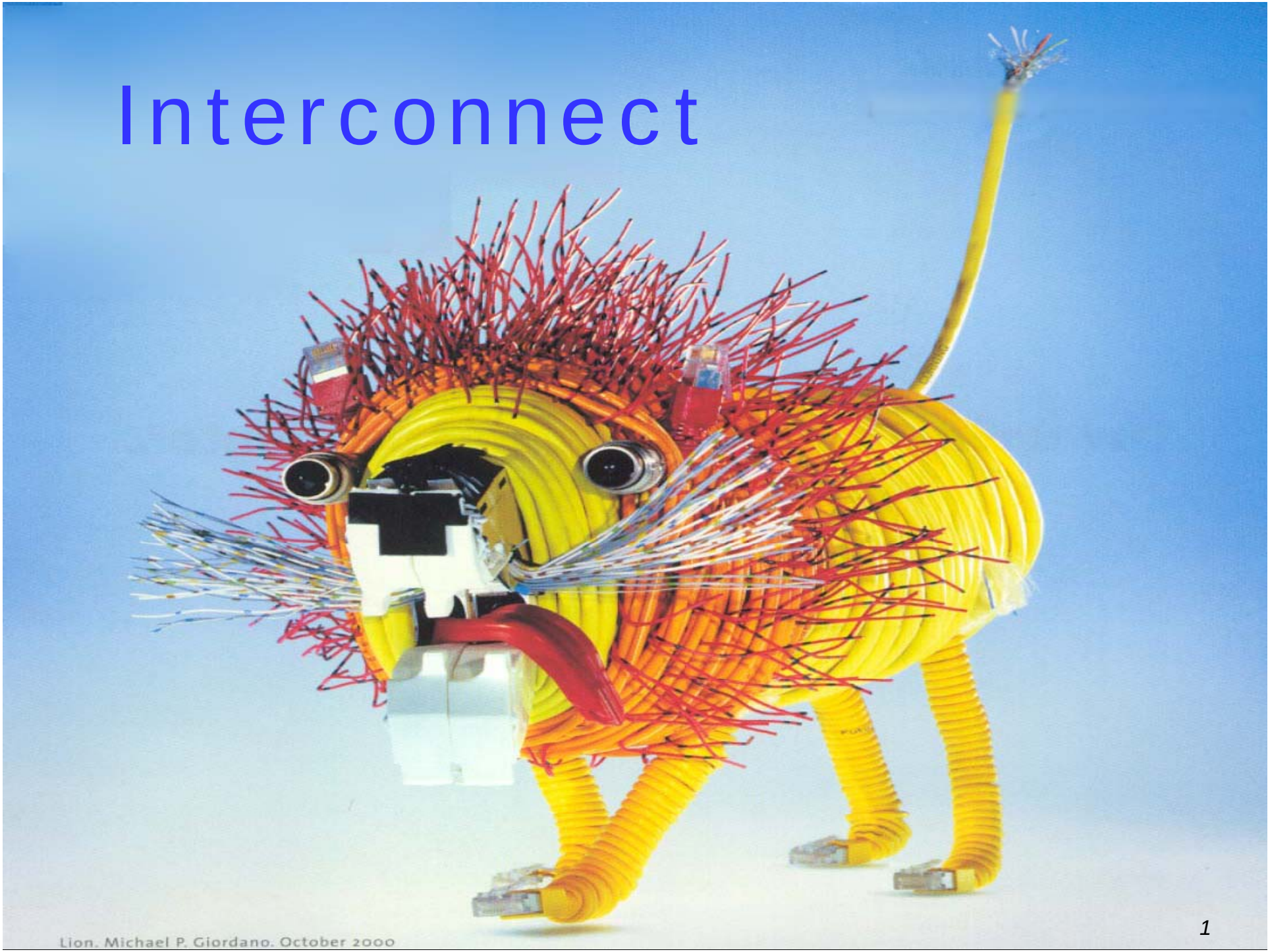
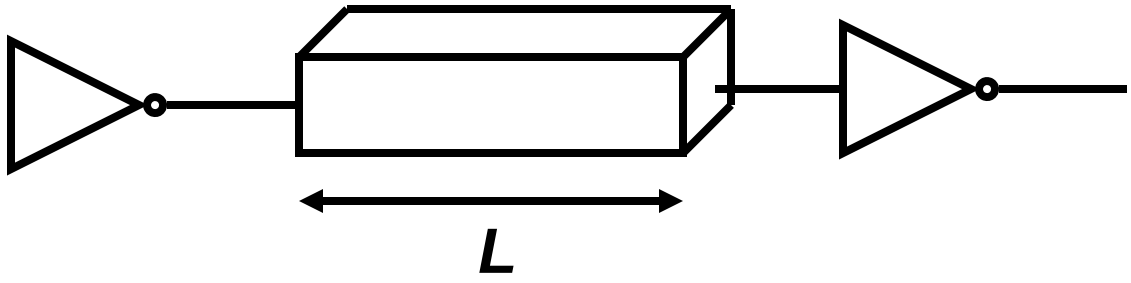
Interconnect



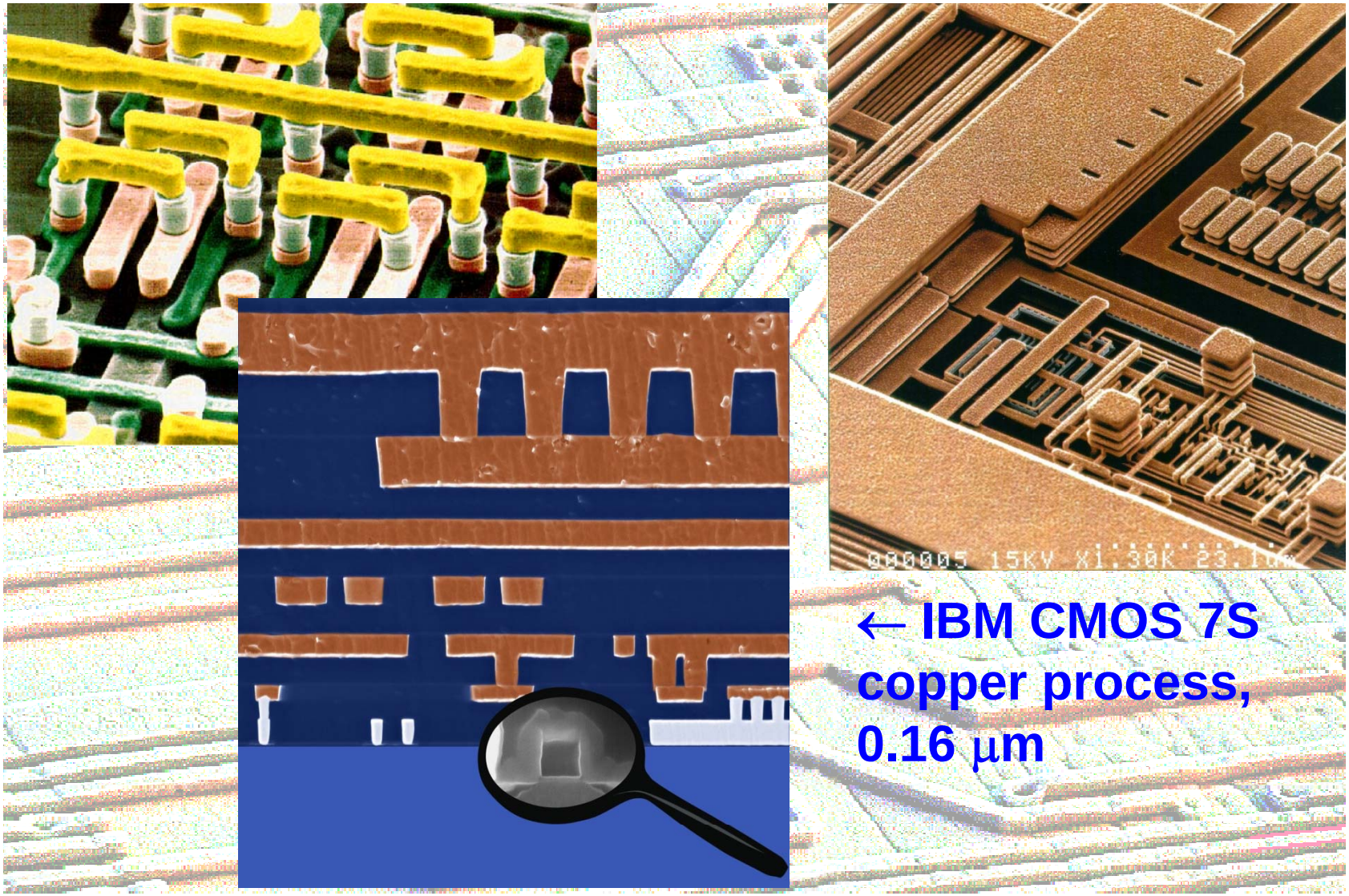
Lion, Michael P. Giordano, October 2000

Interconnect

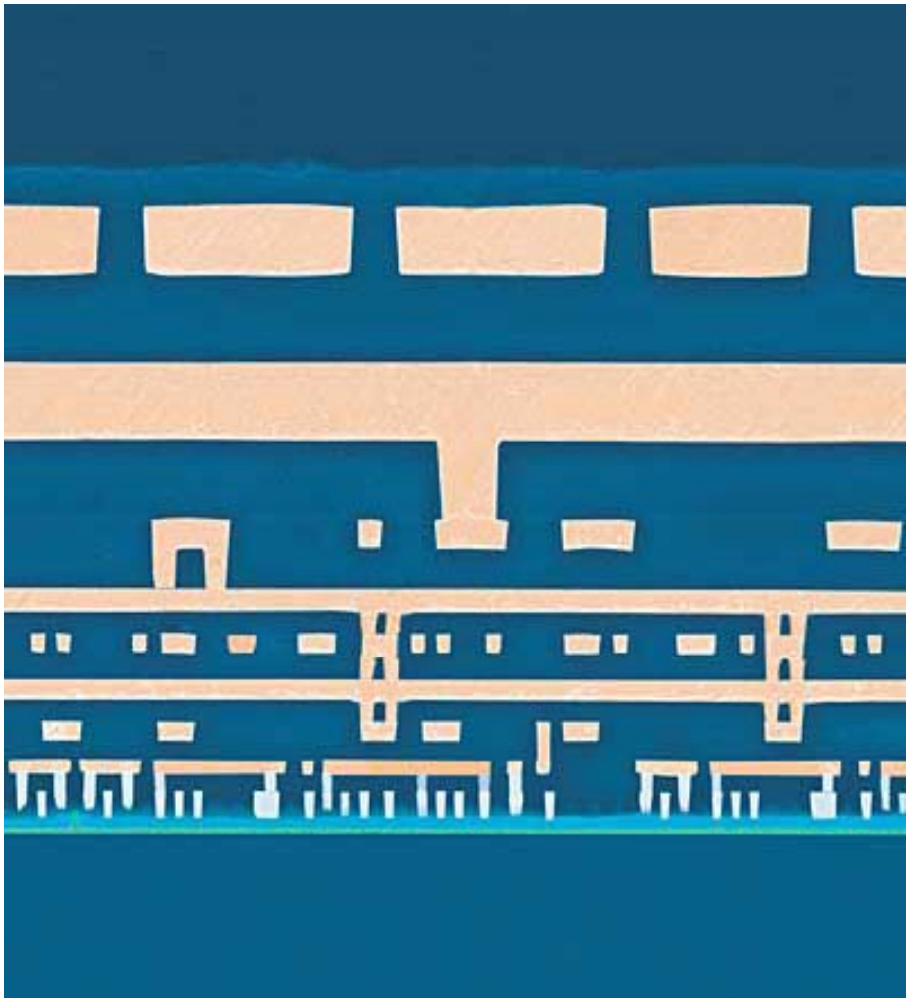


- Wires are not ideal interconnections
- They may have non-negligible capacitance, resistance, inductance
- These are called wire parasitics
- Can dominate performance of chip
- Must be accounted for during design
- Using approximate models
- Detailed post-layout verification also necessary

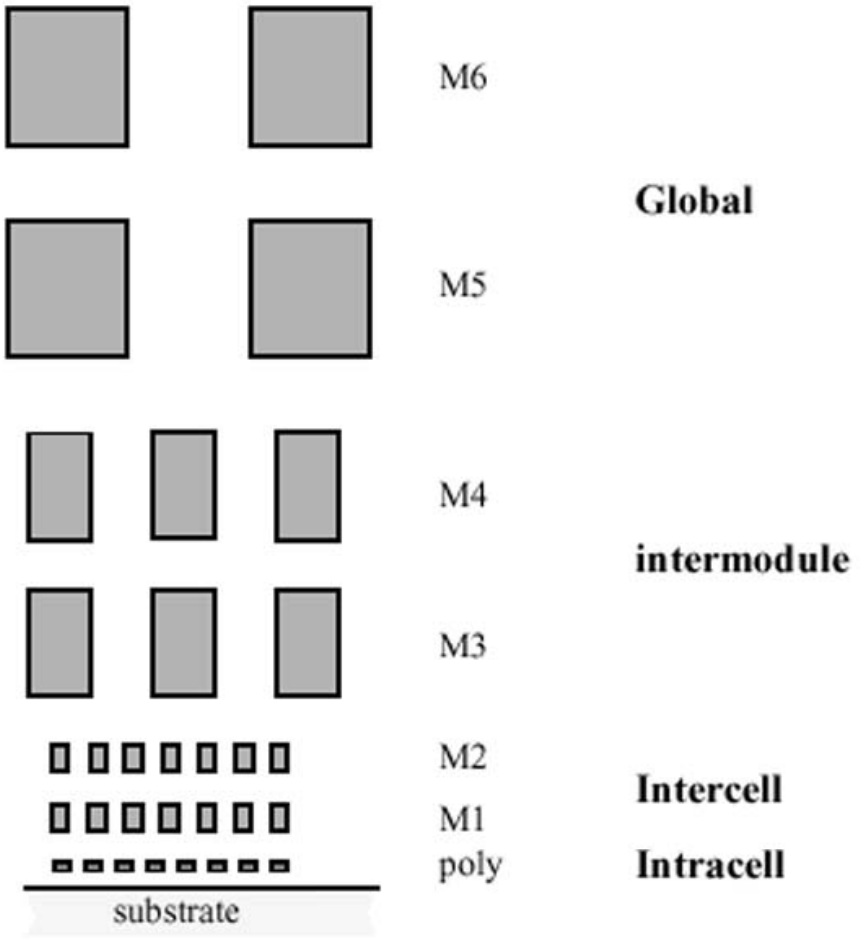
Wires



Interconnect Hierarchy



Cross-section of IBM 0.13 μ process



Example Interconnect Hierarchy for typical 0.25μ process (Layer Stack)

