a cost in energy consumption and an effect on performance:
  1. Power/Perf (➔ Energy), assuming the same technology and voltage.
     - Good for:
       - Battery life, electricity bill.
       - Savings in a fixed *power envelope* – without voltage scaling.
  2. Power/$Perf^2$ (➔ Energy*Delay)
     - Balance between performance and energy consumption.
  3. Power/$Perf^3$ (➔ $Energy*Delay^2$)
     - Independent of voltage scaling
     - Good to evaluate architectural techniques that can go over the benefits of voltage scaling

Example: Adder

- There exist several adder designs:
  - Ripple, select, skip (x2), Look-ahead, conditional-sum.
  - Each on has a different trade-off in delay and power requirements



Power and Performance Figures

- Callaway i Swartzlander*:

	Energy (pJ)	Delay (nSec)
Ripple Carry	117	54.27
Constant Width Carry Skip	109	28.38
Variable Width Carry Skip	126	21.84
Carry Lookahead	171	17.13
Carry Select	216	19.56
Conditional Sum	304	20.05

- Better alternative:
  - Best power – "*constant width carry skip*"
  - Best delay – "*carry look-ahead*"

* "Estimating the power consumption of CMOS adders" - Callaway, T.K.; Swartzlander, E.E., Jr. 11th Symposium on Computer Arithmetic, 1993. Proceedings.

Conclusions

- VLSI Basics
  - Resistance & Capacity are the basis.
- Energy Consumption
  - Important design factor (at the moment more important than area or delay)
- Voltage Scaling: good but not enough
- New metrics for power-aware designs

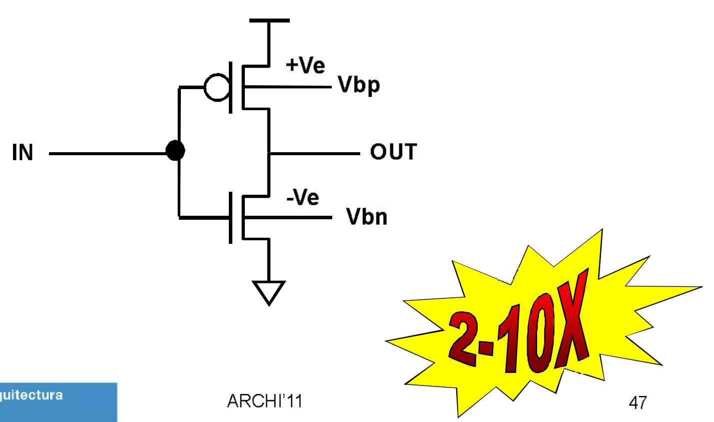
Recent interesting work on power/leakage reduction

Low leakage techniques

- Circuit level
  - Body bias
  - Stack effect
  - Sleep transistors (Vdd-gating)
- Architecture level
  - Drowsy caches
  - Cache Decay
- System level
  - Compiler-directed
  - Voltage-Frequency Scaling

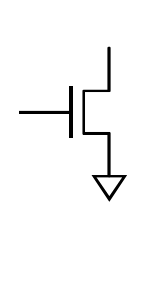
Circuit level techniques

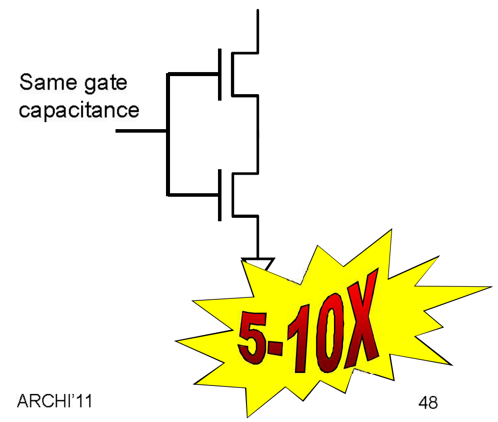
Body bias



Circuit level techniques

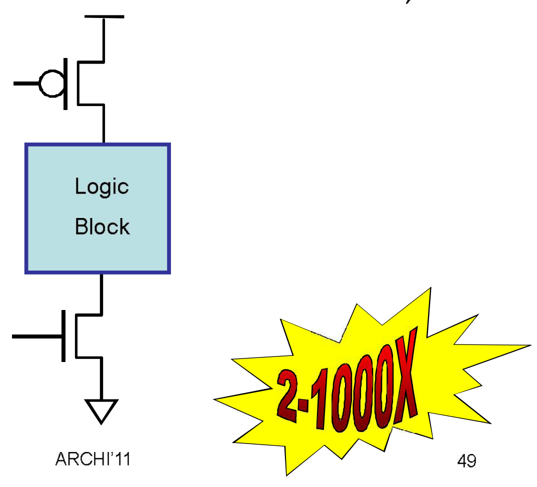
Stack effect (Chen et al. –Hong Kong U. ISLPED 1998)





