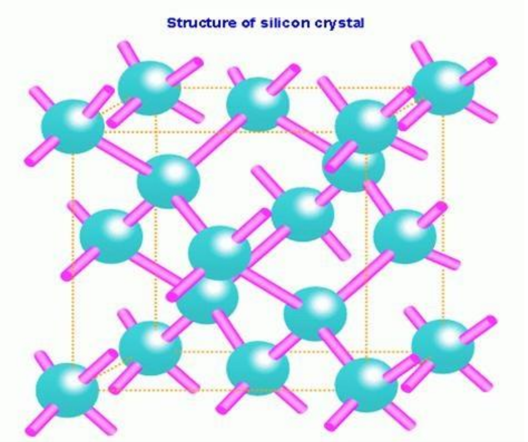
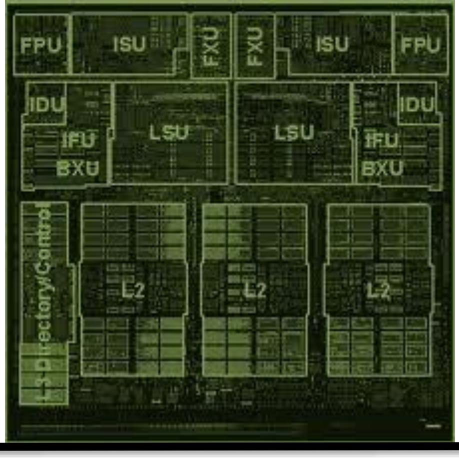
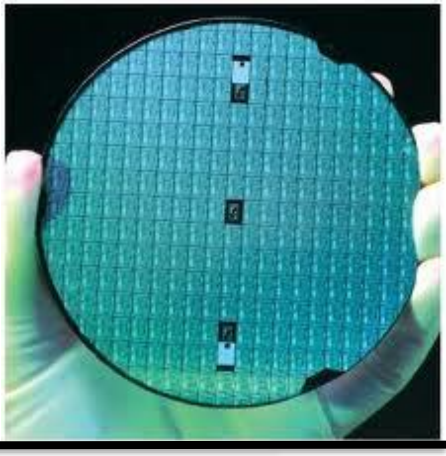