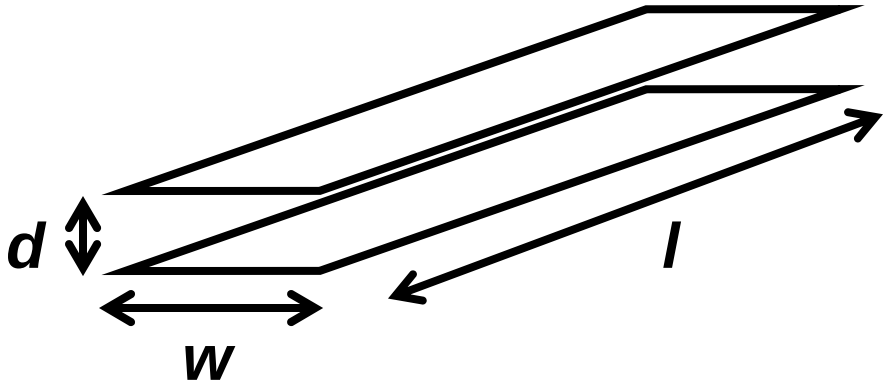
Outline

- **Capacitance**
  - **Area/perimeter model, coupling**
- **Resistance**
  - **Sheet resistance**
- **Interconnect delay**
  - **Delay metrics, rc delay, Elmore delay**

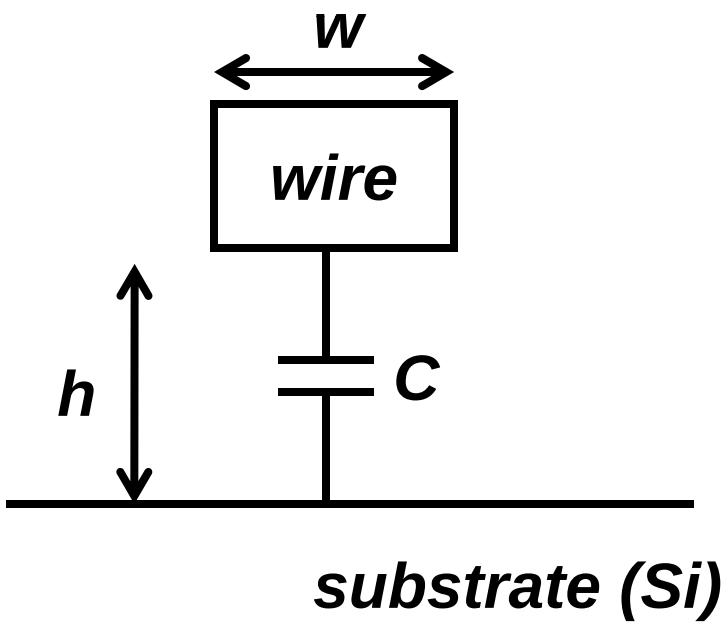
Capacitance

- **Area/perimeter model, coupling**

Wire Capacitance – Parallel Plate



$$C = \frac{\varepsilon_0 \varepsilon_r w l}{d}$$

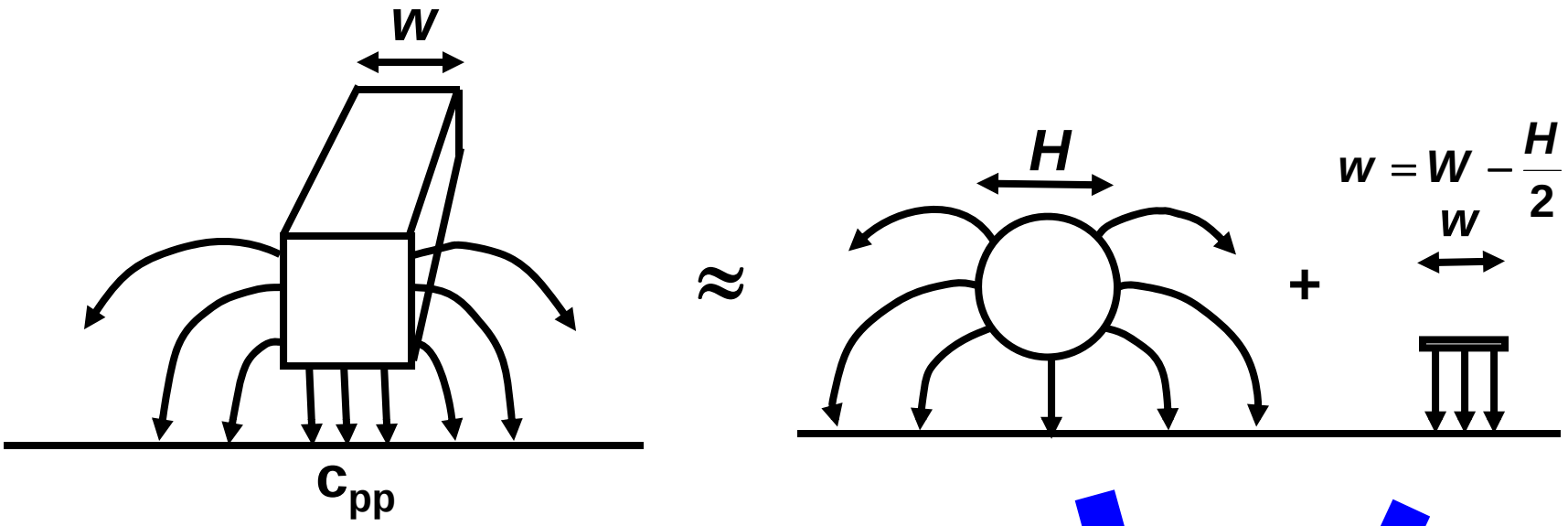


$$\frac{C}{l} = \varepsilon_0 \varepsilon_r \frac{w}{h}$$

$\varepsilon_0 = 8.85$ pF/m

$\varepsilon_r = 3.9$ (SiO_2)

Wire Capacitance – Fringing Fields



$$c_{wire} = c_{fringe} + c_{pp} = \frac{2\pi\varepsilon_{di}}{\log(t_{di}/H)} + \frac{w\varepsilon_{di}}{t_{di}}$$

- **Works reasonably well in practice**
- **Not directly applicable for interconnects with varying widths**

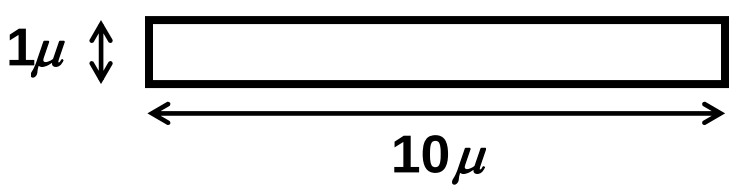
Wire Capacitance – Area/Perimeter Model

- **C_a was calculated with modified wire width**
- **Formula inapplicable for irregular interconnects (non-constant width)**

- **More practical approximation**

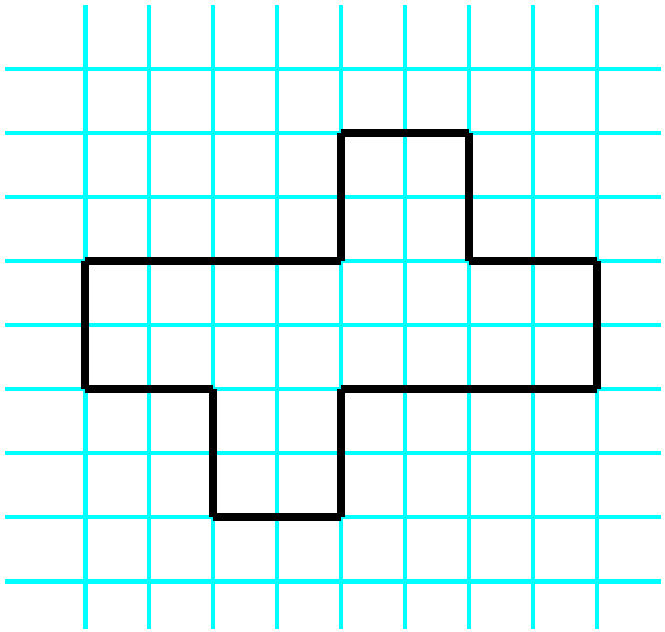
$$C = A \times C_a + P \times C_p$$

	units	*alternative*
A = Area	m^2	μm^2
C_a = Area capacitance	F / m^2	$aF / \mu m^2$
P = Perimeter	m	μm
C_p = Perimeter capacitance	F / m	$aF / \mu m$



$$C = \square \times C_a + \square \times C_p$$

Area / Perimeter Capacitance Model

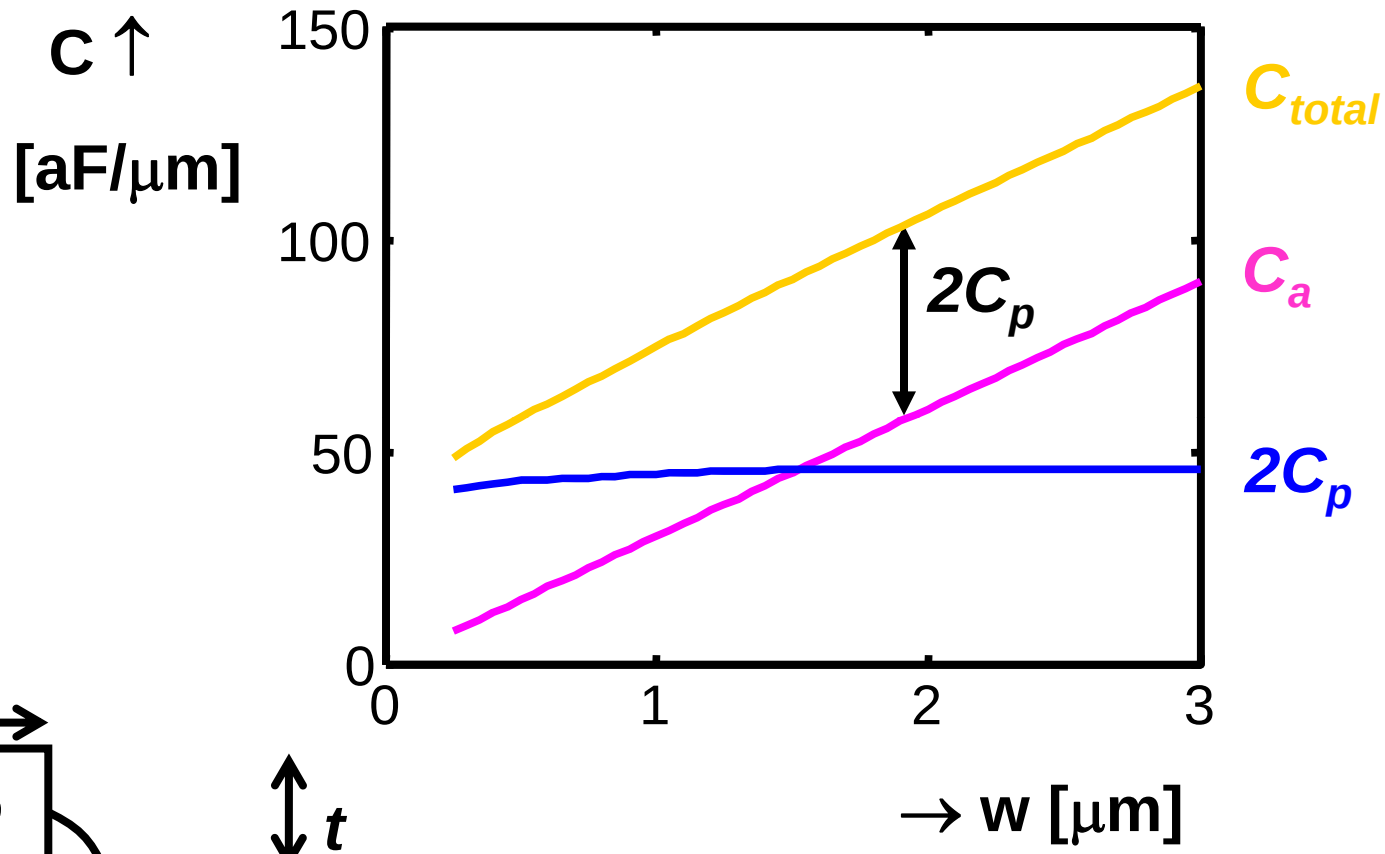


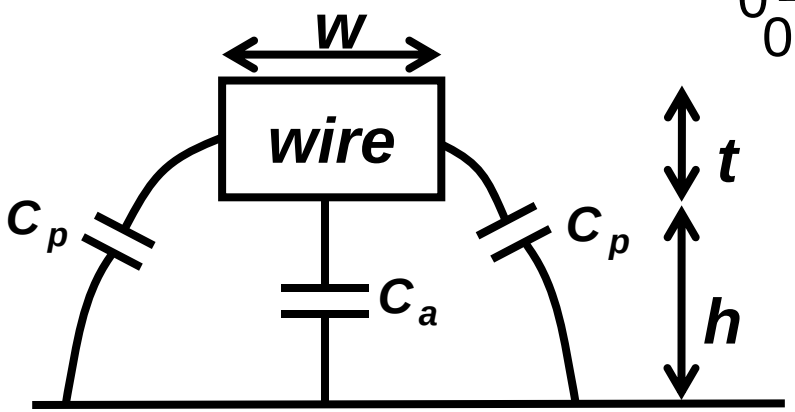
$$C = \square \times C_a + \square \times C_p$$

- **Question: How to derive C_a, C_p ?**

How accurate is this model?

Derivation of C_a, C_p

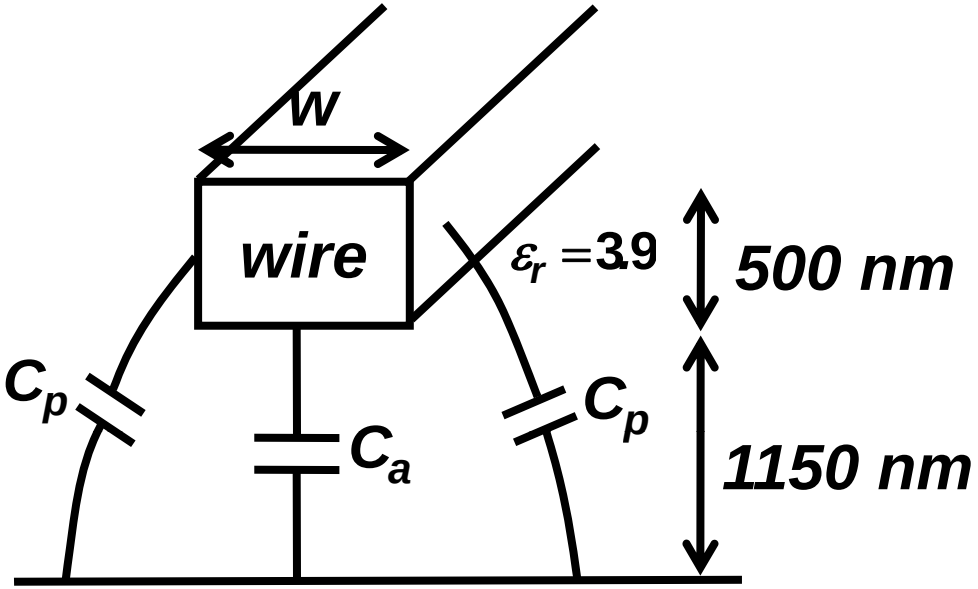


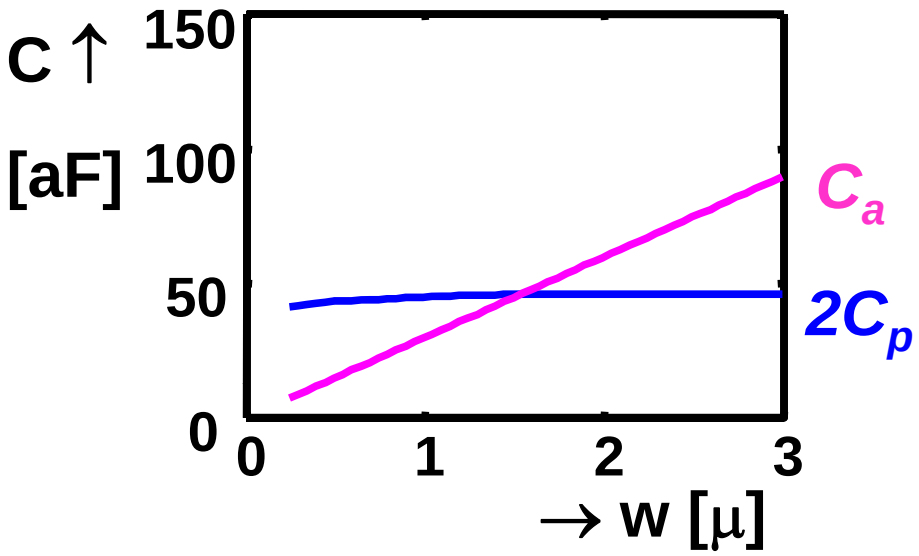


Derivation of C_a, C_p

- **2D (cross-section) numerical computation (or measurement)**
- **C_l : total wire capacitance per unit length**
- $C_a = \varepsilon_0 \varepsilon_r / h$
- $C_p = 1/2 (C_l - C_a \times w)$
- **C_p depends on t, h ➔ determined by technology, layer**
- **C_p would depend slightly on w (see previous graph), this dependence is often ignored in practice**