Circuit level techniques

Sleep transistors (Powell et al. –Purdue U.- ILSPED 2000)

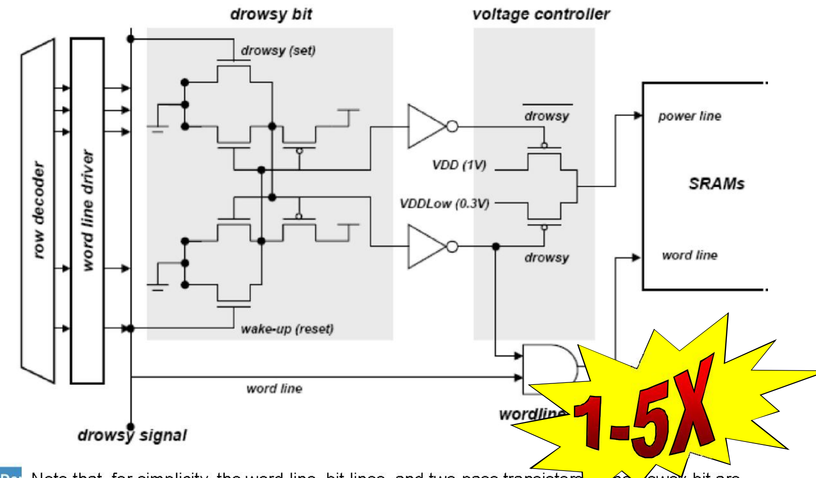


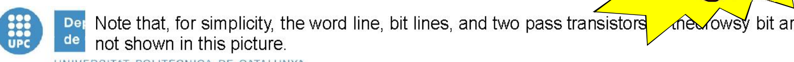
Circuit level techniques

- Circuit level
  - Body bias
  - Stack effect
  - Sleep transistors (Vdd-gating)
- Aims:
  - Reduce leakage
- Consequences
  - Extra transistors
  - Multiple speed transistors
    - Maybe leading to multiple speed units, datapaths…

Architecture level techniques

Drowsy Caches (K. Flautner et al. –ARM, U. Michigan- ISCA 2002)

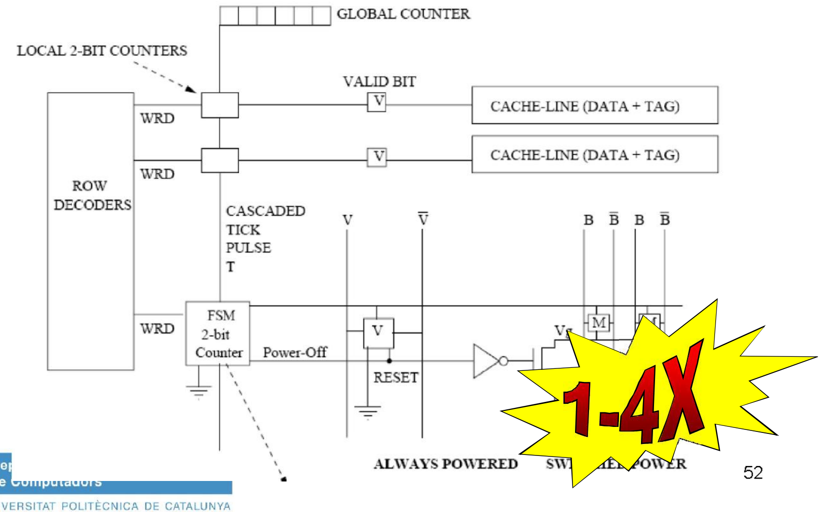




Note that, for simplicity, the word line, bit lines, and two pass transistors in the drowsy bit are not shown in this picture.

