The Abdus Salam International Centre for Theoretical Physics

2499-5

**International Training Workshop on FPGA Design for Scientific Instrumentation and Computing**

***11 - 22 November 2013***

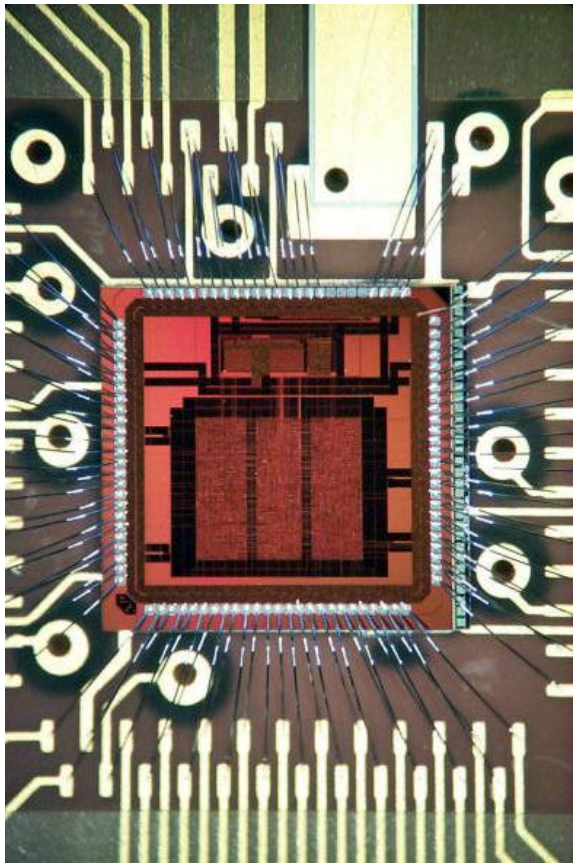
# Introduction to VLSI Digital Design
# The CMOS Inverter

Sandro BONACINI
*CERN, Geneva*
*Switzerland*

Strada Costiera, 11 - 34151 - Trieste - Italy • Tel. +39 0402240111 • Fax. +39 040224163 • sci_info@ictp.it • www.ictp.it
ICTP is governed by UNESCO, IAEA, and Italy, and it is a UNESCO Category 1 Institute

# Outline

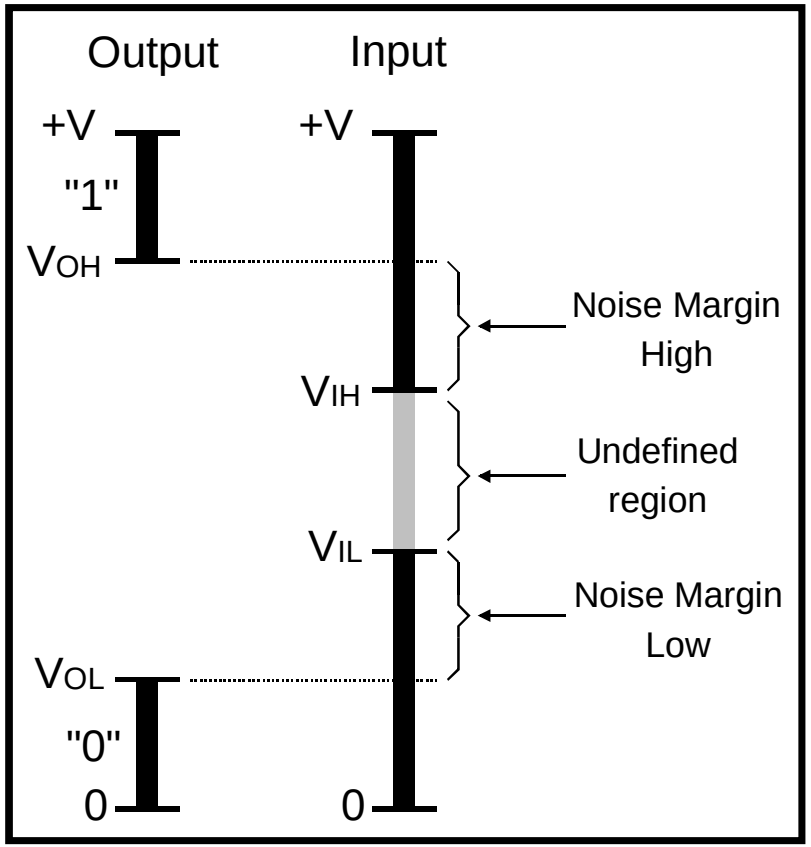
- Introduction
- Transistors
- The CMOS inverter
  - DC operation
  - Dynamic operation
    - Propagation delay
    - Power consumption
  - Layout
  - Delays
- Technology
- Scaling
- Gates
- Sequential circuits
- Storage elements



Many slides are a courtesy of Paulo Moreira

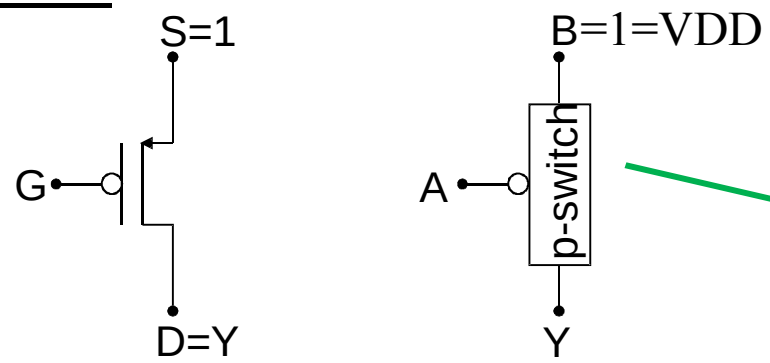
# CMOS logic: "0" and "1"