Area / Perimeter Capacitance

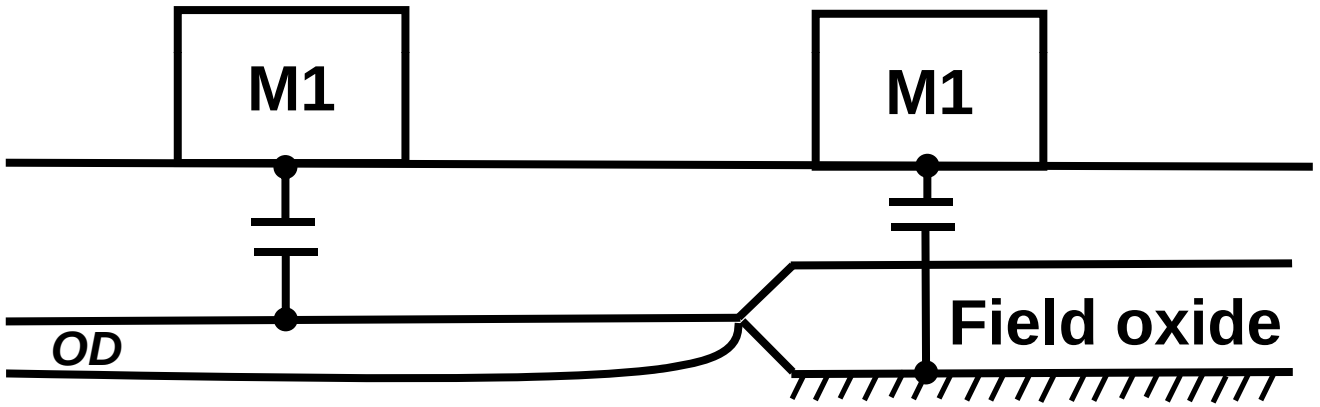




- C_p dominates for many wires
- C_p may not be neglected
- A constant value for C_p is usually a good approximation
- C_p is sometimes called C_f (fringe capacitance)

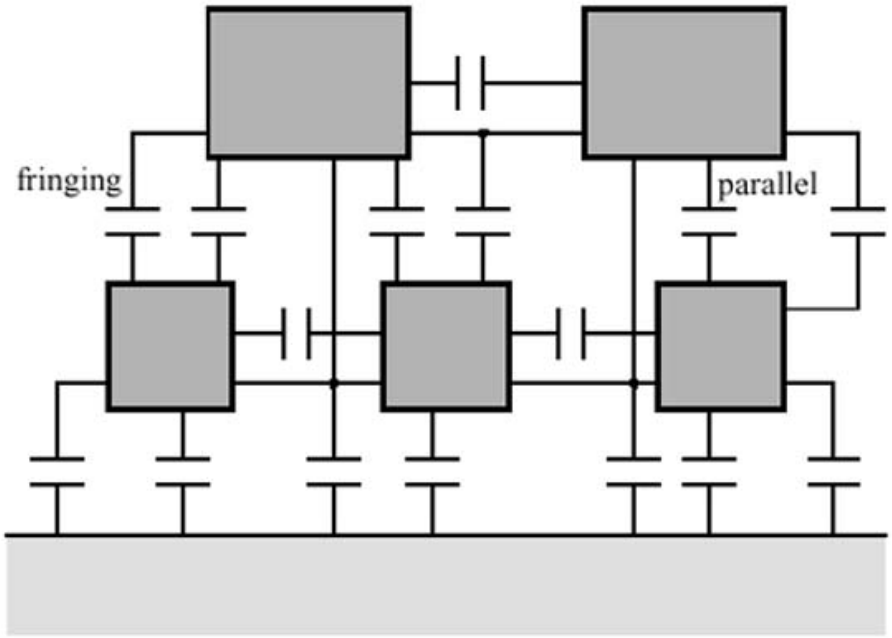
Interconnect Capacitance Design data

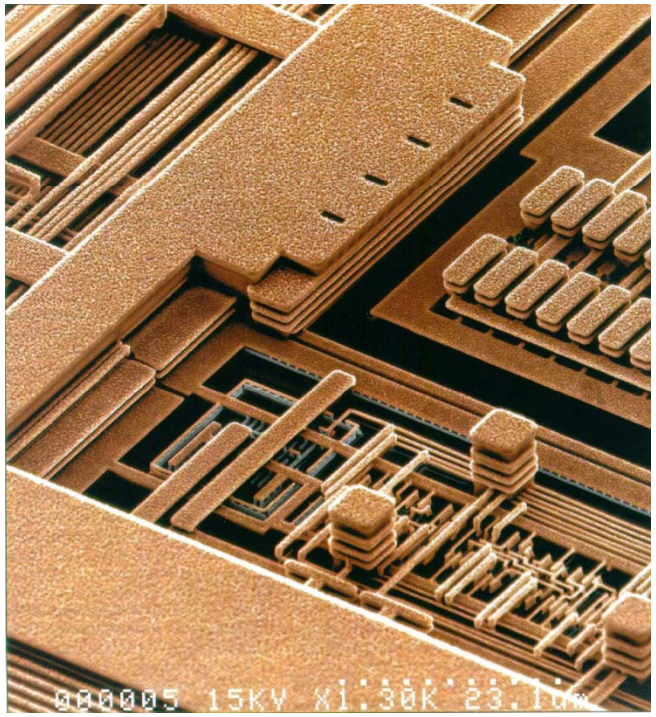
- **See Table 4.2 (or inside backside cover)**
- **Example: M1 over Field vs. M1 over Active (hypothetical)**



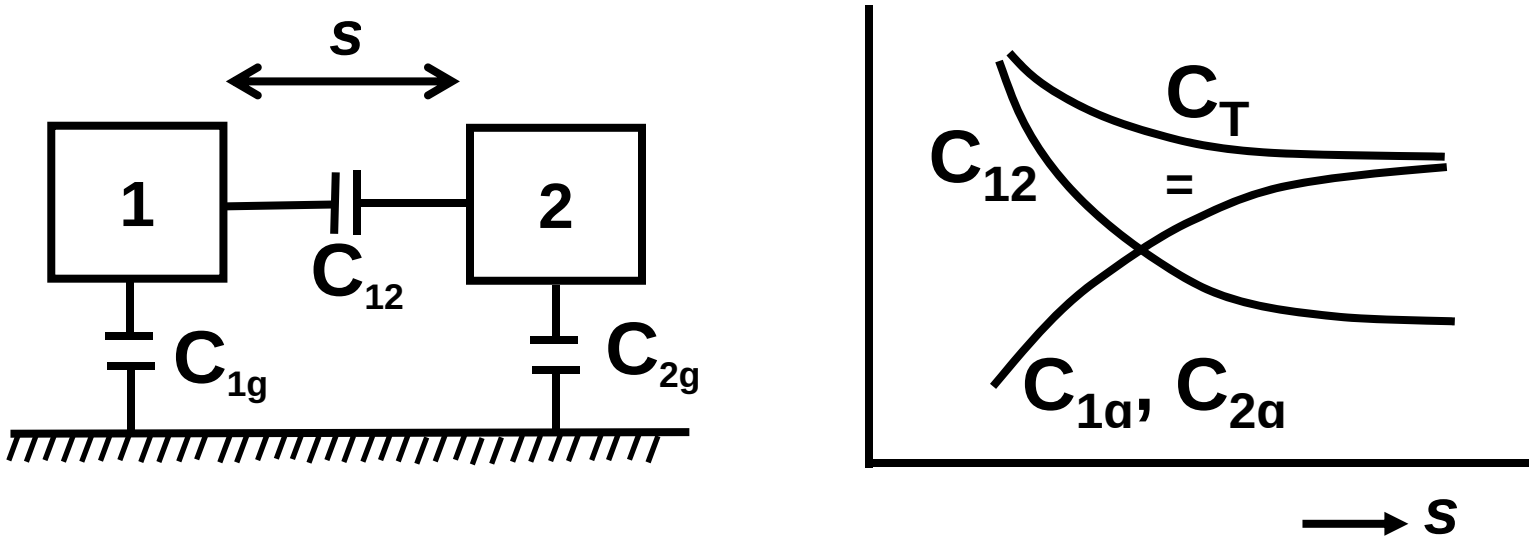
M1 over Active	M1 over Field	Unit
$C_a = 41$	$C_a = 30$	$aF/\mu m^2$
$C_p = 47$	$C_p = 40$	$aF/\mu m$

Coupling Capacitances





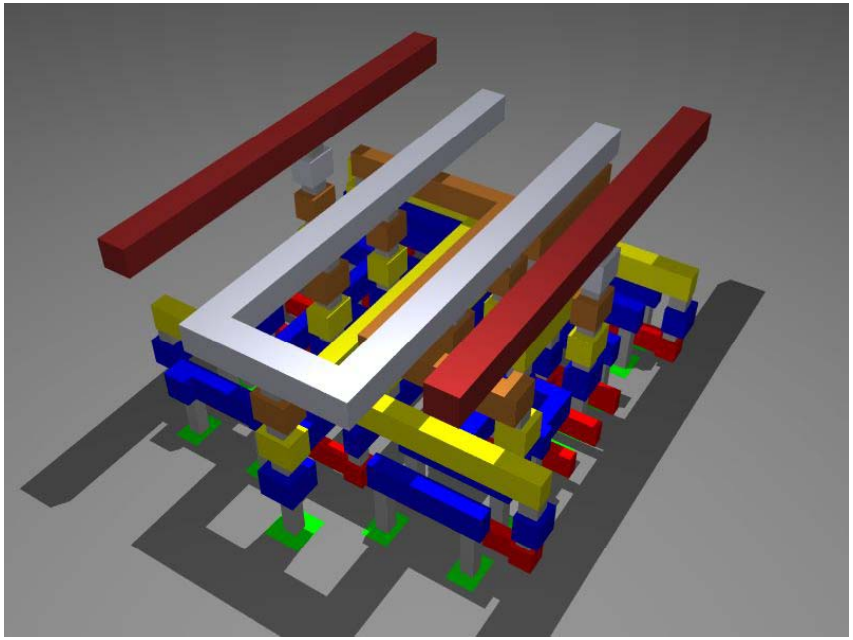
Coupling Capacitances (2)



- $C_T = C_{1g} + C_{12} = C_{2g} + C_{12}$ fairly constant
- Use that as first order model
- Interconnect capacitance design data (e.g. Table 4.3)

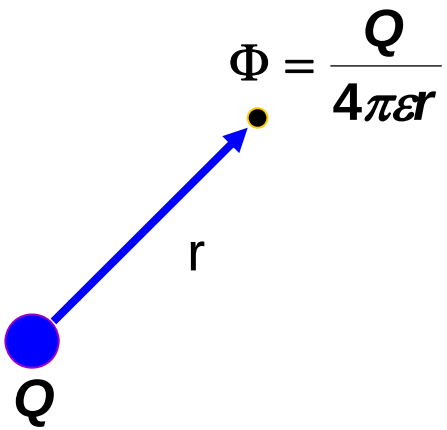
Field Solvers

- **Numerical, physics based technique for accurately computing capacitances**
- **Based on 3D geometry**
- **Finite element method**
- **Finite difference method**
- **Boundary element method**



Solve huge sets of equations, fast

BEM for Capacitance Computation



Electrostatic potential due to point charge

$$\Phi(p) = \int_{\text{all charge}} G(p,q)\xi(q)dq$$

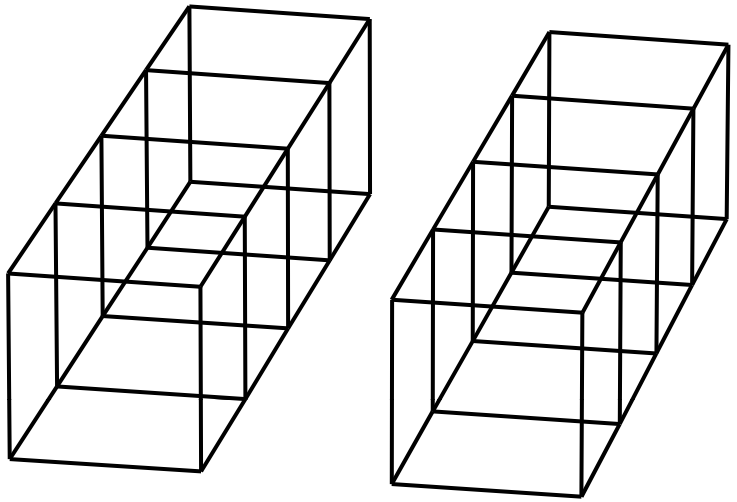
Electrostatic potential due to charge distribution

Green's function G(p,q):
potential at a point in space (x_p,y_p,z_p) due to unit point charge at other point (x_q,y_q,z_q).