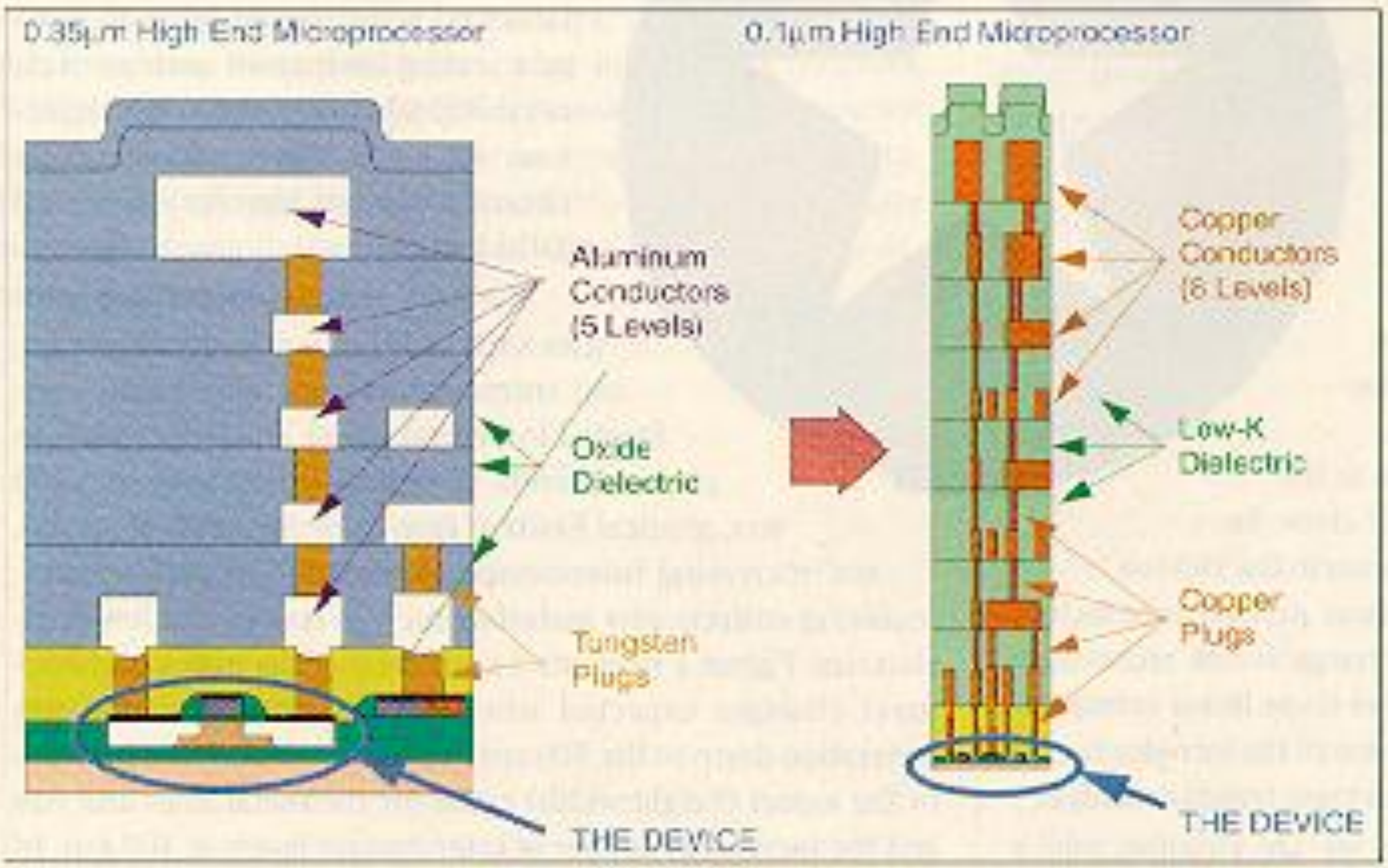
ECE 225
Lecture 5
VLSI Interconnects-I

Prof. Kaustav Banerjee
Electrical and Computer Engineering
University of California, Santa Barbara

Outline

- Interconnect and Technology Scaling
  - Interconnect technology: background and trends
  - Interconnect Parasitics: R-L-C
  - Elmore Delay Formalism
  - Interconnect Scaling: implications for technology and design

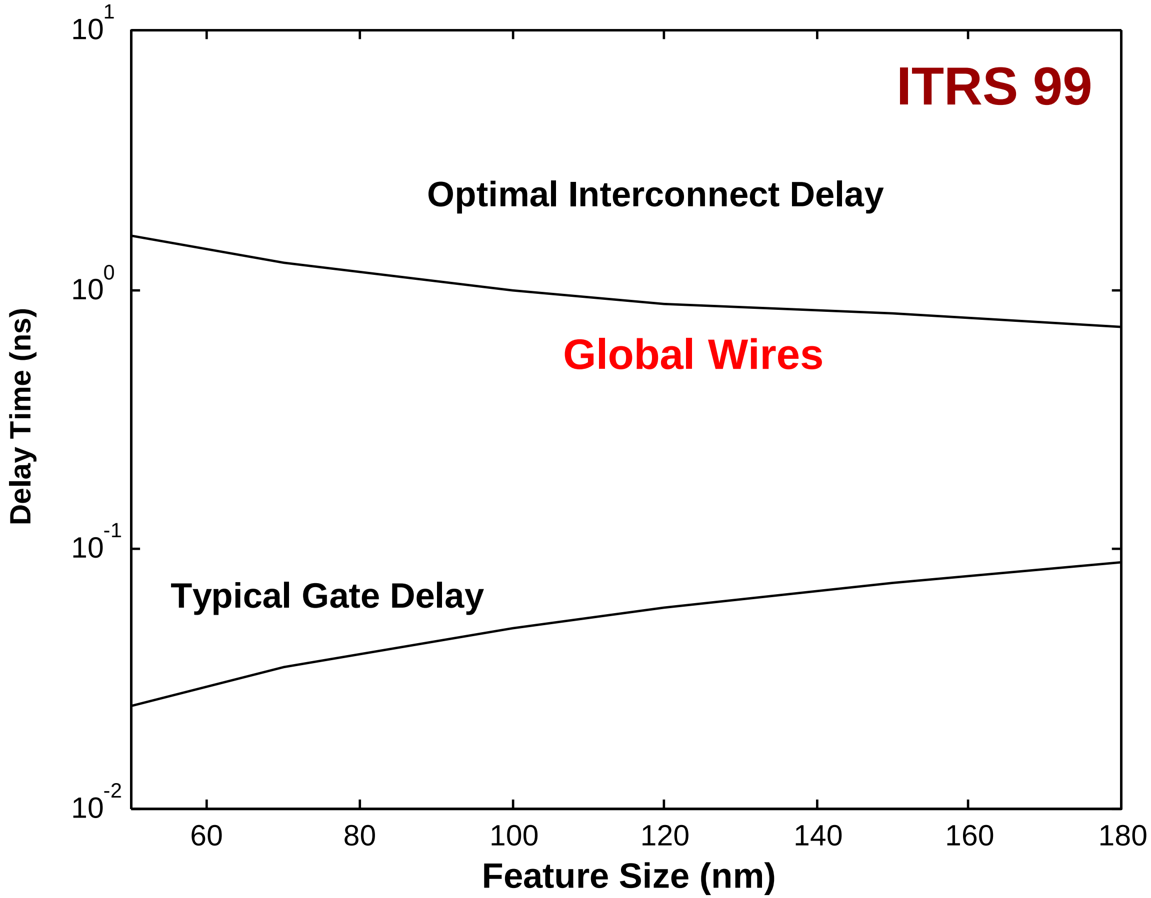
Interconnect Impact on Chip



1. Process architecture challenge.

Interconnect delay has become the dominant factor determining chip performance……..