- Logic circuits process Boolean variables
- Logic values are associated with voltage levels:
  - $V_{IN} > V_{IH} \Rightarrow$ "1"
  - $V_{IN} < V_{IL} \Rightarrow$ "0"
- Noise margin:
  - $NM_H = V_{OH} - V_{IH}$
  - $NM_L = V_{IL} - V_{OL}$



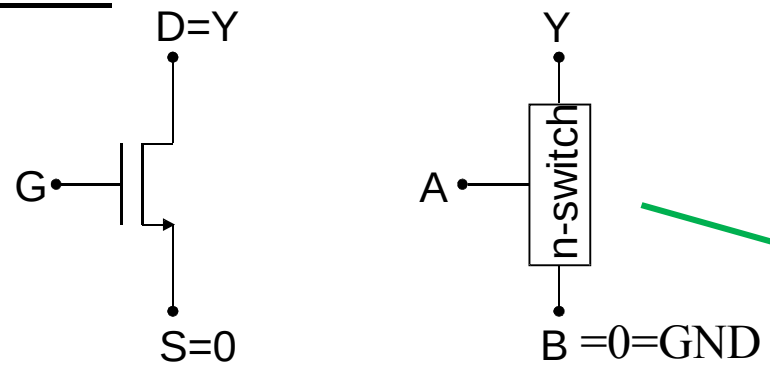
# a simple switch

**p-switch**



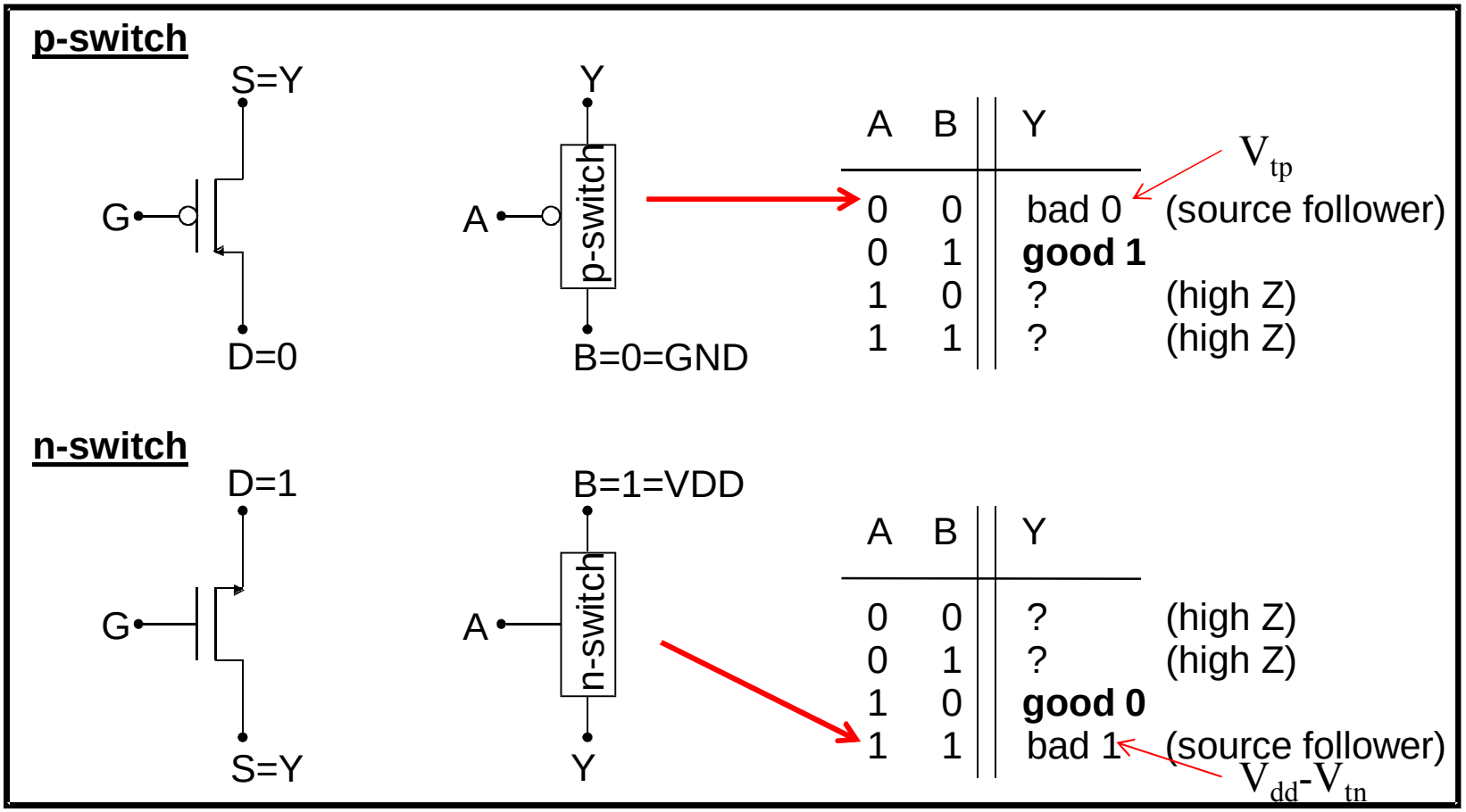

| A | B | Y | |
|---|---|---|---|
| 0 | 0 | bad 0 | (source follower) |
| 0 | 1 | **good 1** | |
| 1 | 0 | ? | (high Z) |
| 1 | 1 | ? | (high Z) |