Boundary Element Method, Discretization

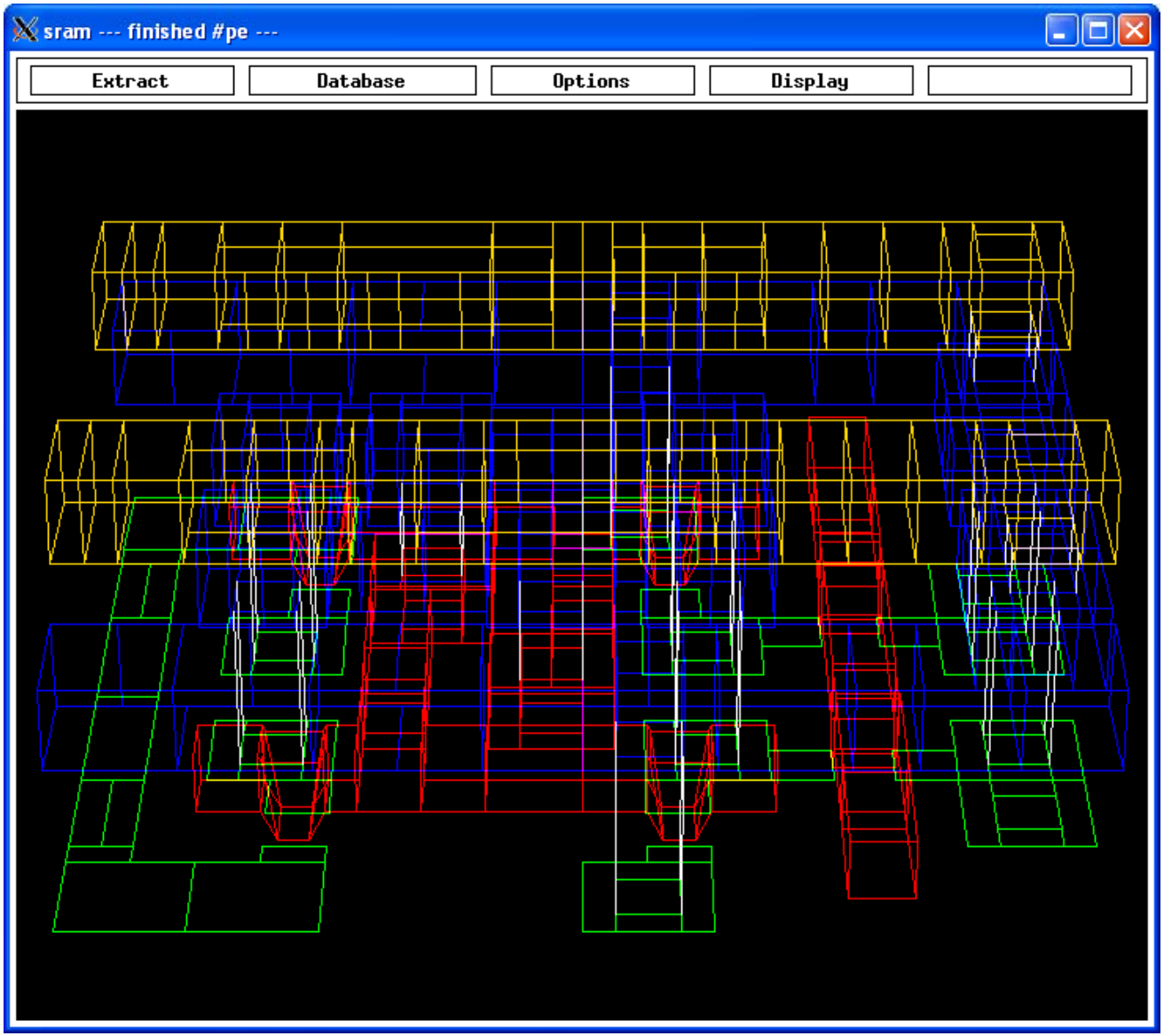
$$Q = F^T G^{-1} F \Phi$$

$$C_S = F^T G^{-1} F$$



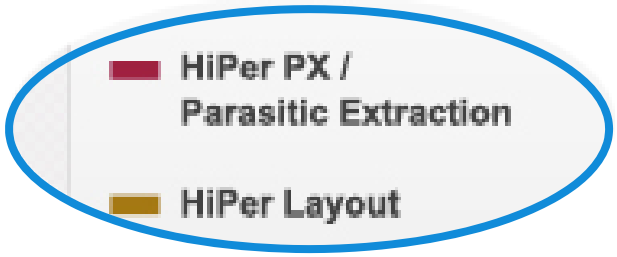
- Q: vector of conductor charges
- Φ: vector of conductor potentials
- Q: vector of panel (discretization element) charges
- ϕ: vector of panel potentials
- F: incidence matrix relating panels to conductors
  ($Q = F^T q$ and $\phi = F\Phi$)
- G_{ij}: potential of panel i due to charge at panel j
- C_S short-circuit capacitance matrix to be obtained

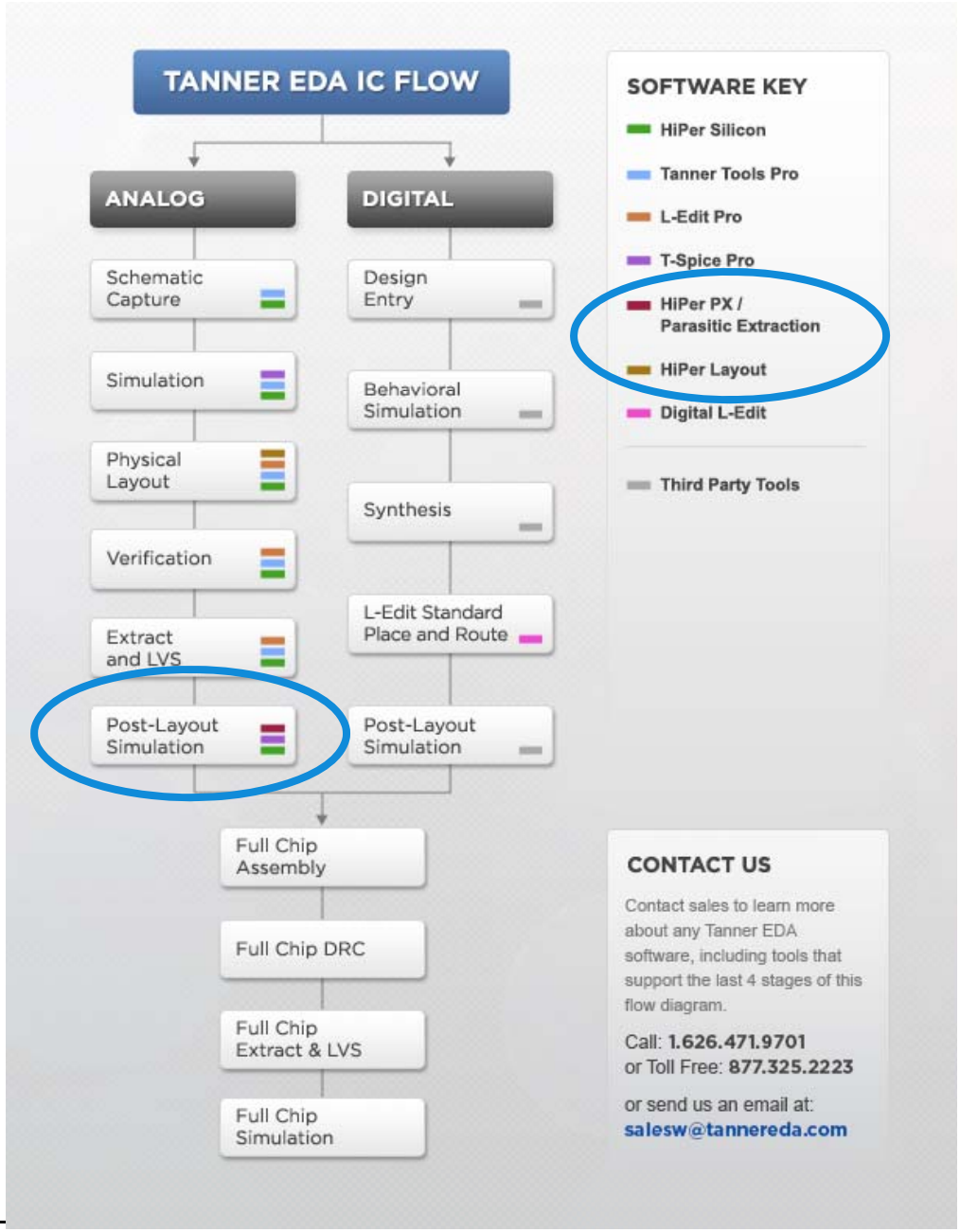
Demo



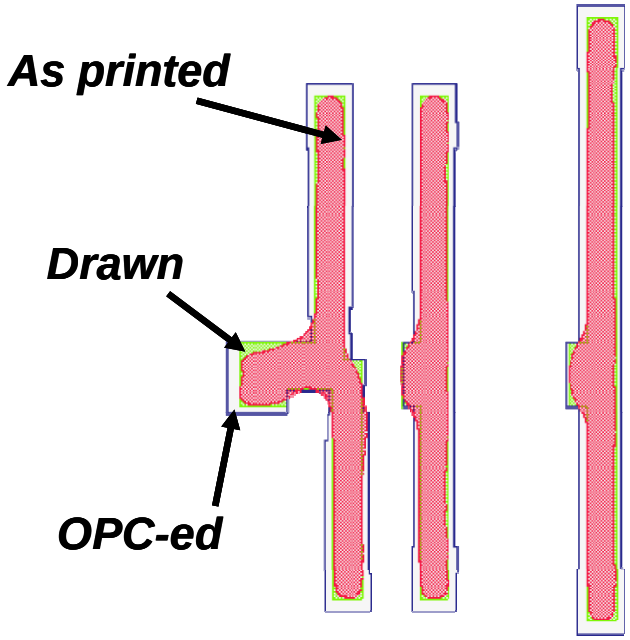
SPACE - OEM

- SPACE aka OptEM Inspector aka Tanner HiPer PX
- 2009 was slow
- Recently signs of more (eval) activity
- Sales can still be easily counted

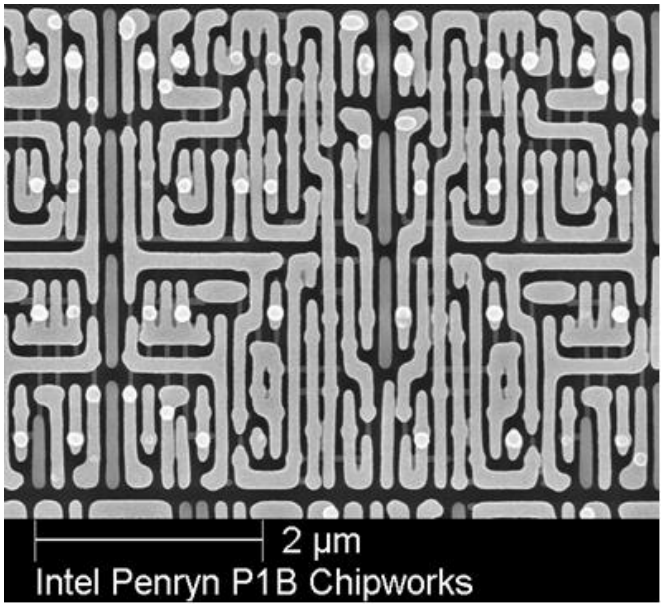


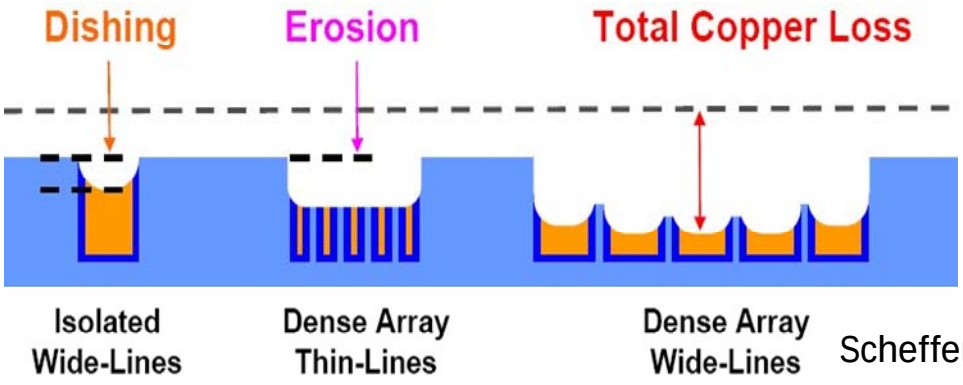


Manufacturing Variability – Field Solving needs Next Level....



Ren, 2006





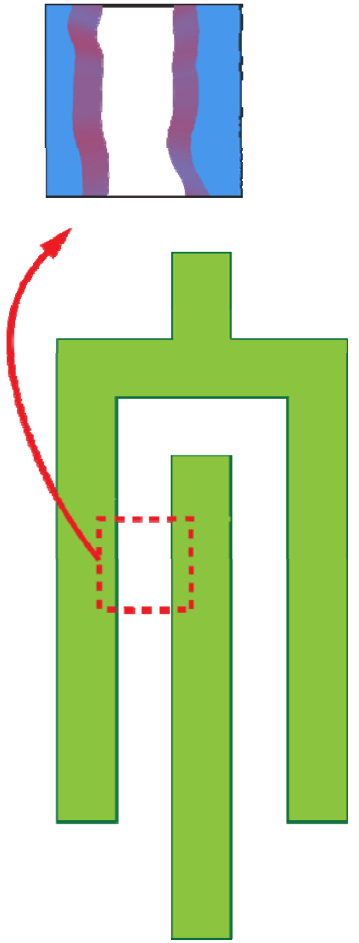
Scheffer, 2006

- Manufactured dimensions ≠ drawn dimensions
- Both **systematic** and **stochastic** variations
- Need for inclusion in **verification flow**

Interconnect Modeling with Variability

- Systematic variations – e.g. litho induced
- Stochastic variations – e.g. line edge roughness

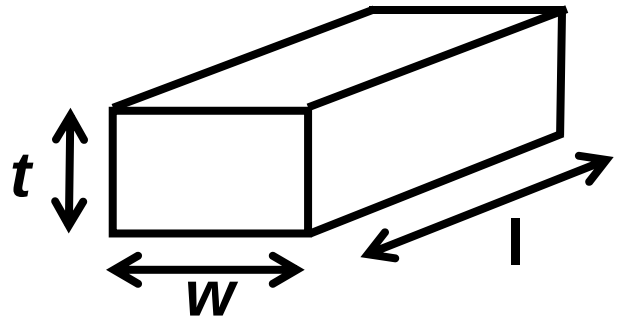
- Systematic aspects of resistance variability
- Stochastic variations of capacitance
- Measurement/validation campaign with/at Holst Centre
- Joint statistical modeling flow together with INESC-ID (affiliated to TU Lisbon) for Parametric Model Order Reduction



Resistance

- **Sheet resistance**

Wire Resistance



- **Proportional to l**
- **Inversely proportional to w and t (cross-sectional area)**
- **Proportional to ρ: specific resistance, material property [Ωm]**
- **$R = \rho l/wt$**
- **Aluminum: $\rho = 2.7 \times 10^{-8}$ Ωm**

  Copper: $\rho = 1.7 \times 10^{-8}$ Ωm

Sheet Resistance

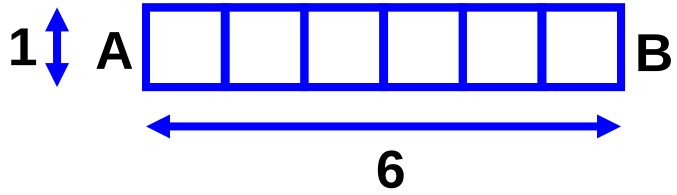
- **$R = \rho l/wt$**
- **t, ρ constant for layer, technology**
- **$R = Rl/w$**
- **R : sheet resistance [Ω/]**

  resistance of a square piece of interconnect

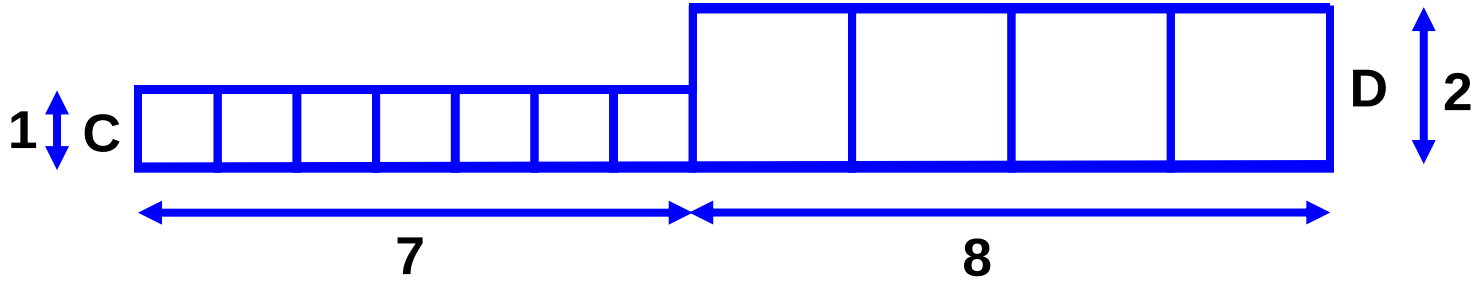
  other symbol: R_s
- **Interconnect resistance design data e.g. Table 4.5 (or inside back-cover)**

Interconnect Resistance

- **Assume $R_{\square} = 40\ \Omega$**
- **Estimate the resistance between A and B in the wire below.**



$R_{AB} \approx 6 \times 40 = 240\ \Omega$



$R_{CD} \approx 11 \times 40 = 440\ \Omega$

Engineering is about making controlled approximations to something that is too complicated to compute exactly, while ensuring that the approximate answer still leads to a working (functional, safe, …) system

Exercise

An interconnect line is made from a material that has a resistivity of $\rho = 4\ \mu\Omega$-cm. The interconnect is 1200 Å thick, where 1 Angstrom (Å) is 10^{-10} m. The line has a width of 0.6 μm.