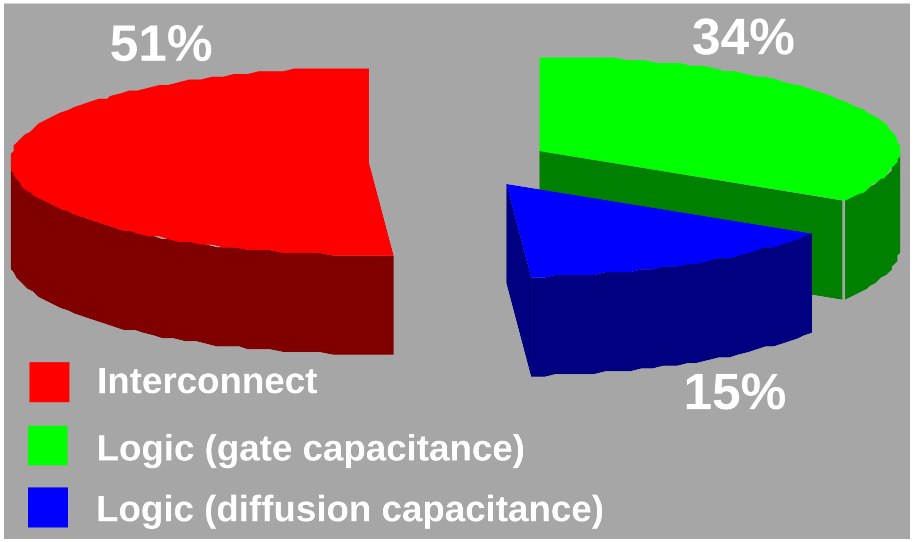
Wire Vs Gate Delay…..



Interconnect delay has become the dominant factor determining chip performance…

K. Banerjee et al., Proc. IEEE, May 2001.

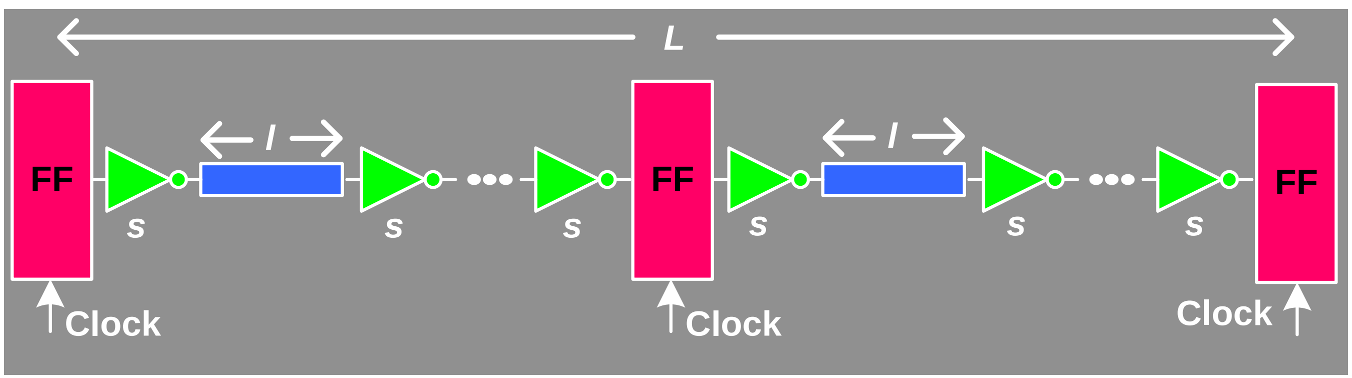
Interconnect Power Dissipation



$$Power = \underbrace{CV_{DD}^{2}f}_{\text{dynamic}} + \underbrace{I_{leak}V_{DD}}_{\text{static}}$$

Interconnect Power Dissipation

*N. Magen et al., (**INTEL**), SLIP, pp. 7-13, 2004.*



*K. Banerjee et al., **IEEE T-ED**, vol. 49, no. 11, pp. 2001-2007, 2002.*

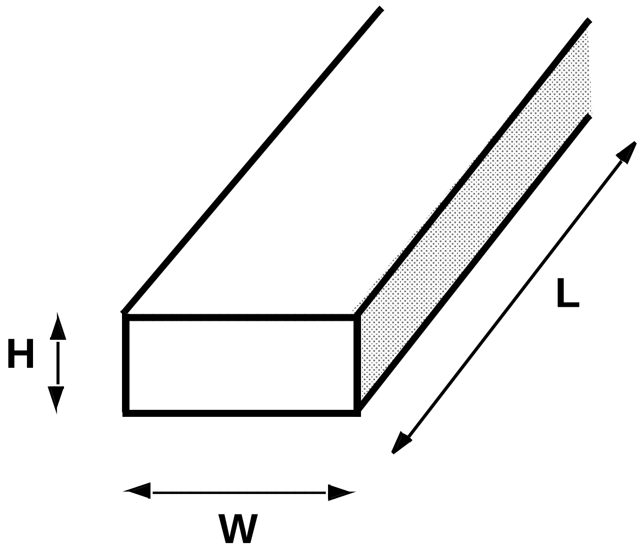
Long (Global) interconnects must be optimally pipelined to meet clock period and power criteria