**n-switch**

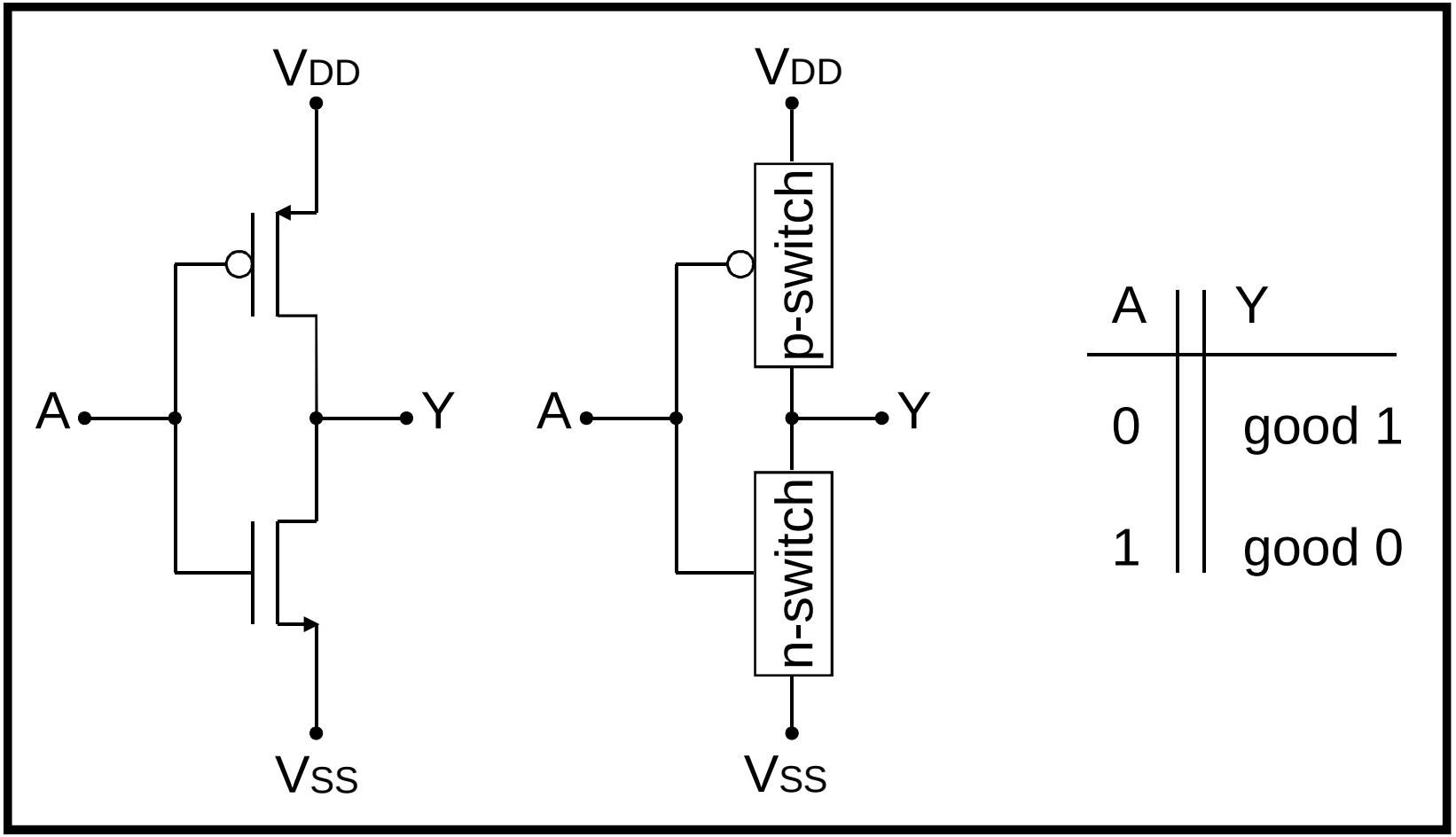



| A | B | Y | |
|---|---|---|---|
| 0 | 0 | ? | (high Z) |
| 0 | 1 | ? | (high Z) |
| 1 | 0 | **good 0** | |
| 1 | 1 | bad 1 | (source follower) |

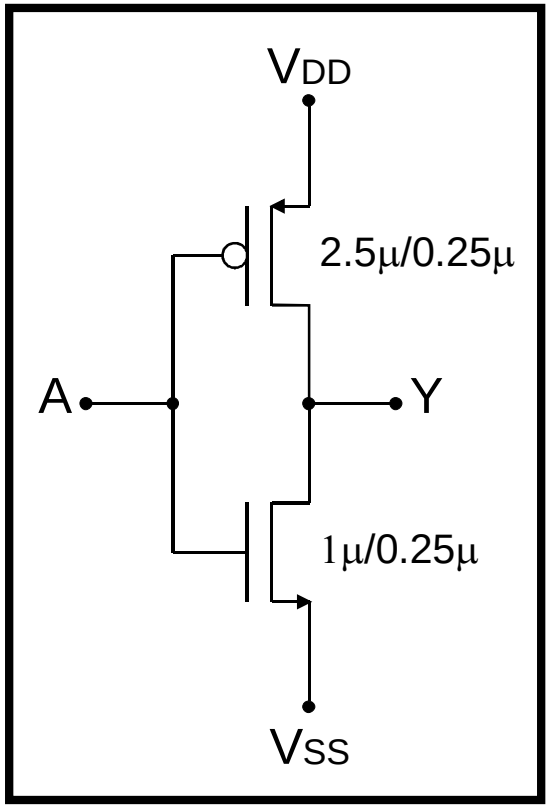
# a simple switch – "bad configuration"