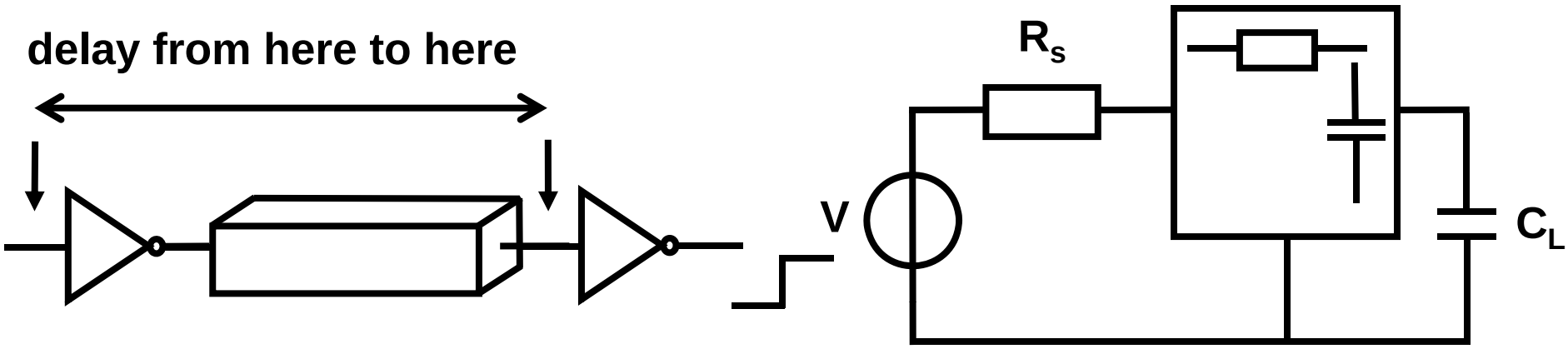
a) Calculate the sheet resistance $R_{\square}$ of the line.

b) Find the line resistance for a line that is 125 μm long.

Interconnect delay

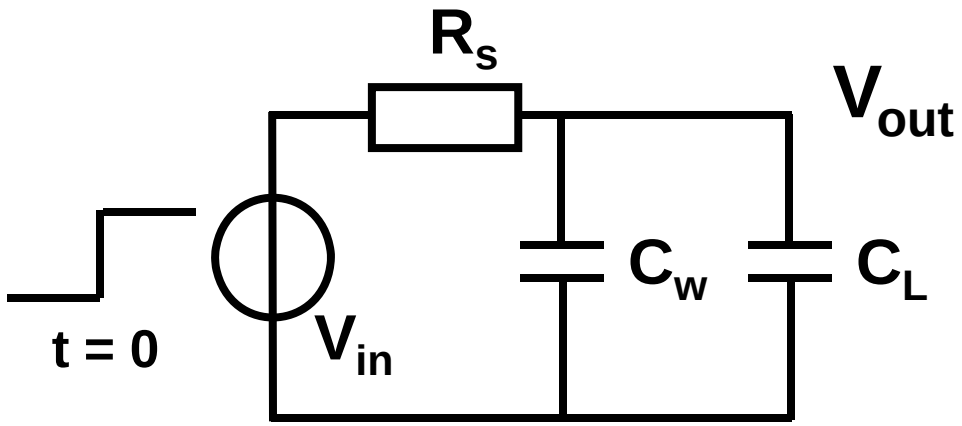
- **Delay metrics, rc delay, Elmore delay**

Delay



- **Model driver as linearized Thevenin source V, R_s, assume step input**
- **Model load as C_L**
- **Wire is an RC network (two-port)**

Wire Capacitance



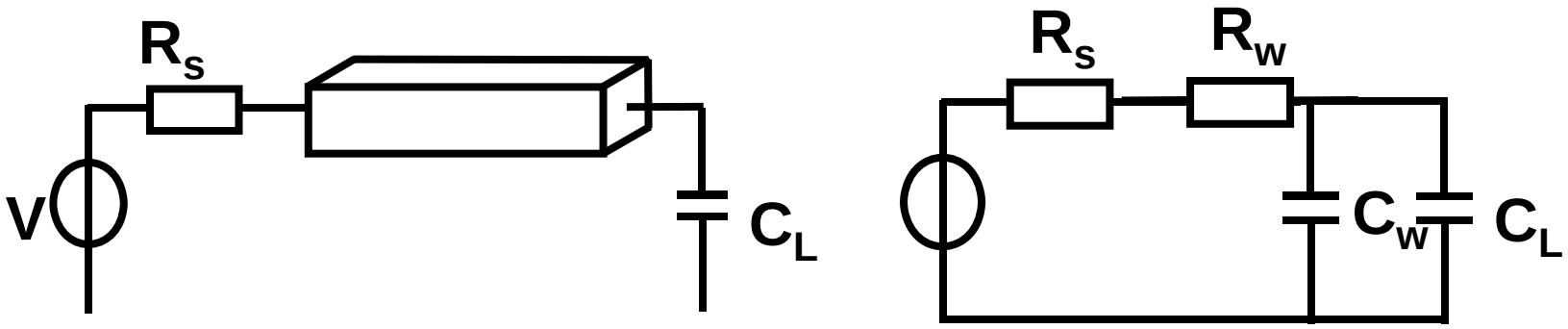
Assume wire behaves purely capacitive

$$(C_w + C_L)\frac{dV_{out}}{dt} + \frac{V_{out} - V_{in}}{R_s} = 0$$

$$V_{out} = V_{in} - \tau\frac{dV_{out}}{dt} \qquad \tau = R_s(C_w + C_L)$$

$$V_{out} = \left(1 - e^{-t/\tau}\right)V_{in}$$

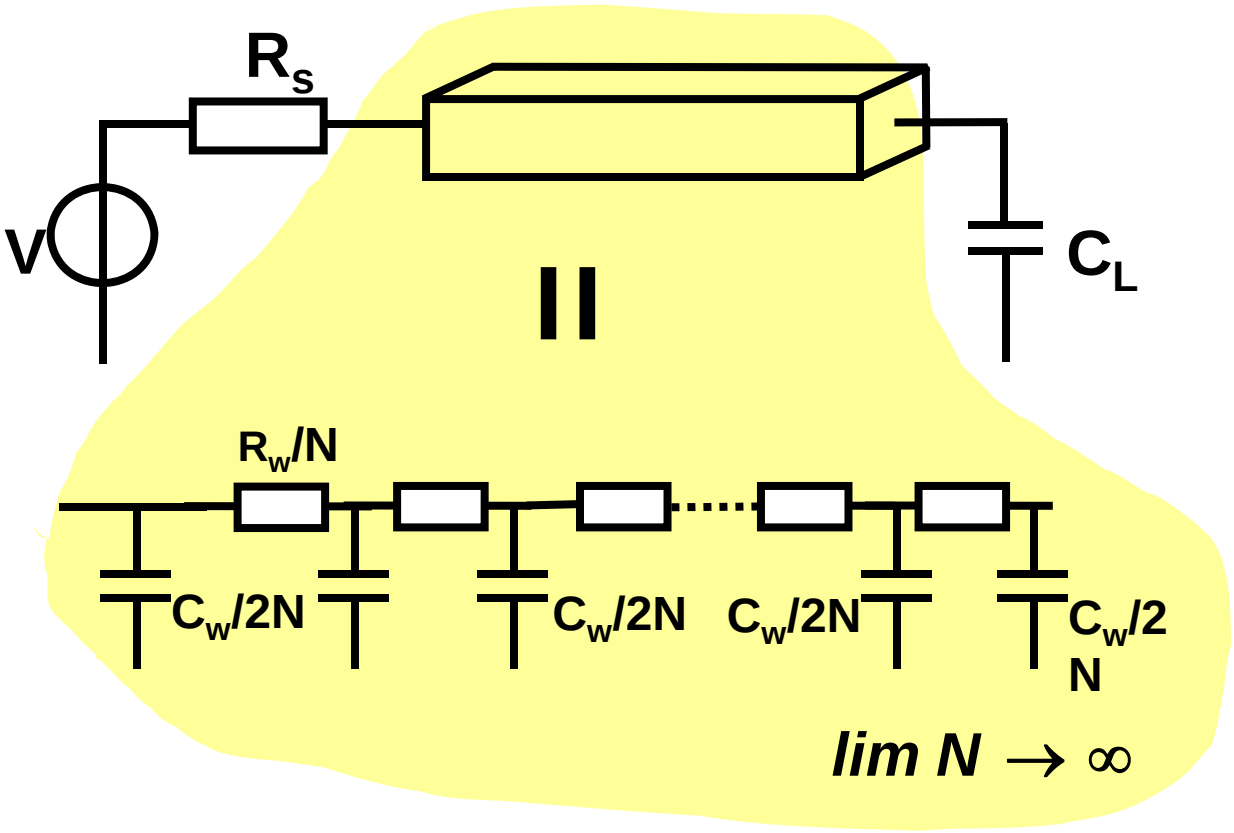
Wire Resistance



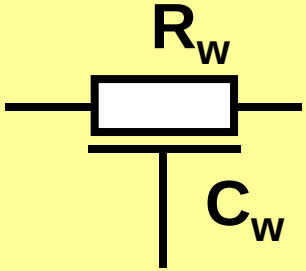
Now, assume wire capacitance *and resistance*

- $\tau = (R_s + R_w)(C_w + C_L)$
- **Not a a good model**
- **R and C are distributed along the wire**

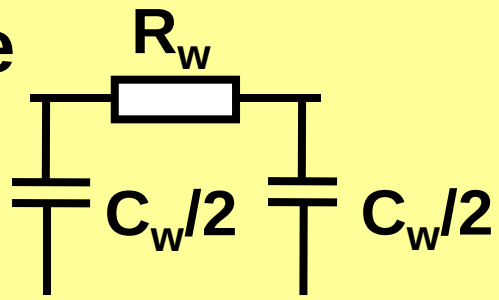
Uniform RC Line



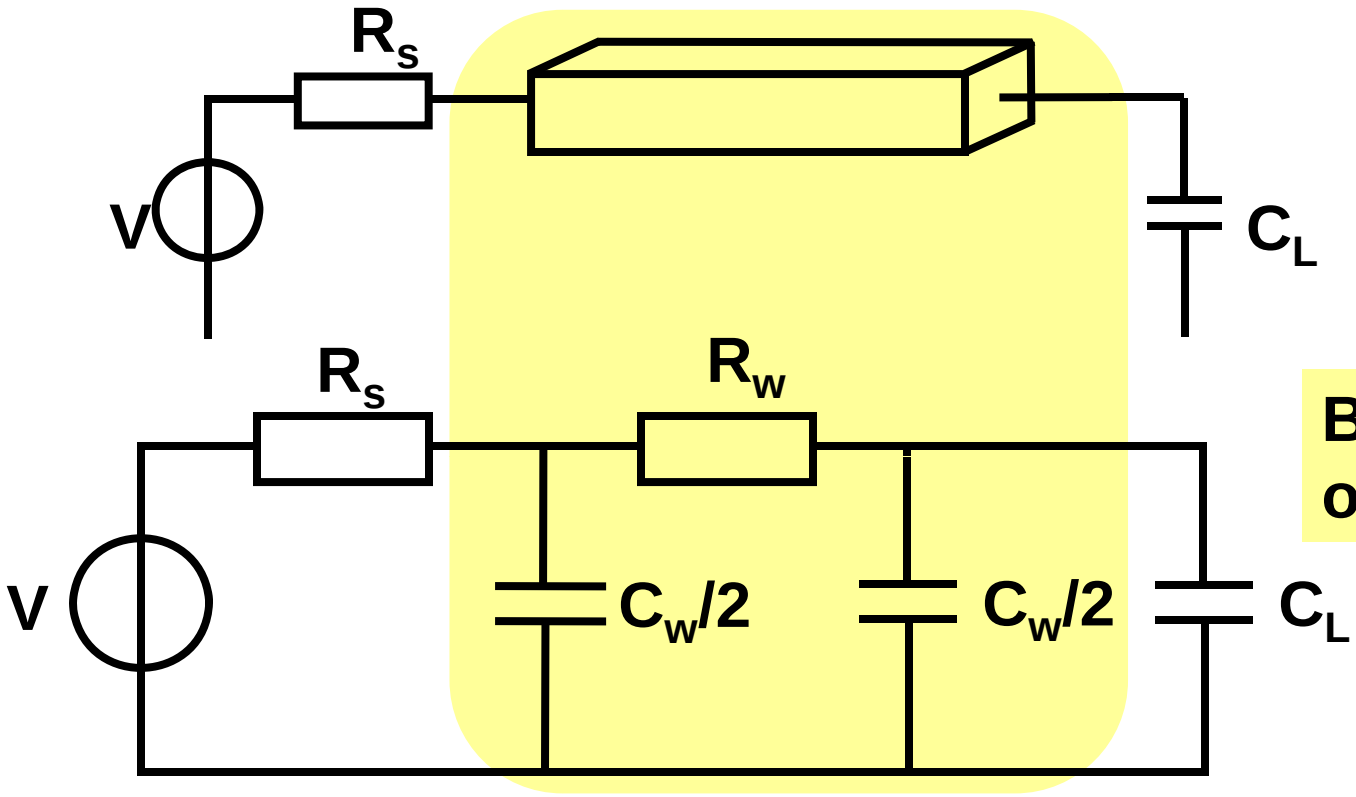
Symbol



[Always] use first order model



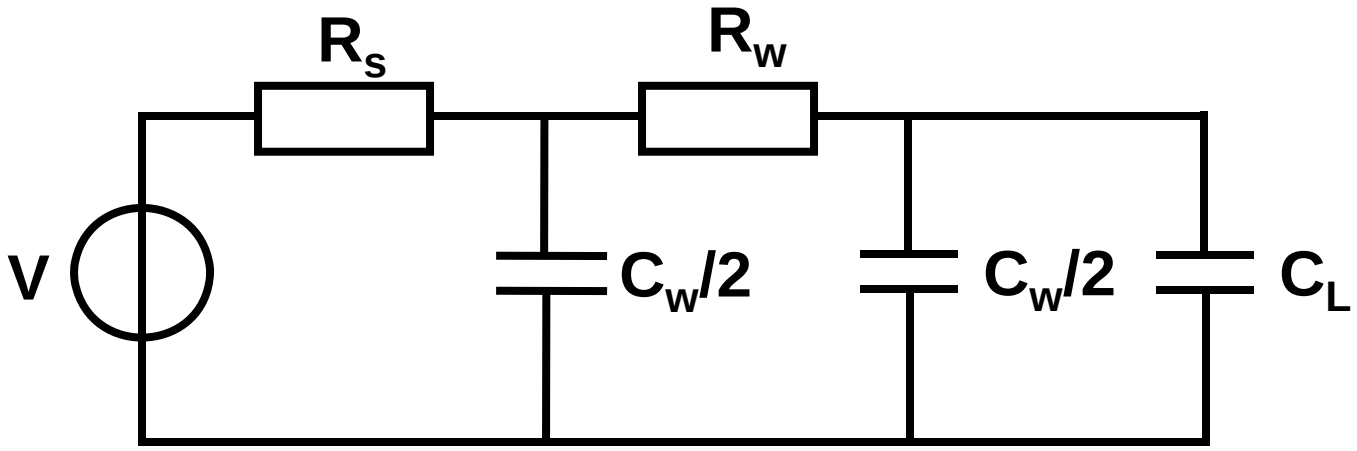
RC Delay (Uniform RC Line)