Impact of Interconnect Parasitics

- Interconnect parasitics
  - affect performance, noise immunity and power consumption
  - affect reliability

- Classes of parasitics
  - Resistive
  - Capacitive
  - Inductive

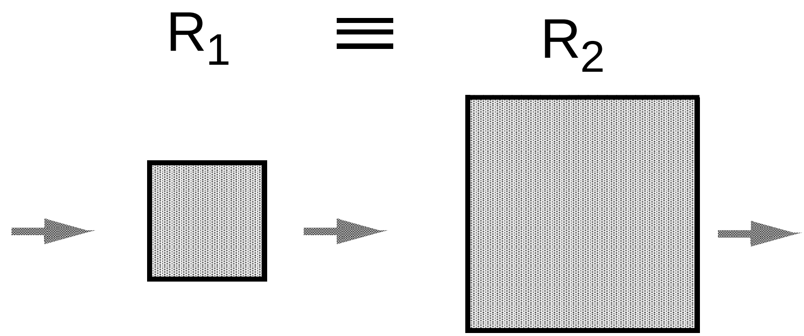
Wire Resistance



$$R = \frac{\rho \, L}{H \, W}$$

Sheet Resistance R_o

Unit of "Ω/sq" or "Ω/□"



Easier to compare wires with different (but uniform) thicknesses…

Interconnect Material Resistivity

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

Sheet Resistance

Material	Sheet Resistance (Ω/❑)
n- or p-well diffusion	1000 – 1500
n^+, p^+ diffusion	50 – 150
n^+, p^+ diffusion with silicide	3 – 5
n^+, p^+ polysilicon	150 – 200
n^+, p^+ polysilicon with silicide	4 – 5
Aluminum	0.05 – 0.1

Dealing with Resistance

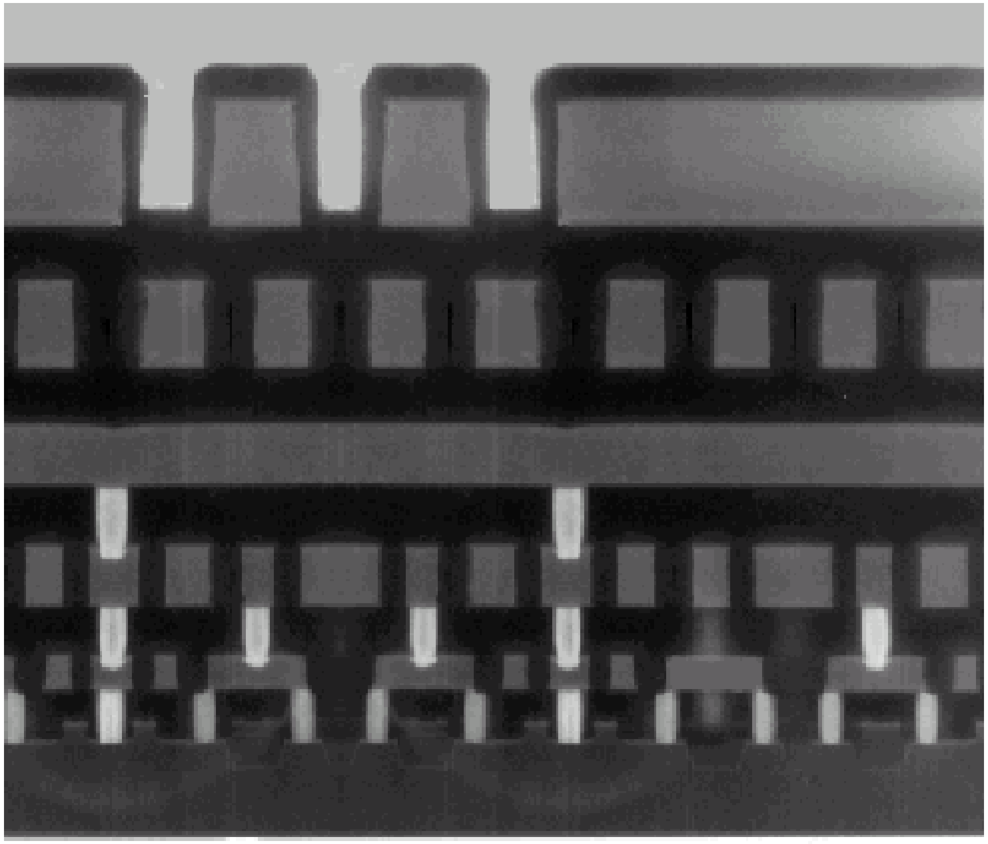
- **Selective Technology Scaling**
- **Use Better Interconnect Materials**
  - reduce average wire-length
  - e.g. copper, silicides
- **More Interconnect Layers**
  - reduce average wire-length