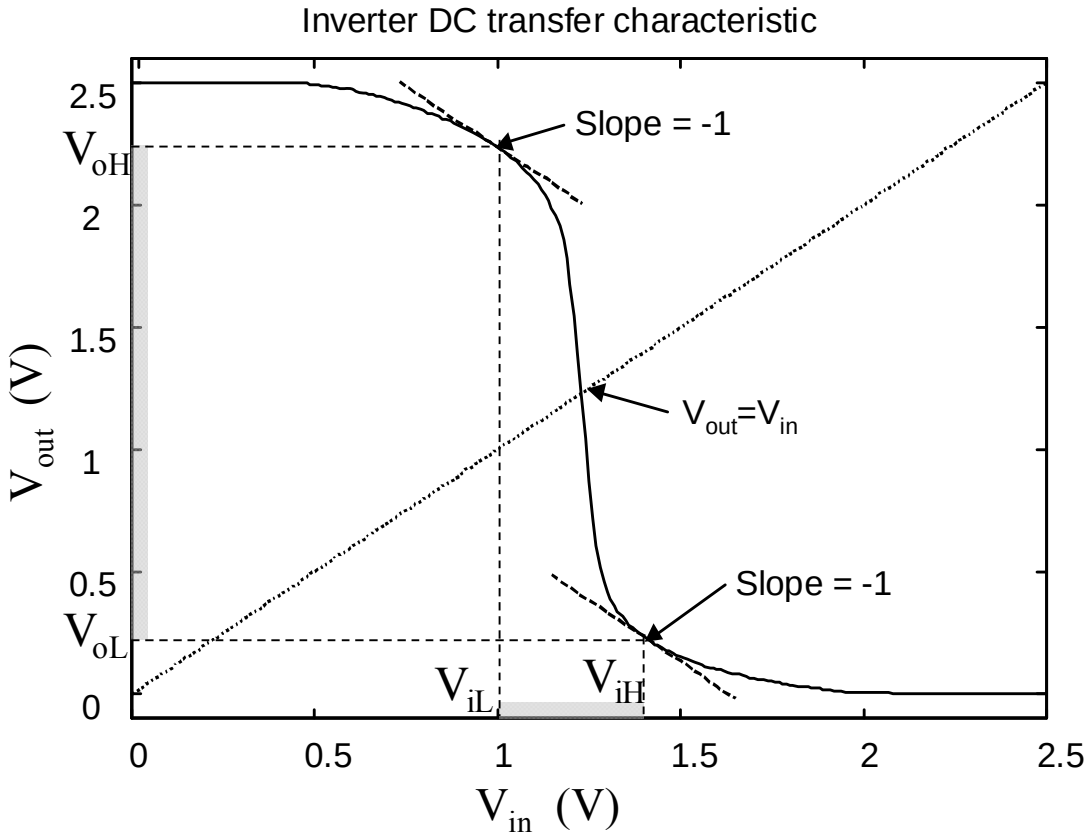


# The CMOS inverter



# DC operation



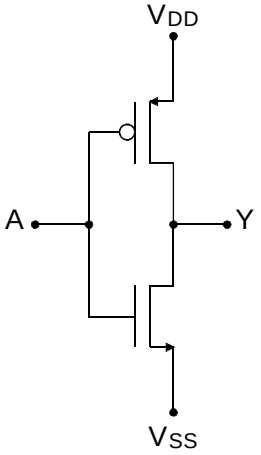



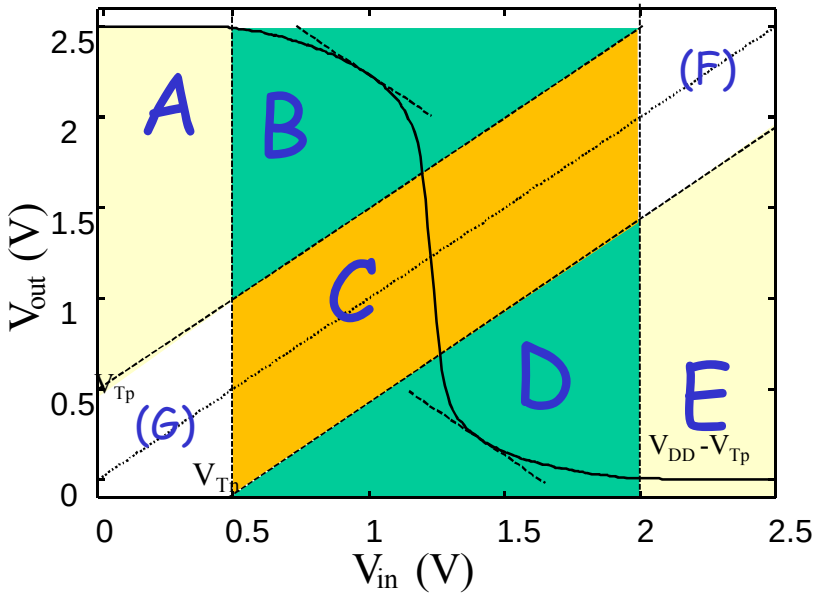
# DC operation

## Regions of operation (balanced inverter):

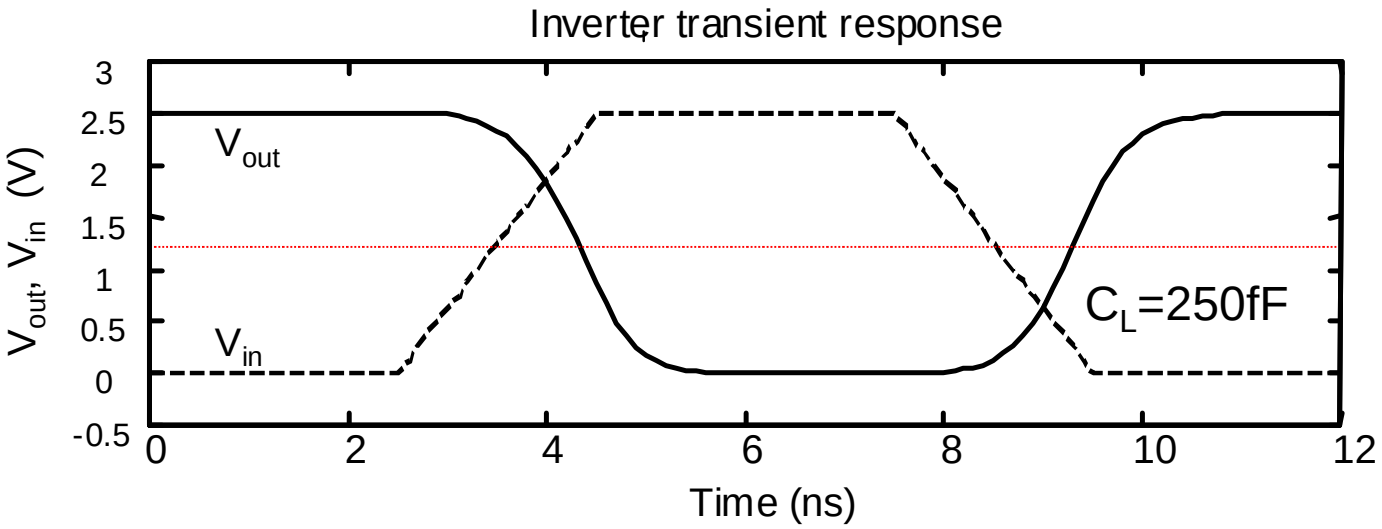
| | $V_{in}$ | n-MOS | p-MOS | $V_{out}$ |
|---|---|---|---|---|
| A | $V_{in} < V_{TN}$ | cut-off | linear | $V_{dd}$ |
| B | $V_{TN} < V_{in} < V_{out}-V_{TP}$ | saturation | linear | $\sim V_{dd}$ |
| C | $V_{out}-V_{TP} < V_{in} < V_{out}+V_{TN}$ | saturation | saturation | $\sim V_{dd}/2$ |
| D | $V_{out}+V_{TN} < V_{in} < V_{dd}-V_{TP}$ | linear | saturation | $\sim 0$ |
| E | $V_{in} > V_{dd}-V_{TP}$ | linear | cut-off | 0 |

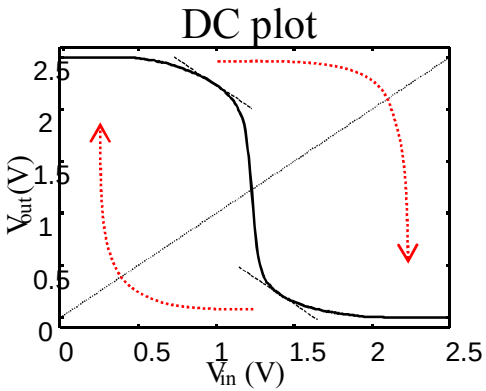