Basic π-model of wire

- **Not 1 τ**
- **Elmore delay: equivalent effective τ**

Equivalent Time Constant



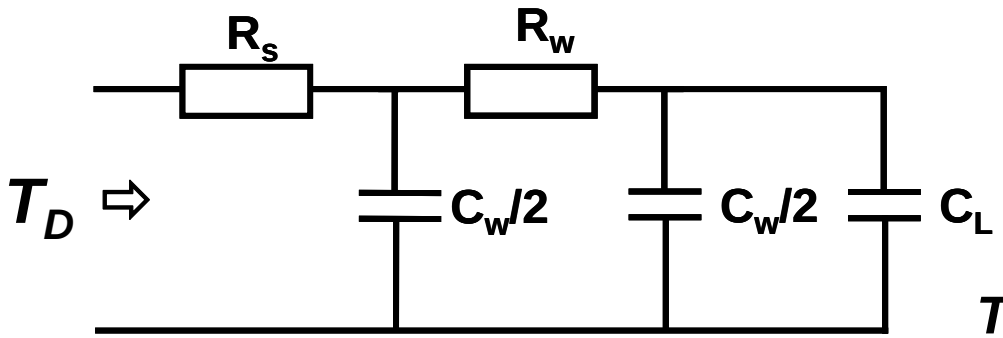
- Multiple time-constants
- Need for one “equivalent” number
- Offered by ***Elmore Delay*** T_D

$$T_D = R_S C_W / 2 + (R_S + R_W)(C_W / 2 + C_L)$$

- Effective “one number” model for delay

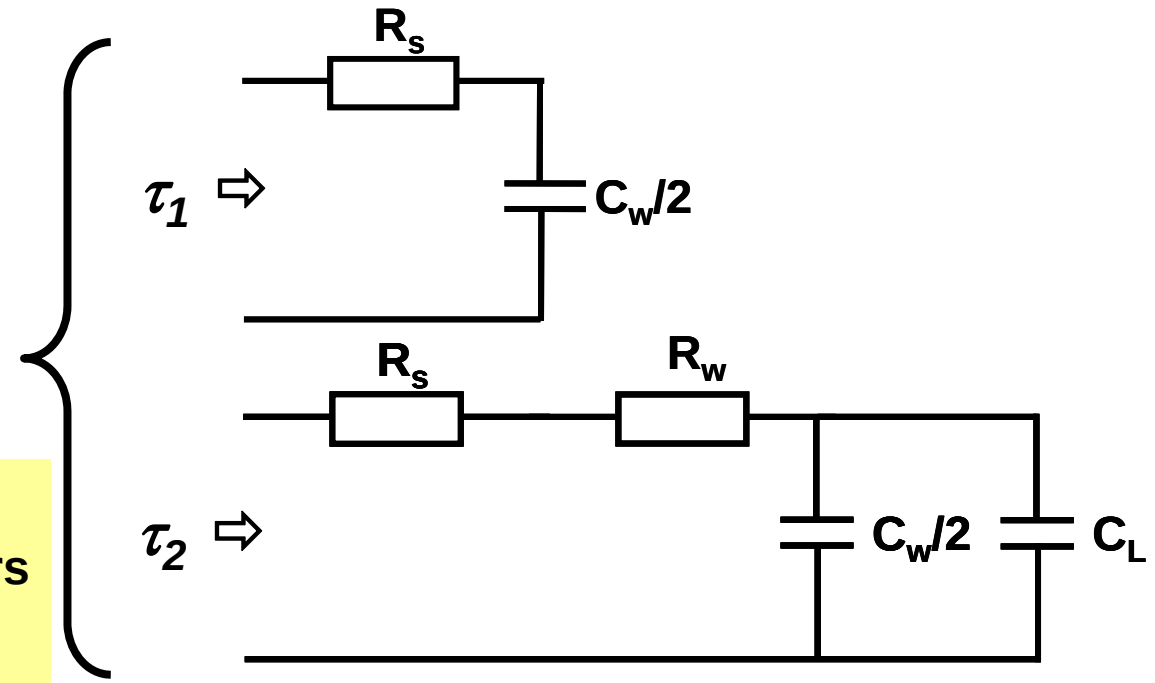
How to compute Elmore Delay?

Equivalent Time Constant



$$T_D = R_S C_W / 2 + (R_S + R_W)(C_W / 2 + C_L)$$

T_D = Σ

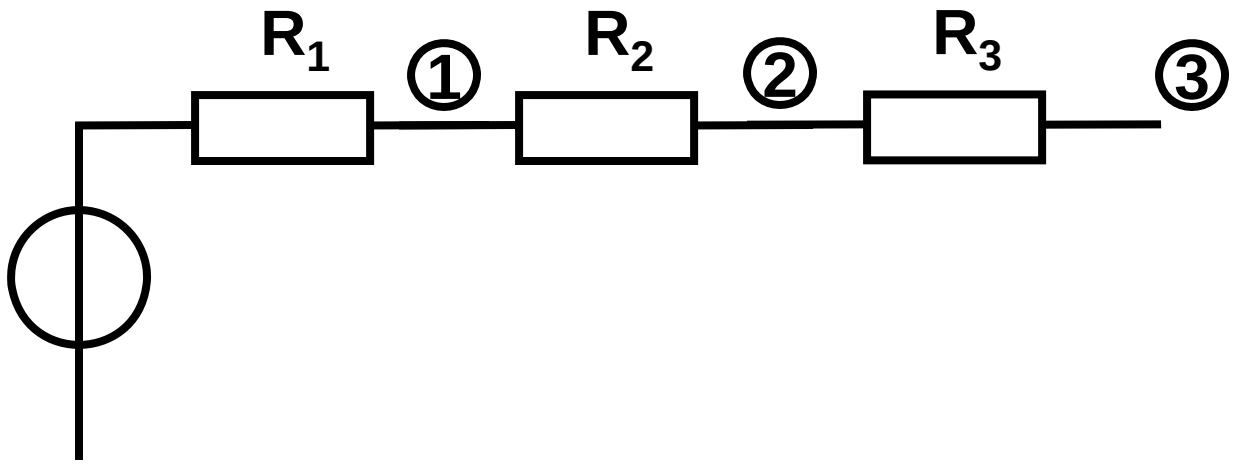


For each capacitor *i*,

Determine τ_i from resistors that (dis)charge C_i,

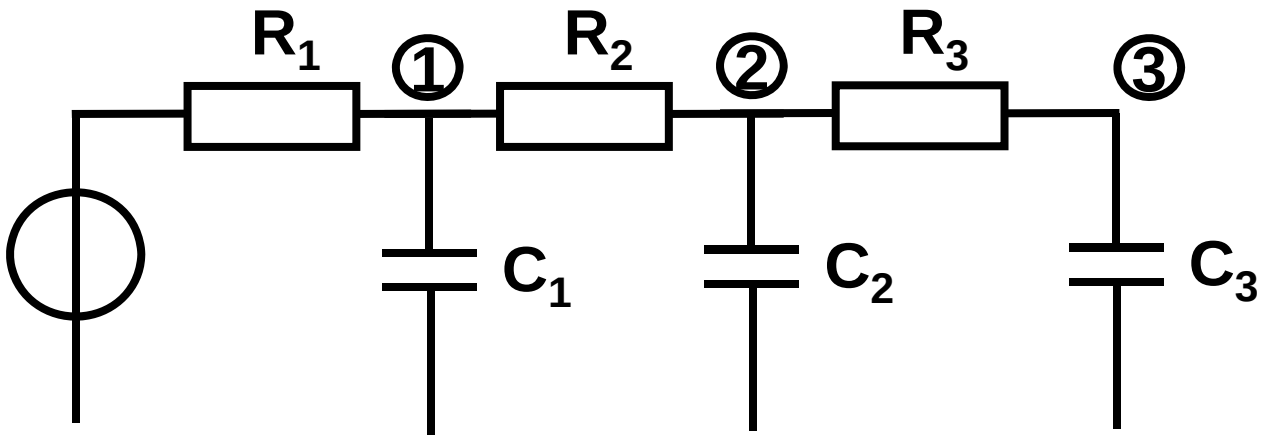
Sum these τ_i

Shared Path Resistance



- **Define: R_{ii} = Resistance from node i to input**
- **Example:** $R_{11} = R_1 \quad R_{22} = R_1 + R_2 \quad R_{33} = R_1 + R_2 + R_3$
- **Define: R_{ik} = Shared path resistance to input for node i and k**
- $R_{12} = R_1 \quad R_{13} = R_1 \quad R_{23} =$ 

Elmore Delay for RC ladders



- **Define:** $T_{Di} = \sum_{k=1}^{N} R_{ik} C_k$

Elmore Delay

We will use 0.69 x T_{di} as approximation of wire delay ($t_{50\%}$)

- $T_{D1} = R_{11}C_1 + R_{12}C_2 + R_{13}C_3 =$

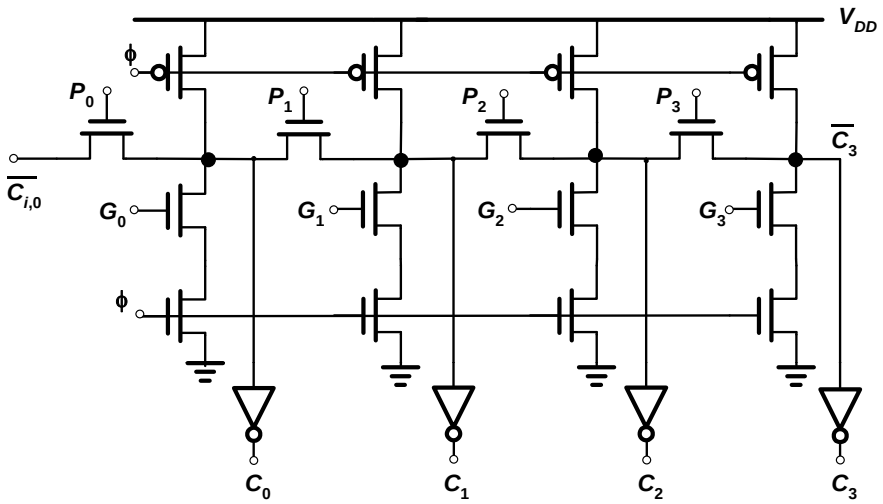
$= R_1C_1 + R_1C_2 + R_1C_3$

$T_{D3} = R_{31}C_1 + R_{32}C_2 + R_{33}C_3 =$

$= R_1C_1 + (R_1 + R_2)C_2 + (R_1 + R_2 + R_3)C_3$

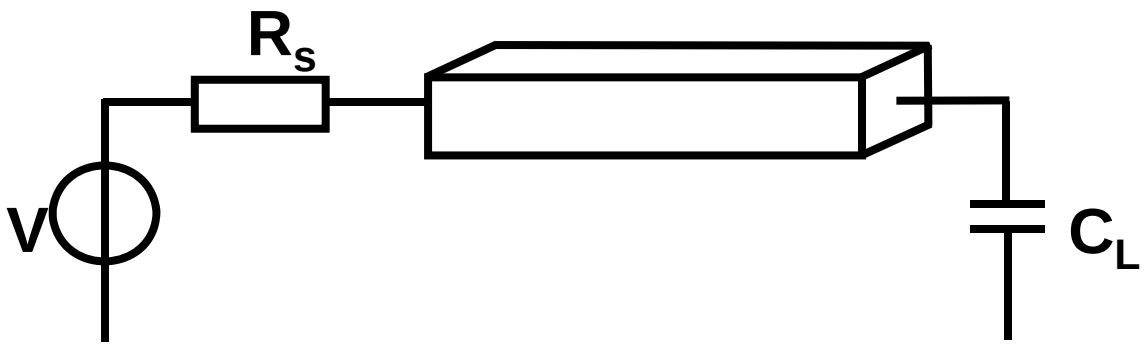
- $T_{D2} =$

Manchester Carry Chain Delay

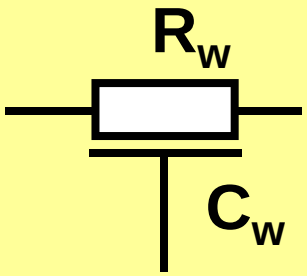


Given an expression of delay (symbols, not numbers) as a function of the number of bits