Example: Intel 0.25 micron Process

5 metal layers
Ti/Al - Cu/Ti/TiN
Polysilicon dielectric

LAYER	PITCH	THICK	A.R.
Isolation	0.67	0.40	-
Polysilicon	0.64	0.25	-
Metal 1	0.64	0.48	1.5
Metal 2	0.93	0.90	1.9
Metal 3	0.93	0.90	1.9
Metal 4	1.60	1.33	1.7
Metal 5	2.56	1.90	1.5
	µm	µm	

Layer pitch, thickness and aspect ratio $\boxed{=h/w}$



Interconnect (RC) Delay

On-Chip VLSI interconnects can be modeled as RC elements

- *R is the wire resistance =* $\rho \frac{L}{A} = \rho \frac{L}{w \cdot h}$

    ρ is the resistivity of the metal

    L is the wire length

    A is the cross sectional area = wh (w is the width and h is the height of the wire)

- *C is the wire capacitance. For a parallel plate capacitor*

$$C = \varepsilon_d \frac{A}{d}, \ \varepsilon_d = k\,\varepsilon_0$$

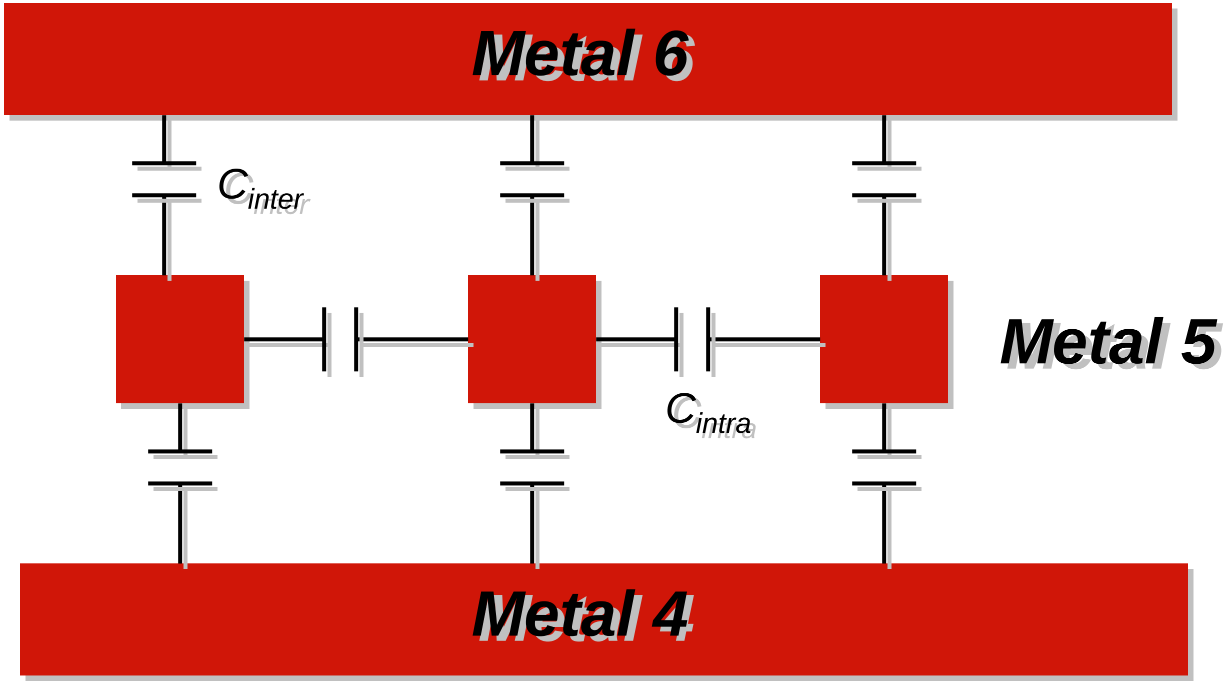
A is the area of the plate

d is the distance between the plates