In the C zone lies most of the short-circuit current.
-> to be avoided in transients!

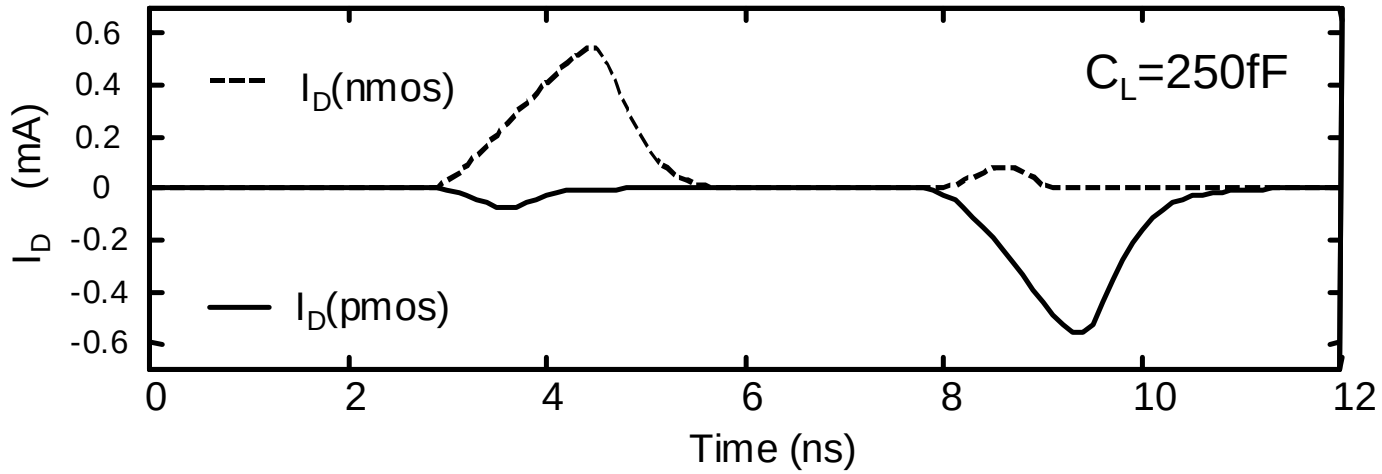


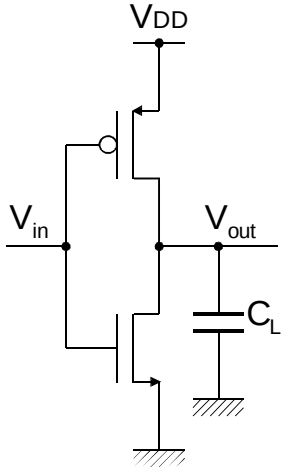



# Dynamic operation



# Dynamic operation

- Propagation delay
  - Saturation region (approx.) $I_{ds} = \frac{\mu \cdot C_{ox}}{2} \cdot \frac{W}{L} \cdot \left(V_{gs} - V_T\right)^2$
  - Main origin: load capacitance

$$t_p = \frac{C_L V_{DD}}{I_{ds}}$$

$$t_{pLH} = \frac{2 \cdot C_L \cdot L_p \cdot V_{dd}}{\mu_p \cdot C_{ox} \cdot W_p \cdot \left(V_{dd} - \left|V_{TP}\right|\right)^2} \approx \frac{2 \cdot C_L}{k_p \cdot V_{dd}}$$

$$t_{pHL} = \frac{2 \cdot C_L \cdot L_n \cdot V_{dd}}{\mu_n \cdot C_{ox} \cdot W_n \cdot \left(V_{dd} - \left|V_{TN}\right|\right)^2} \approx \frac{2 \cdot C_L}{k_n \cdot V_{dd}}$$