Architecture level techniques

Cache Decay (S. Kaxiras et al. –Agere, Princeton U.- ISCA 2001)

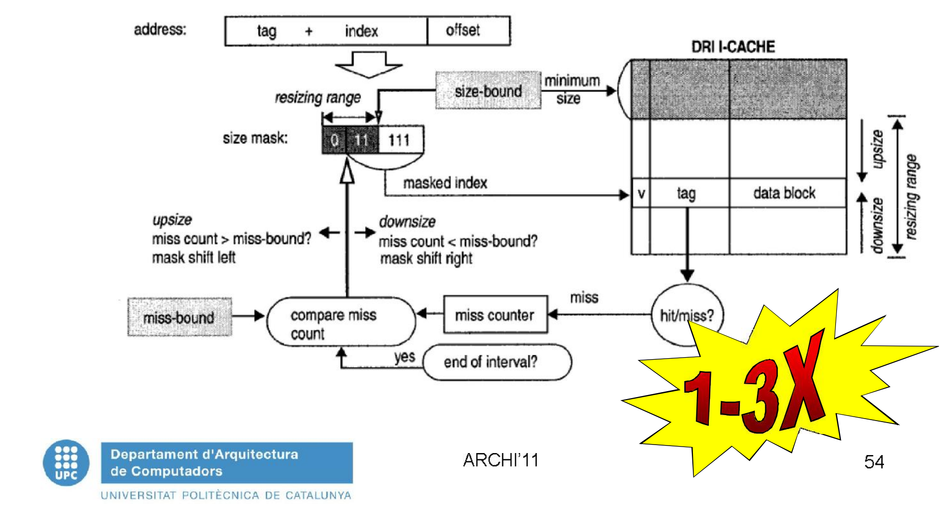


Architecture level techniques

- Architecture level
  - Drowsy caches
  - Cache Decay
- Aims:
  - Reduce leakage by turning-off/"low powering" unused (or possibly unused) cache lines
- Consequences
  - Extra circuitry
  - Access delays
  - Execution penalties

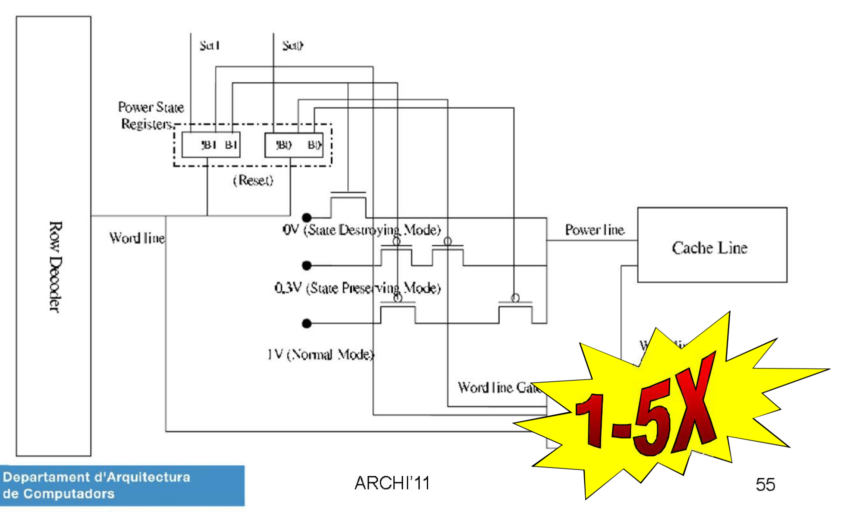
System level techniques

Adaptative Cache (Powell et al. –Purdue U.- IEEE Trans. VLSI 2/2001)



System level techniques

Compiler-directed (Zhang et al. –Penn. State U.- MICRO 2002)



System level techniques

- System level
  - Compiler-directed
  - Voltage-Frequency Scaling
- Aims:
  - Reduce leakage by turning-off/"low powering" unused (or possibly unused) cache lines
- Consequences
  - Difficult compile-time decisions
  - Extra circuitry
  - Access delays
  - Execution penalties

Conclusions

- Need for power, energy and thermal awareness.
  - (TDP) Thermal Design Power
  - Even commercially advertised power efficiency (Cool’n quiet –AMD)
- Extensive use of power/performance metrics such as Energy Delay
  - MIPS/W (Intel)