k is the dielectric constant of the insulating material between the plates

ε_0 is the permittivity of free space

VLSI Interconnect Structure



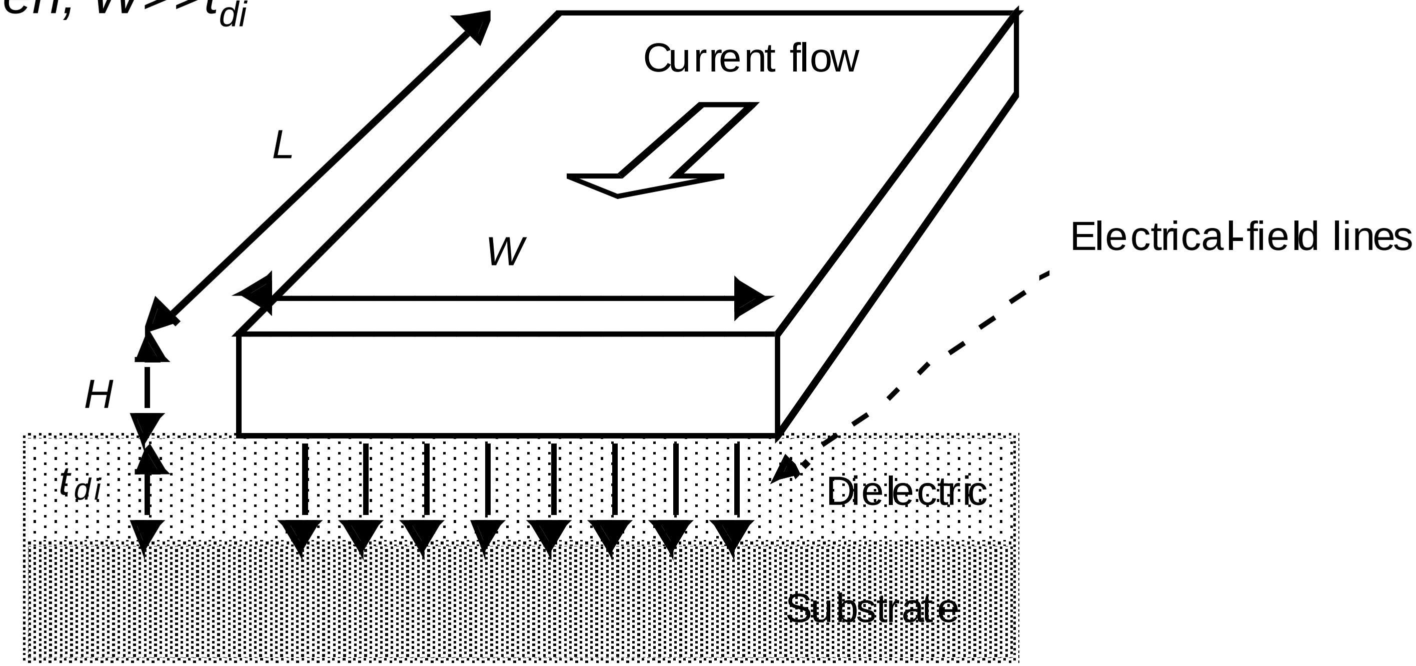
Adjacent wires are orthogonal….why?

C_{intra} is the dominating capacitance……why?

Capacitance: The Parallel Plate Model

1-dimensional model

Valid when, $W >> t_{di}$

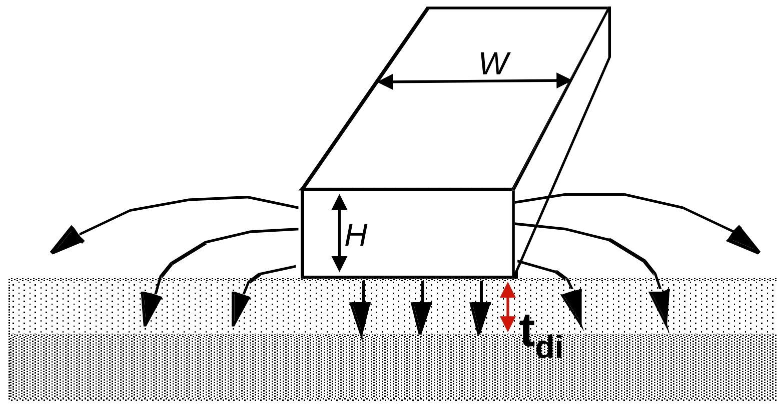


$$c_{int} = \frac{\varepsilon_{di}}{t_{di}} WL$$

Permittivity

Material	ε_r
Free space	1
Aerogels	~1.5
Polyimides (organic)	3-4
Silicon dioxide	3.9
Glass-epoxy (PC board)	5
Silicon Nitride (Si_3N_4)	7.5
Alumina (package)	9.5
Silicon	11.7

Fringing Capacitance Model



Approximates the capacitance as the sum of two components:

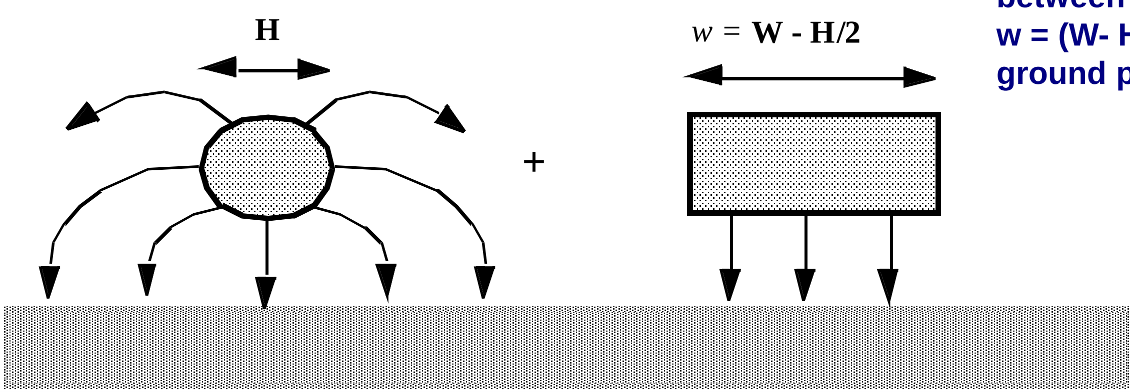
$$c_{wire} = c_{pp} + c_{fringe} = \frac{w\varepsilon_{di}}{t_{di}} + \frac{2\pi\varepsilon_{di}}{\operatorname{arccosh}\left(\frac{2t_{di}}{H}+1\right)}$$