  - To reduce the delay:
    - Reduce $C_L$
    - Increase $k_n$ and $k_p$. That is, increase W/L

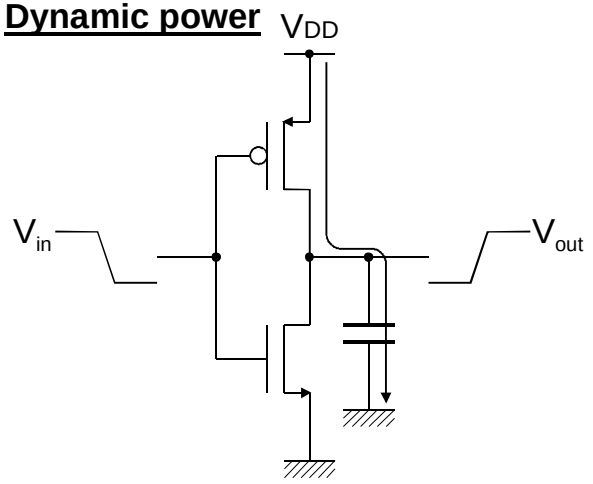
# Dynamic operation

- CMOS power budget:
  - Dynamic power consumption:
    - Charging and discharging of capacitors
  - Short circuit currents:
    - Short circuit path between power rails during switching
  - Leakage
    - Leaking diodes and transistors

# Dynamic operation

- The dynamic power dissipation is a function of:
  - Frequency
  - Capacitive loading
  - Voltage swing
- To reduce dynamic power dissipation
  - Reduce: $C_L$
  - Reduce: f
  - Reduce: $V_{dd}$ ⇐ The most effective action

**Dynamic power**




In the capacitor we store:

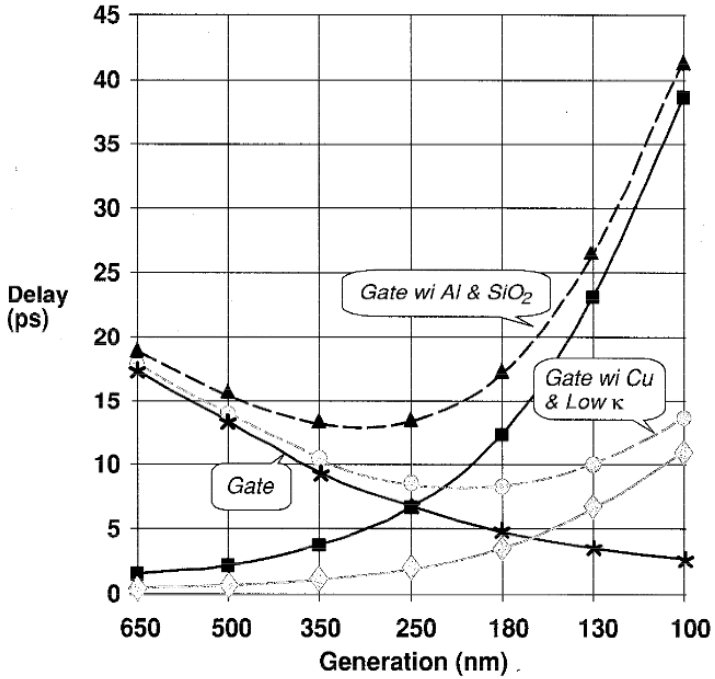
$$E = \text{Energy/transition} = \frac{1}{2} \cdot C_L \cdot V_{dd}^2$$

Another $C_L V_{dd}^2/2$ is dissipated on the PMOS when charging

$$P = \text{Power} = 2 \cdot f \cdot E = f \cdot C_L \cdot V_{dd}^2$$

# Load capacitance

- Reduce: $C_L$
  - $C_L$ is the sum of interconnect and transistor capacitances
  - $C_L = n_{fanout}C_{gate} + C_{interconnect} + C_{drain}$
  - An ***interconnect*** is a thin-film wire that electrically connects 2 or more components in an integrated circuit.
    - As transistors are scaled down in size and the number of metal wiring layers increases, the impact of interconnect parasitics increases



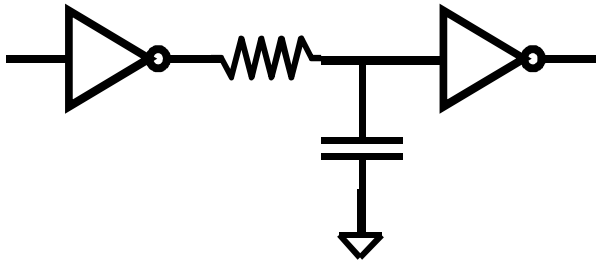
# Repeaters

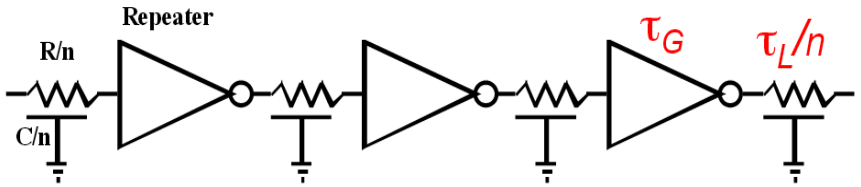
Metal lines run over thick oxide covering the substrate