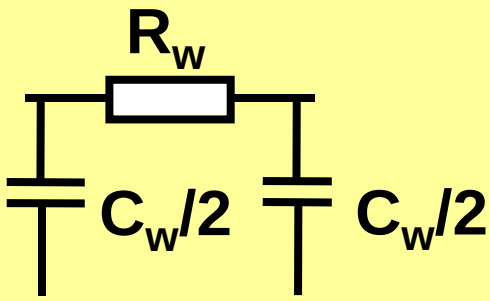
Elmore Delay for Distributed RC Lines



Symbol

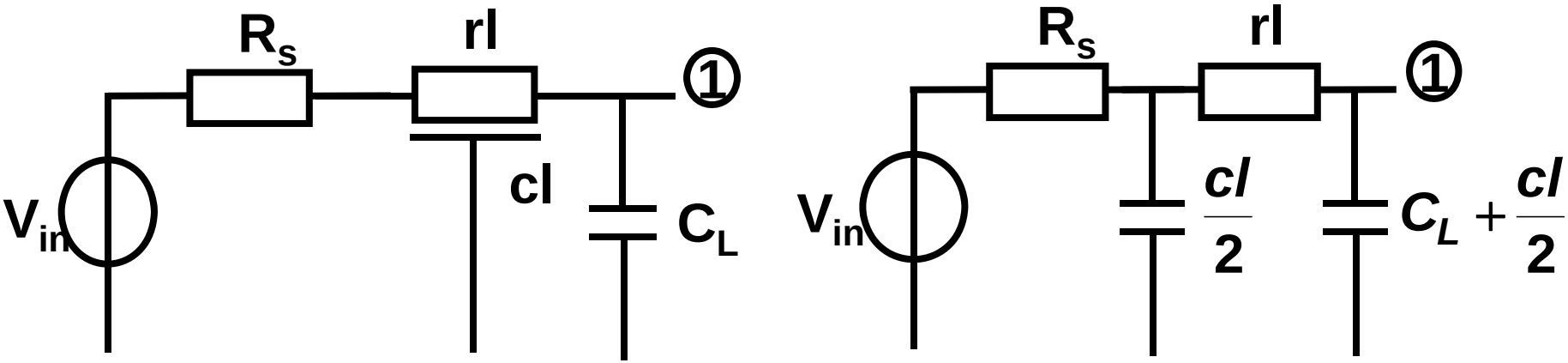


Symmetric π-model



- **Theorem: For Elmore Delay calculations, each uniform distributed RC section is equivalent to a symmetric π-model**

Canonical Driver-Line-Load



$$T_{D_1} = R_s \frac{cl}{2} + (R_s + rl)\left(C_L + \frac{cl}{2}\right)$$

$$= R_s(cl + C_L) + rlC_L + \frac{1}{2}rcl^2$$

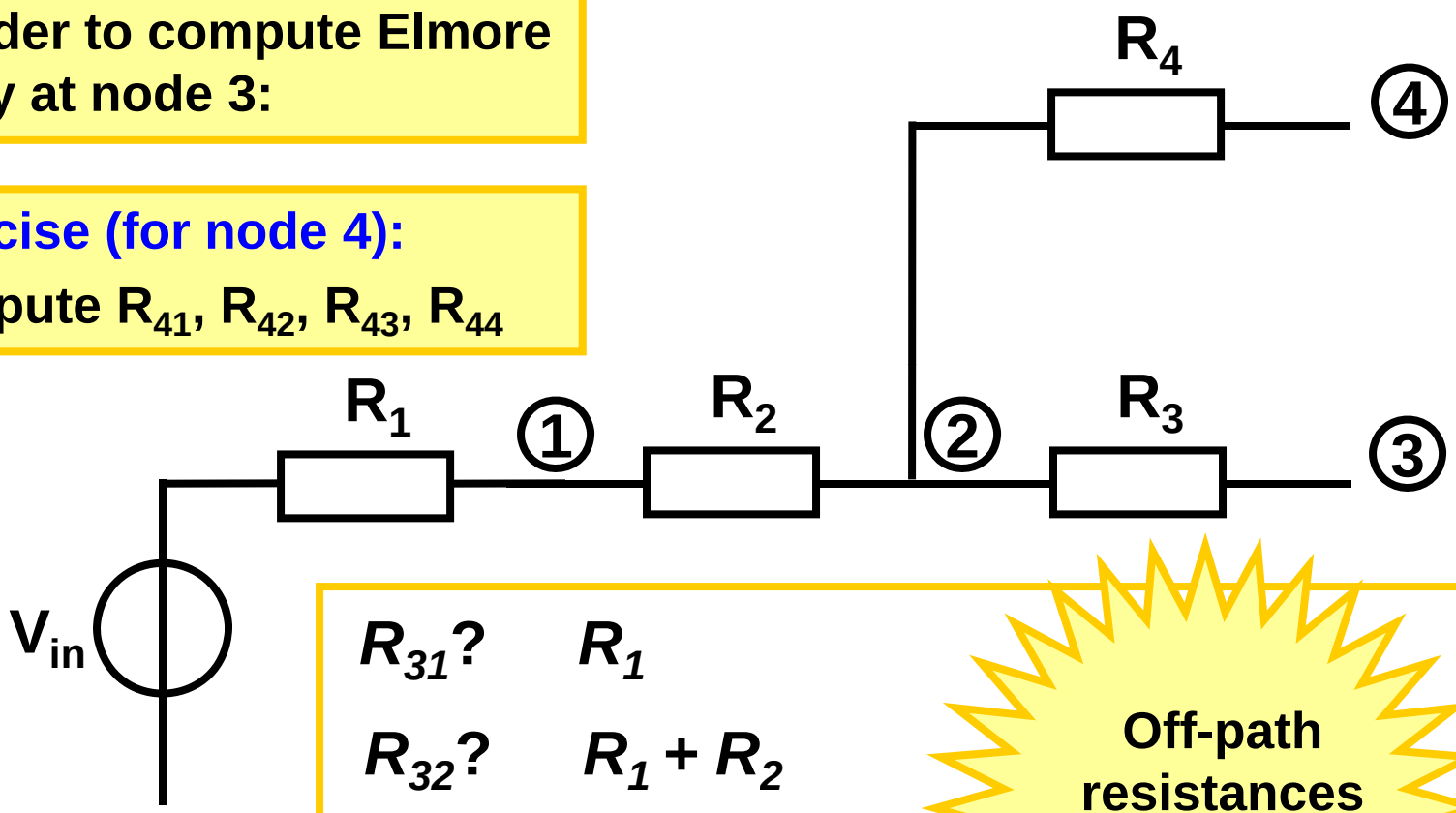
- **Delay quadratic in line length**

Shared Path Resistance for Tree Structures

In order to compute Elmore Delay at node 3:

Exercise (for node 4):

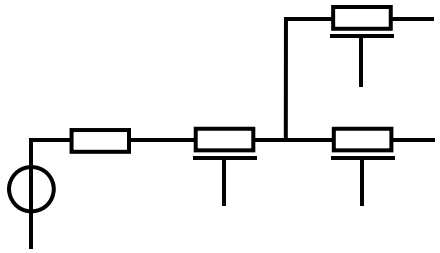
Compute R_{41}, R_{42}, R_{43}, R_{44}



R_{31}?	R_1
R_{32}?	$R_1 + R_2$
R_{33}?	$R_1 + R_2 + R_3$
R_{34}?	$R_1 + R_2$

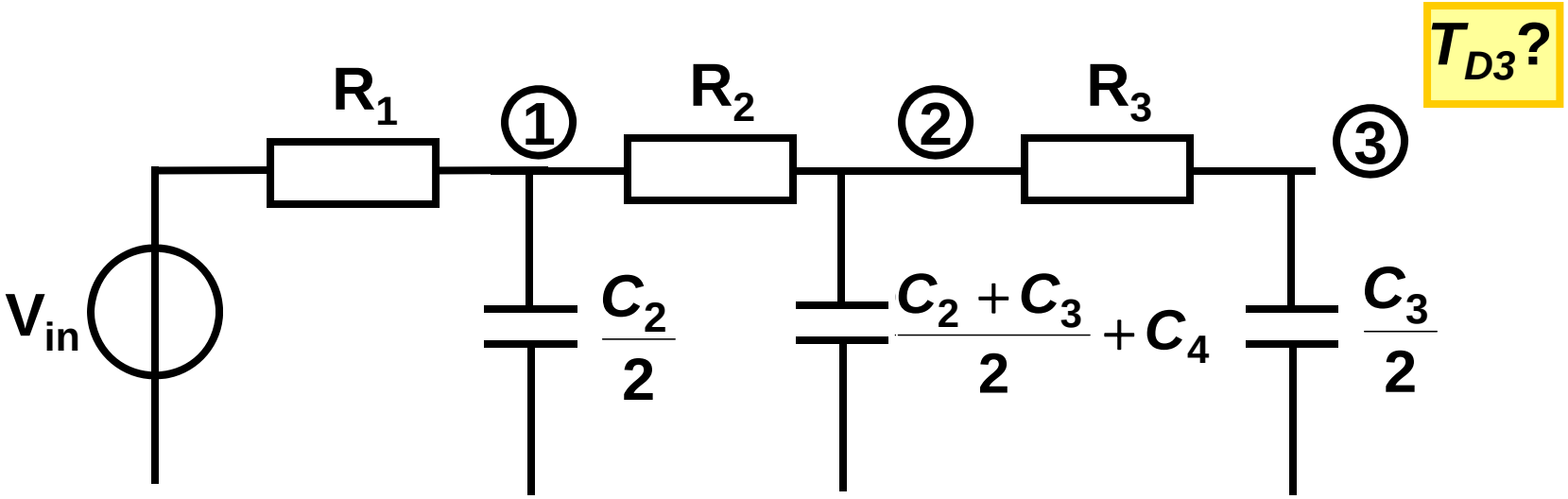
Off-path resistances don't count

Elmore Delay for Tree Structures



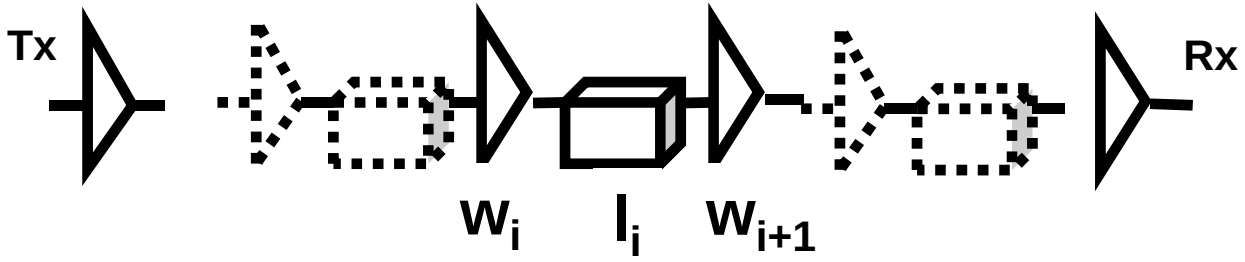
Exercise: Compute T_{D1}, T_{D2}, T_{D3}, T_{D4}

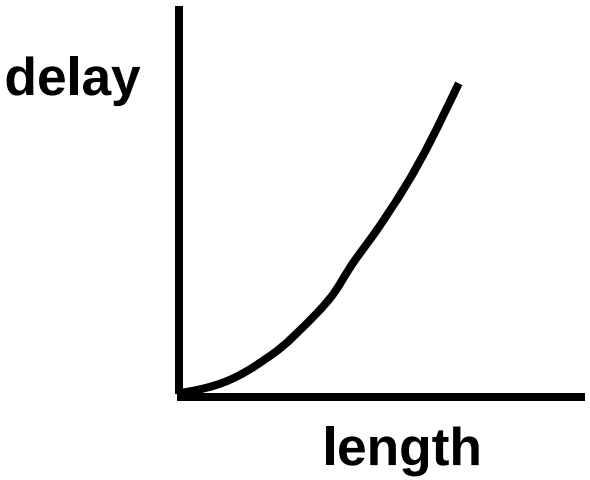
- **Replace RC lines by π-sections**
- **Given observation node i, then only resistances along the path from input to node i can possibly count**
- **Make others zero**
- **Compute as if RC ladder**

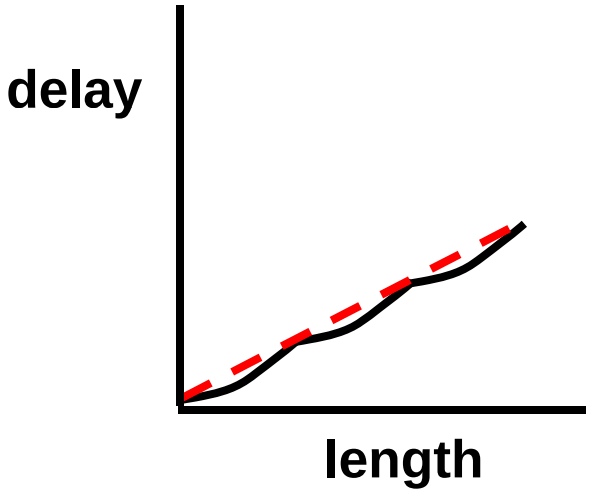


Quadratic Wire Delay

Becomes linear with repeaters/buffers at fixed intervals







Delay ~ rcL^2

Delay ~ $nrcl^2 = rcLl$

rc: independent of length

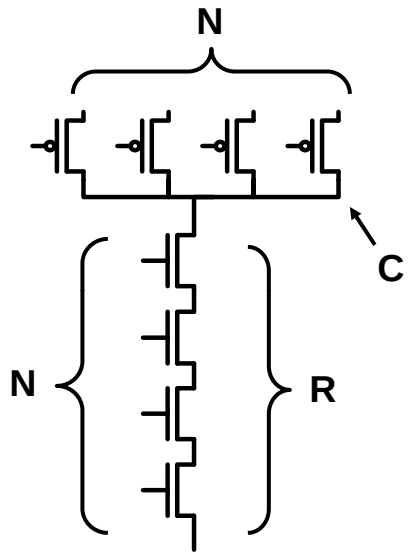
L: length of wire

l: length of segment

n: number of segments

Quadratic RC Delay Issues also in Logic

Avoid Large Fan-In



C linear in N

R linear in N

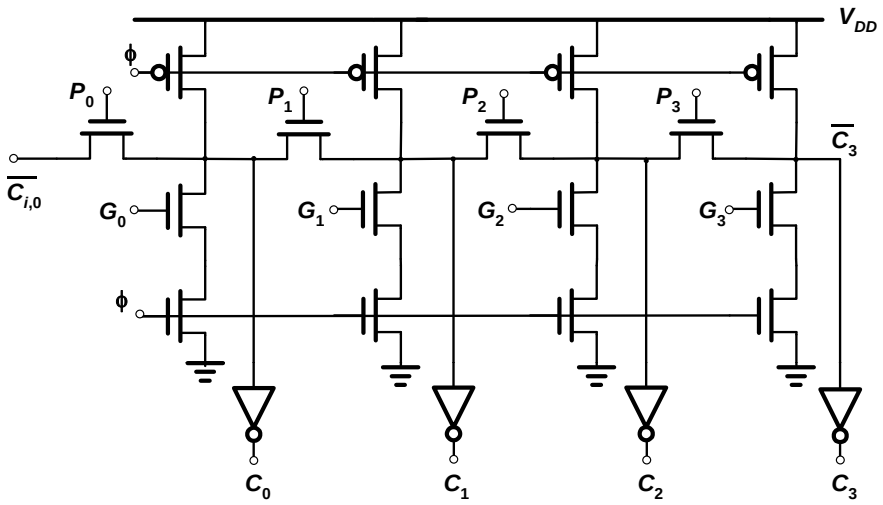
Delay ∝ RC quadratic in N

- **Transistor Sizing**
- **Progressive Transistor Sizing**
- **Input Re-Ordering**
- **Logic Restructuring**

Empirical

Delay = $a_1FI + a_2FI^2 + a_3FO$

Manchester Carry Chain Delay



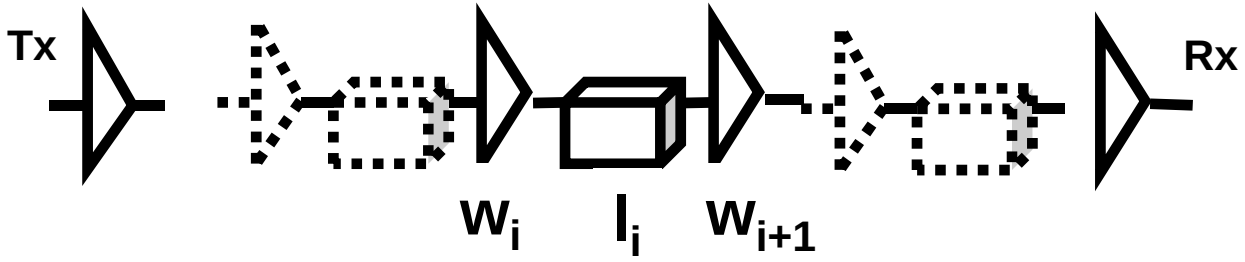
Given an expression of delay (symbols, not numbers) as a function of the number of bits -> $O(N^2)$

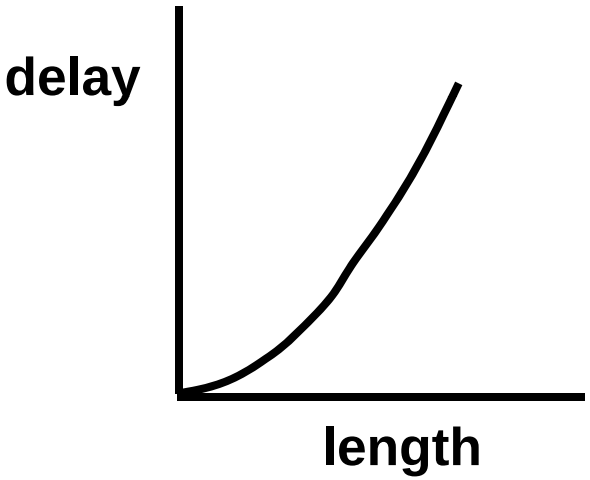
How to break this relationship?