Orthogonal field between a wire of width w = (W- H/2) and the ground plane

Modeled by a cylindrical wire of diameter= H

$$\operatorname{arccosh}\left(\frac{2t_{di}}{H}+1\right) \approx \ln\left(\frac{4t_{di}}{H}+2\right)$$

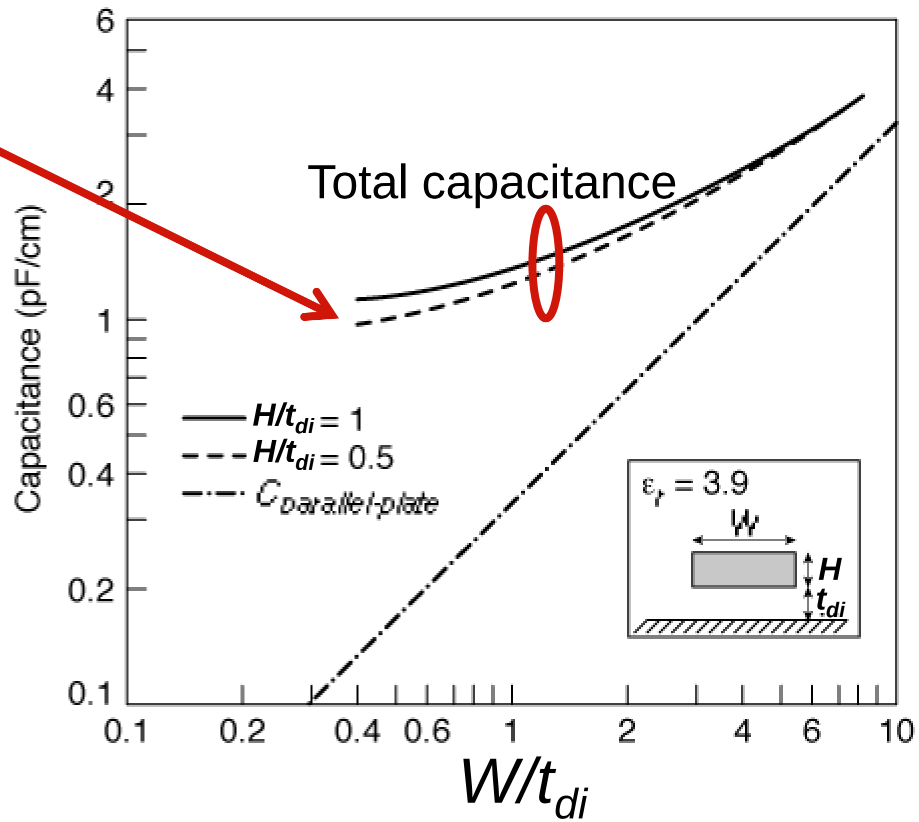
$$\text{as } t_{di} > 0.5H$$



Fringing Cap. **In parallel with..** *Parallel plate Cap.*

Fringing versus Parallel Plate

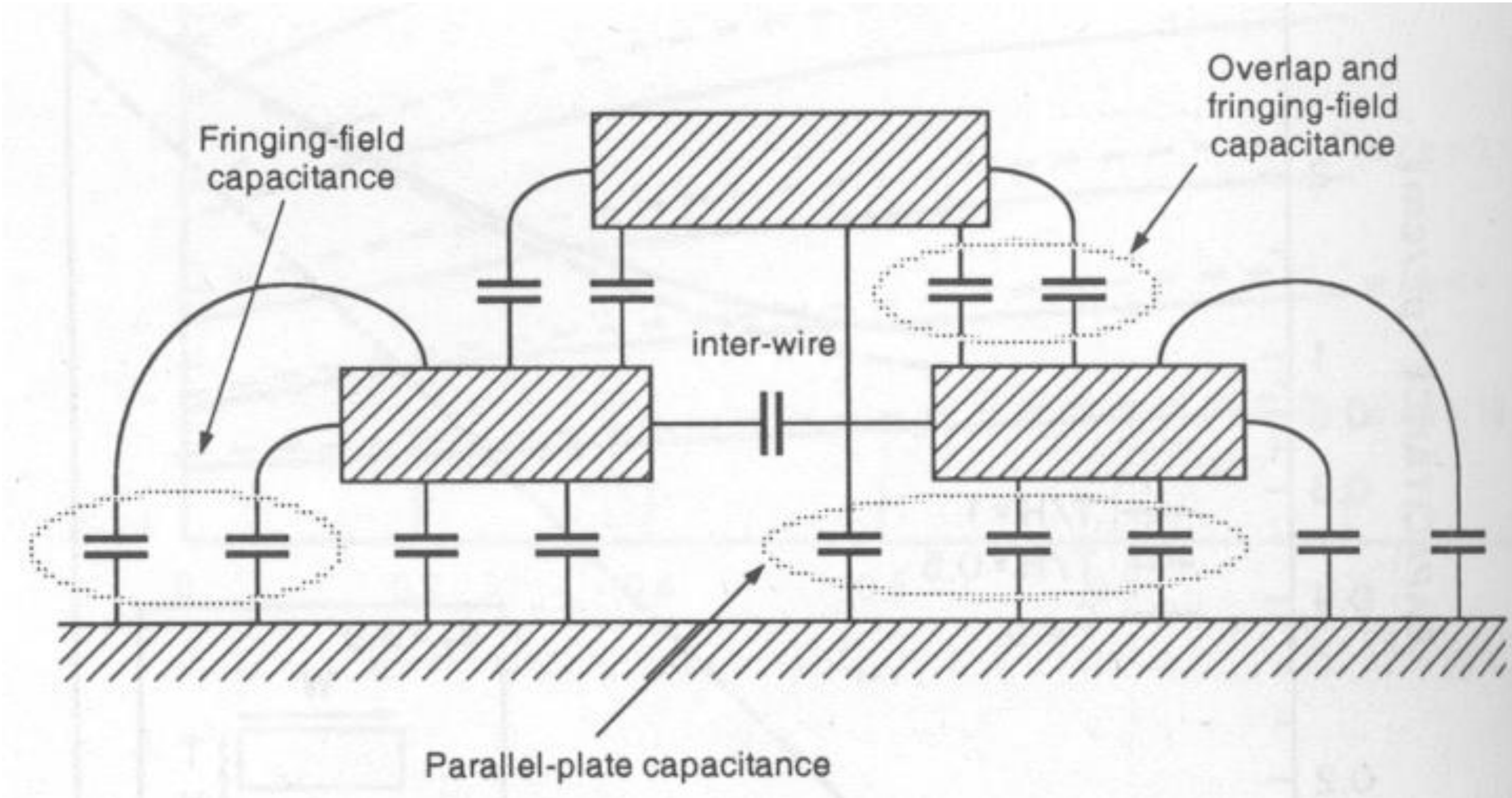
Saturates to ~1pF/cm due to fringing fields