→ contribute RESISTANCE & CAPACITANCE to the output node of the driving logic gate

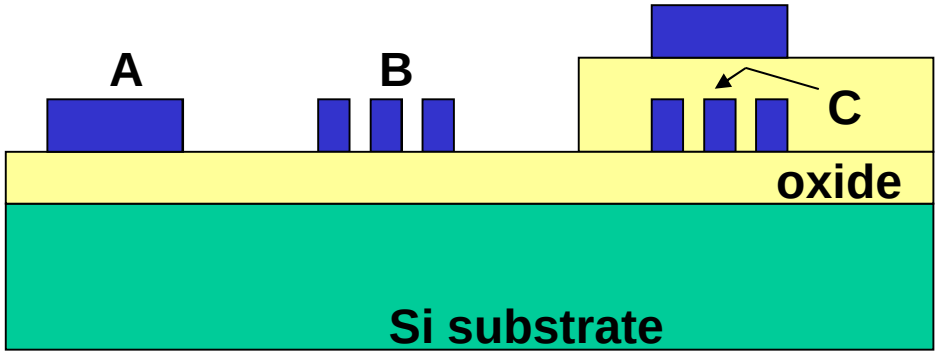
R and C are normally proportional to the length of the wire L, so the delay is proportional to $L^2$

- A first option for reducing RC delays is to use better interconnect materials when they areavailable and appropriate.
  - Copper vs Aluminum
- For very long wires the delay can be substantially larger than the gate delay
- It is possible to reduce the propagation delay by introducing intermediate buffers

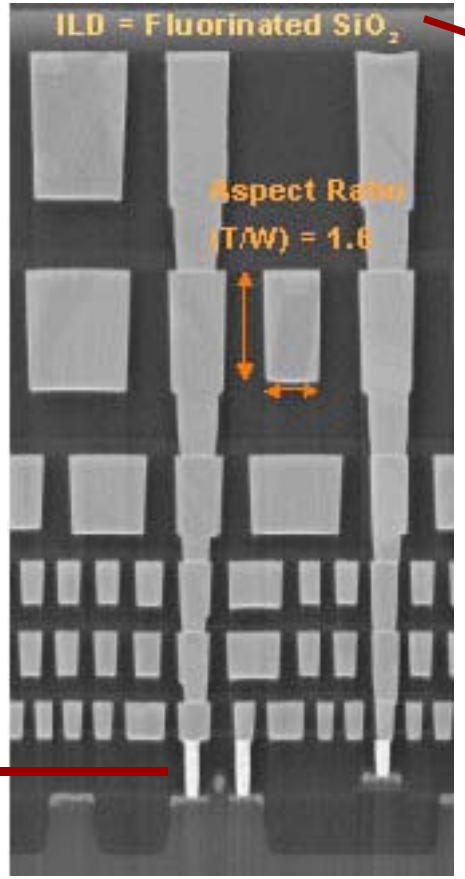




# Interconnect Wire-to-Wire Capacitance

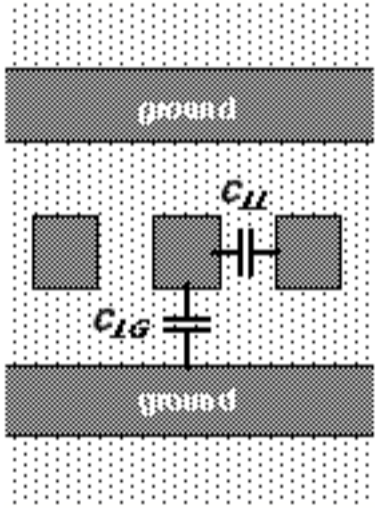


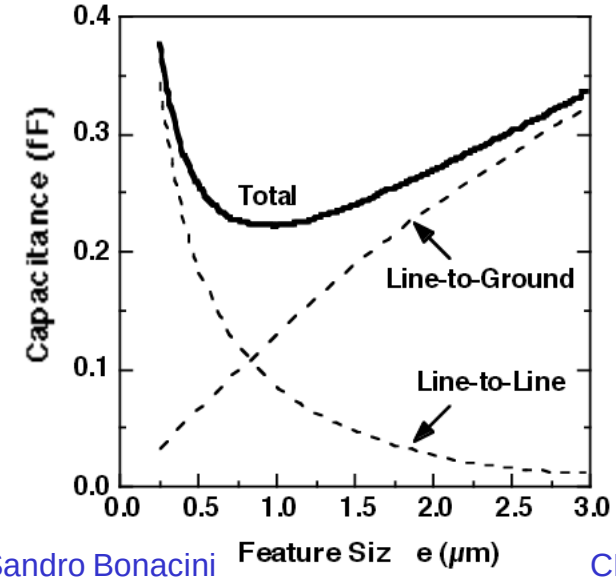

- Wire A simply has capacitance to substrate
- Wire B has additional sidewall capacitance to neighboring wires
- Wire C has additional capacitance to the wire above it




0.13µm Process (Cu)