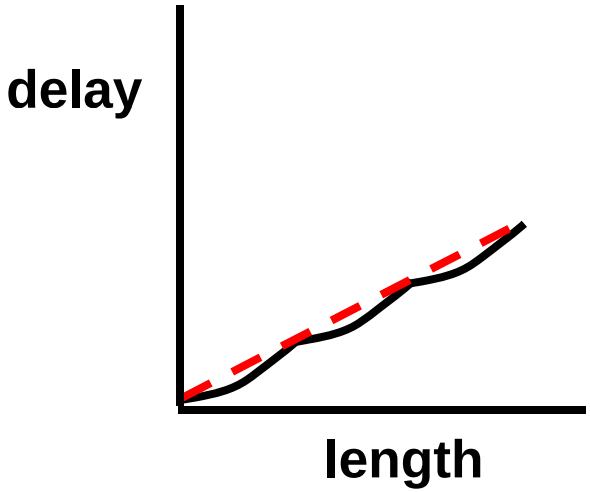
Insert restoring gate at regular intervals!

Quadratic Wire Delay

Becomes linear with repeaters/buffers at fixed intervals







Delay ~ rcL^2

Delay ~ $nrcl^2 = rcLl$

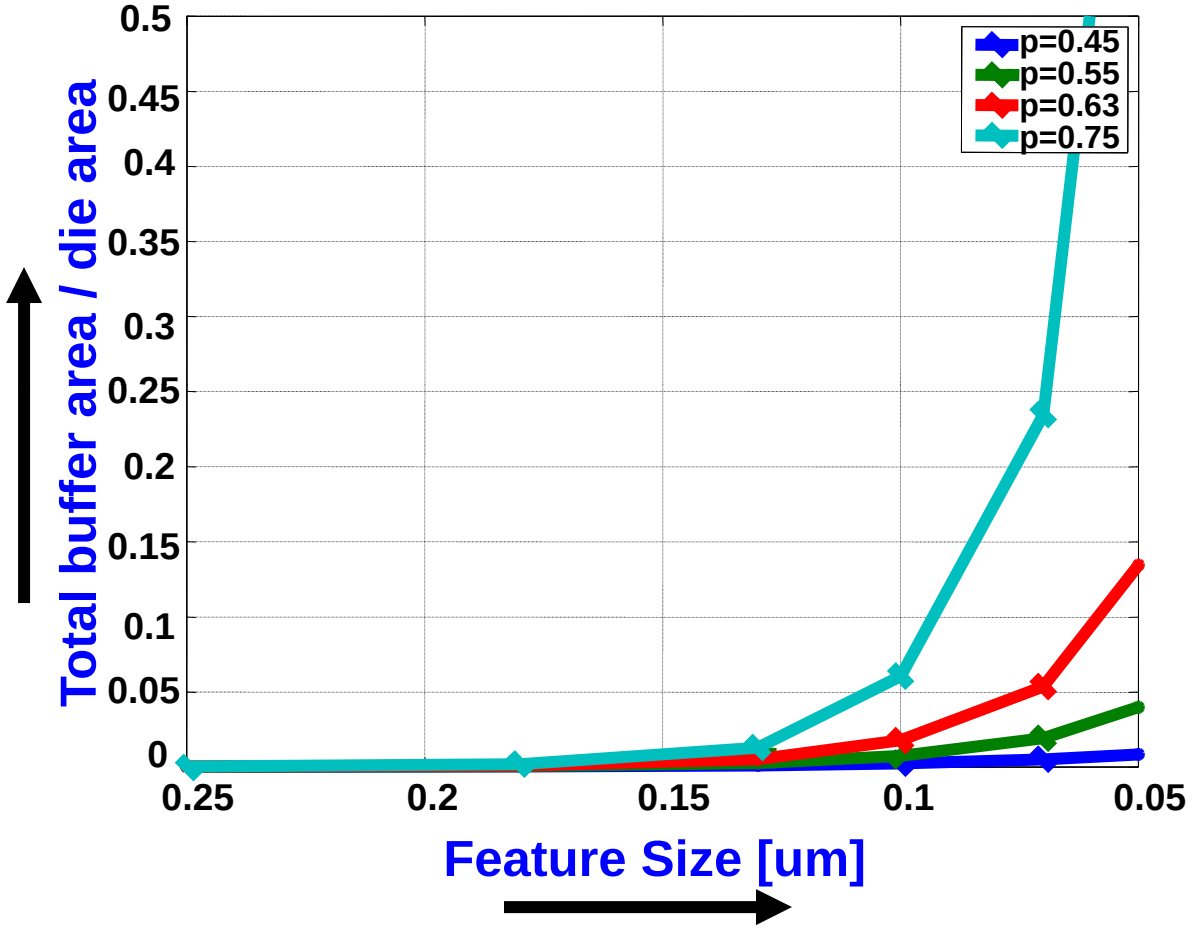
rc: independent of length

L: length of wire

l: length of segment

n: number of segments

Area Requirements for Optimal Buffering

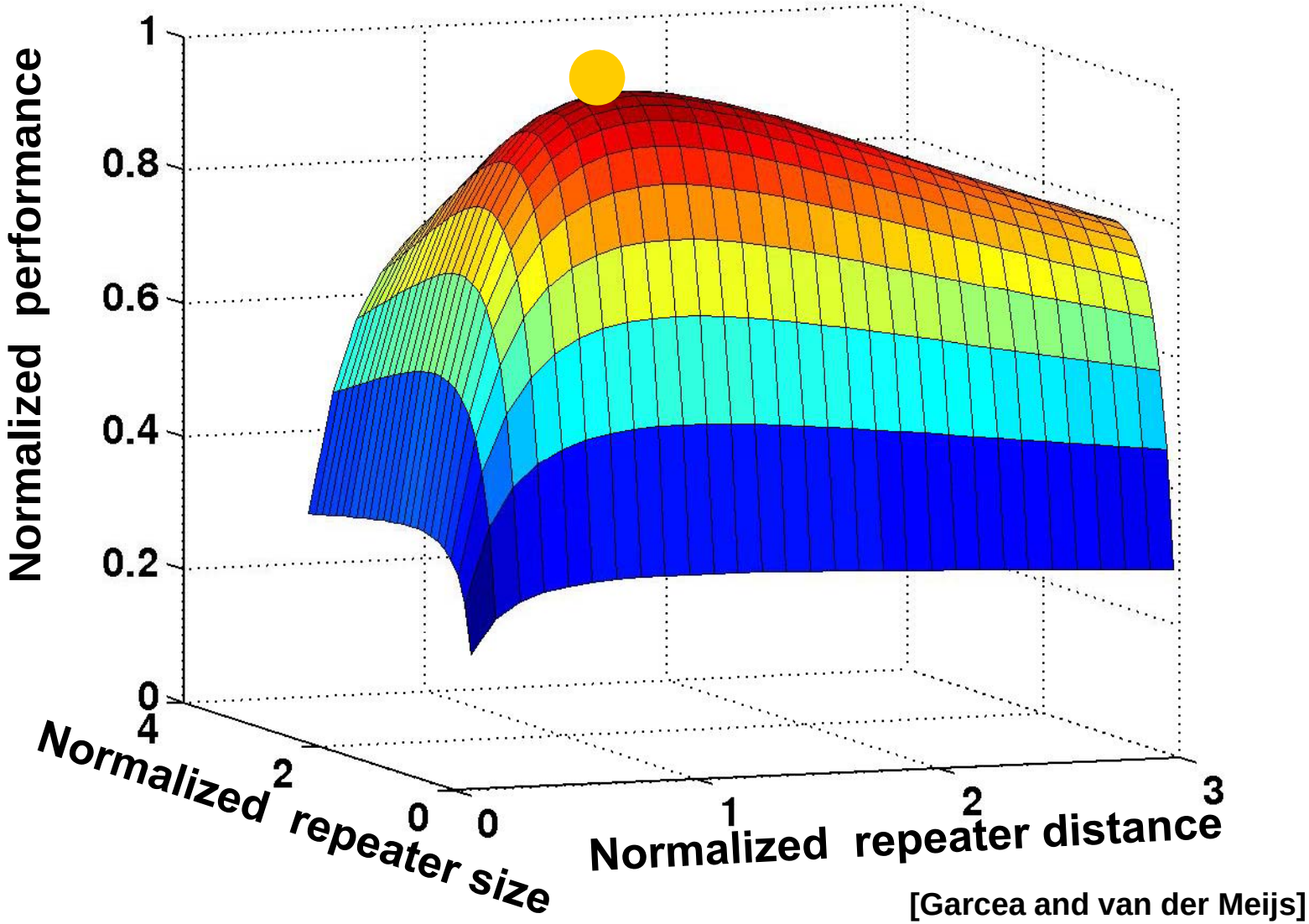


- **p** is the **Rent's exponent**
- **wire length distribution** according to **Davis**
- interconnect **topology** and **number of gates** taken from **ITRS**.

[Garcea and van der Meijs]

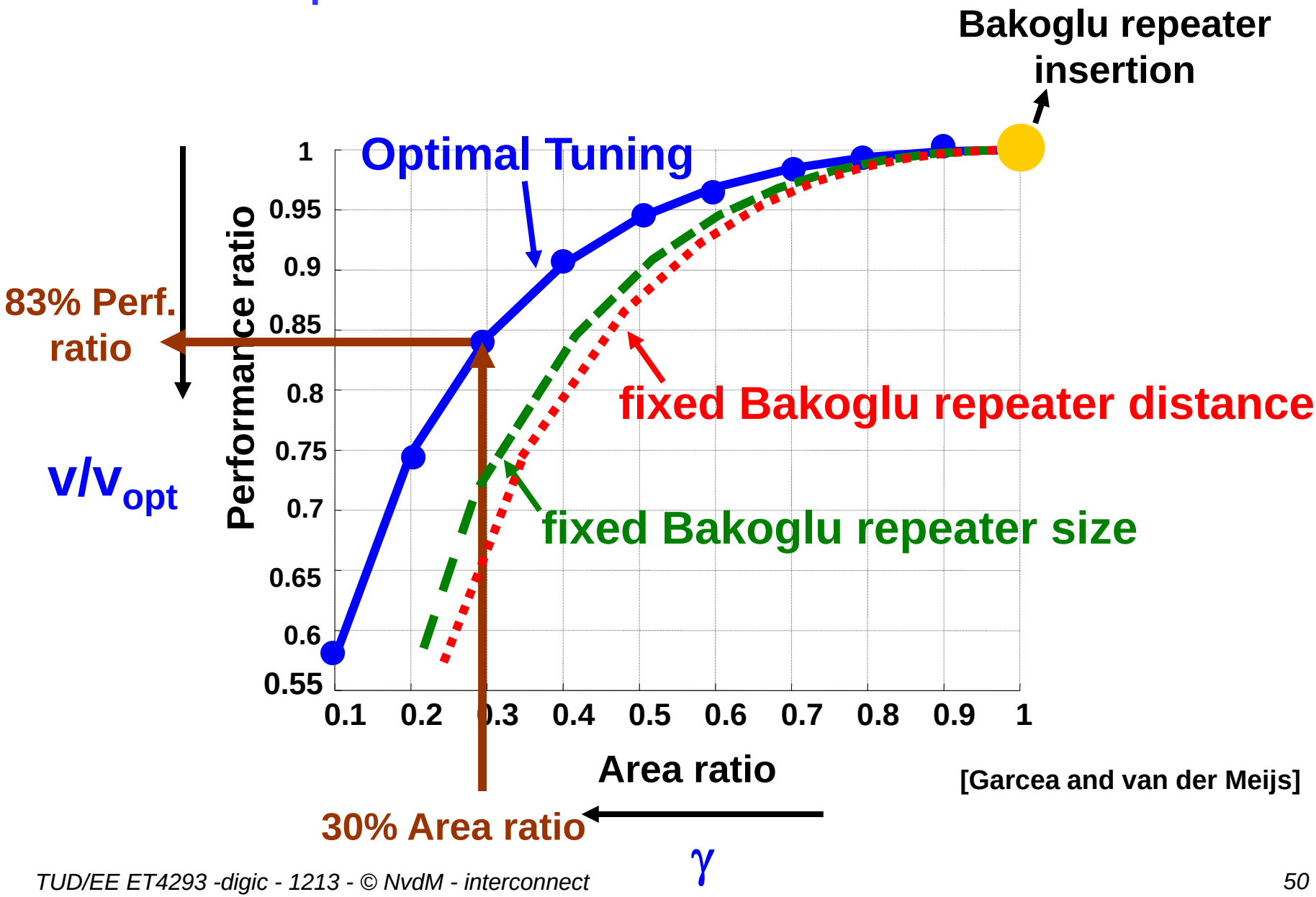
Optimal buffering is expensive in terms of area, up to 75% of the die area for a very complex circuit.

Performance vs. Repeater Size and Distance



[Garcea and van der Meijs]

Comparision to ad-hoc Insertion



[Garcea and van der Meijs]

Other Interconnect/Repeater issues

- **Optimal repeater sizing for power**
- **Optimal repeater sizing for minimizing effects of interconnect variability**
- **Optimize throughput of busses under area and/or power constraints**
- **Throughput of busses under variability**
- **….**

Summary

- **Capacitance**
  - **Area/perimeter model, coupling**
- **Resistance**
  - **Sheet resistance**
- **Interconnect delay**
  - **Delay metrics, rc delay, Elmore delay**