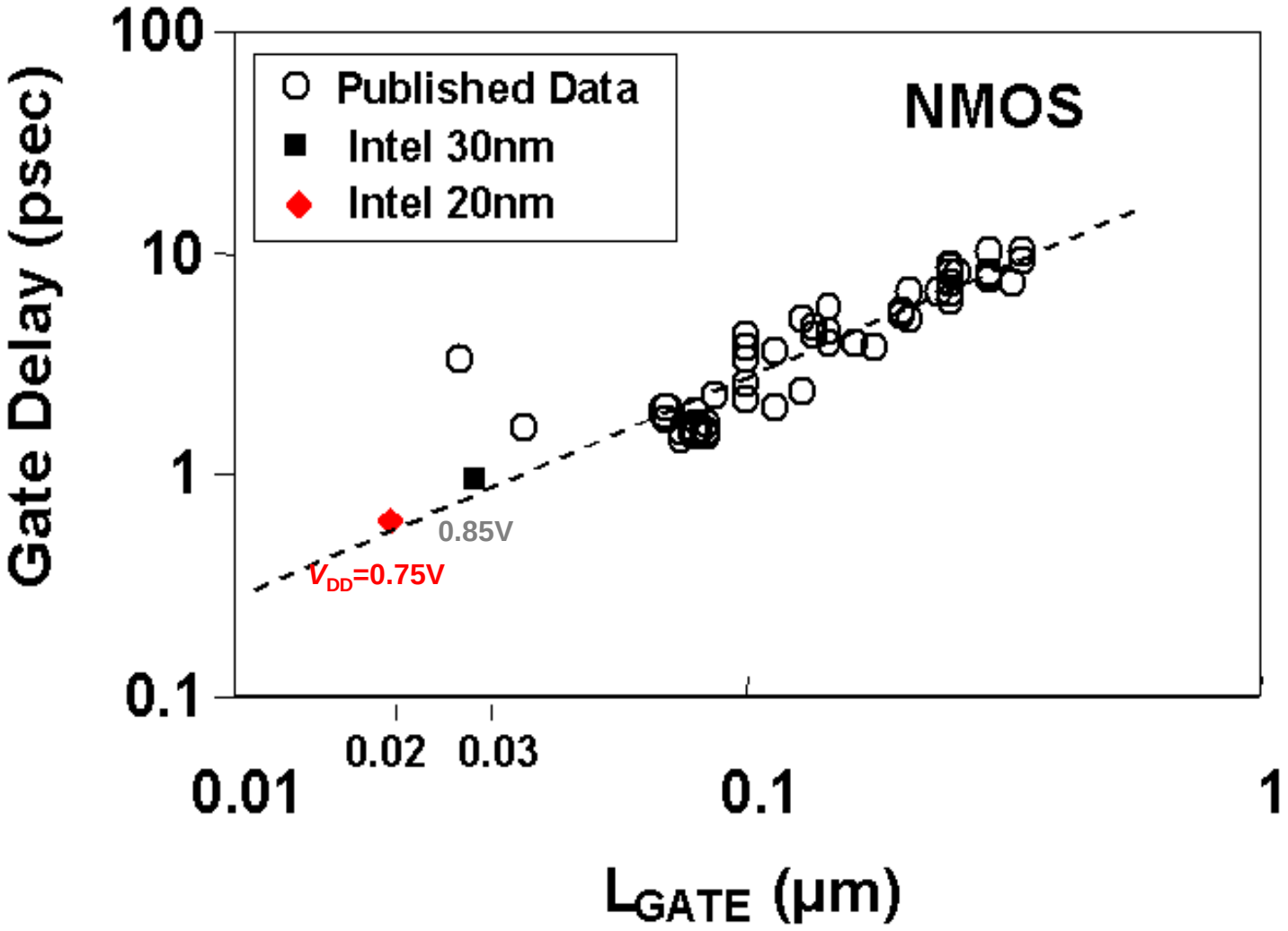
# Crosstalk



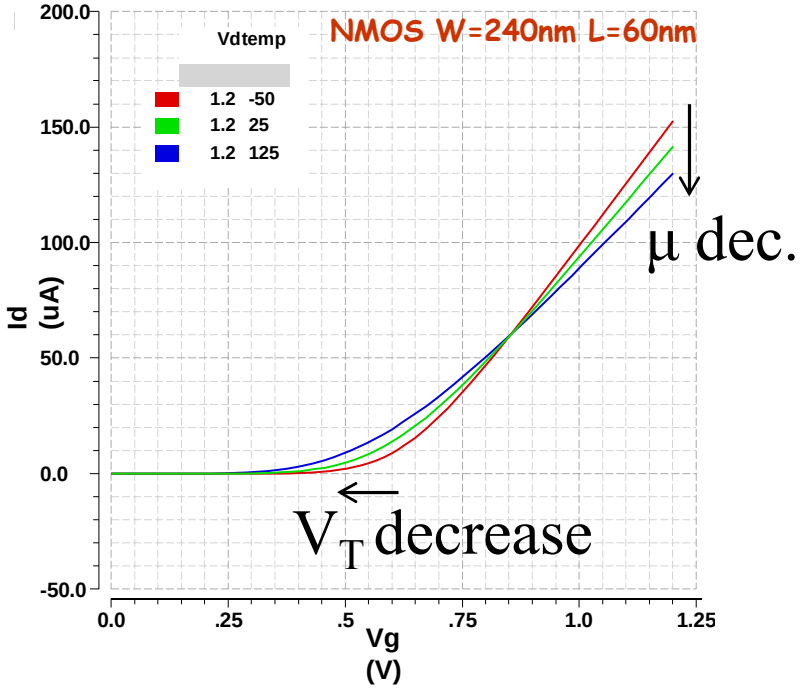

- Crosstalk: unwanted transfer of signals from one place to another through capacitive coupling
- A possible solution is to reduce the dielectric constant of the insulator
  - Low-k materials
- Increase distance between lines for critical nets

# Intrinsic Gate Delay ($C_{gate} V_{DD} / I_{DSAT}$)



# Temperature dependence

- **Transistor characteristics are influenced by temperature (T)**
  - $V_T$ decreases linearly with T
  - μ decreases with T
  - $I_{leakage}$ increases with T
- Transistor performances are worst at high temperature




- Gate performance is worse at high temperature
- Interconnect resistance increase with temperature
- Interconnect capacitances normally decrease with temperature
  - -> Logic blocks are slower with increasing temperature