Note: aspect ratio = H/W

(from [Bakoglu89])

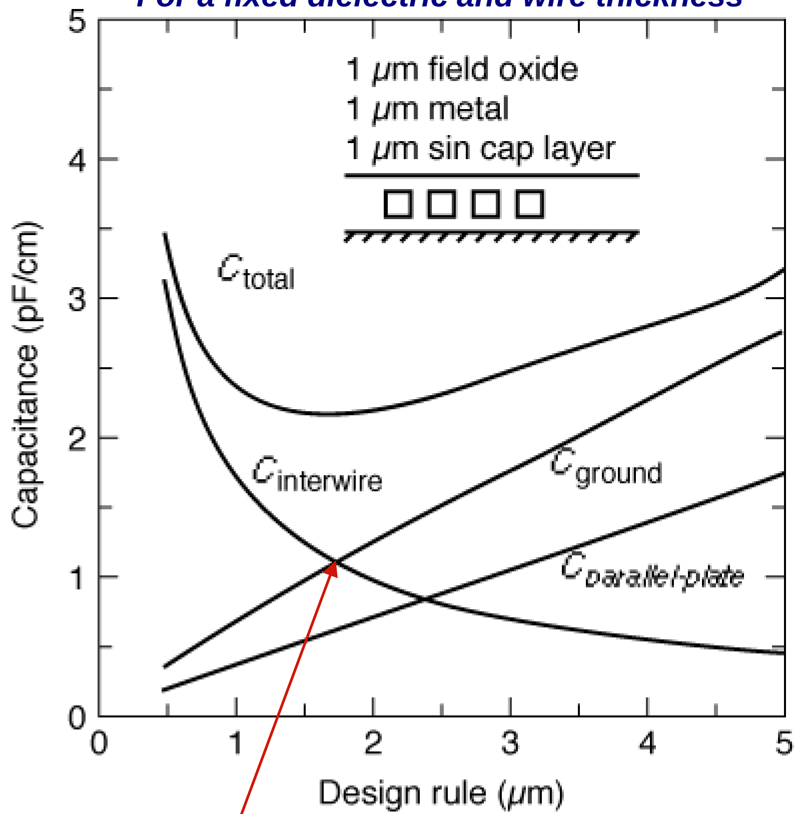
Interconnect



- Usually interconnect layers alternate wire direction

Impact of Interwire Capacitance

For a fixed dielectric and wire thickness



W/H =1.75

- *Reduce wire width, w*
- *Reduce spacing (=w)*
- $C_{total} = C_{ground} + C_{interwire}$
- $C_{ground} = C_{p\text{-}p} + C_{fringing}$

Interwire cap. (crosstalk) becomes important in a multi-level structure

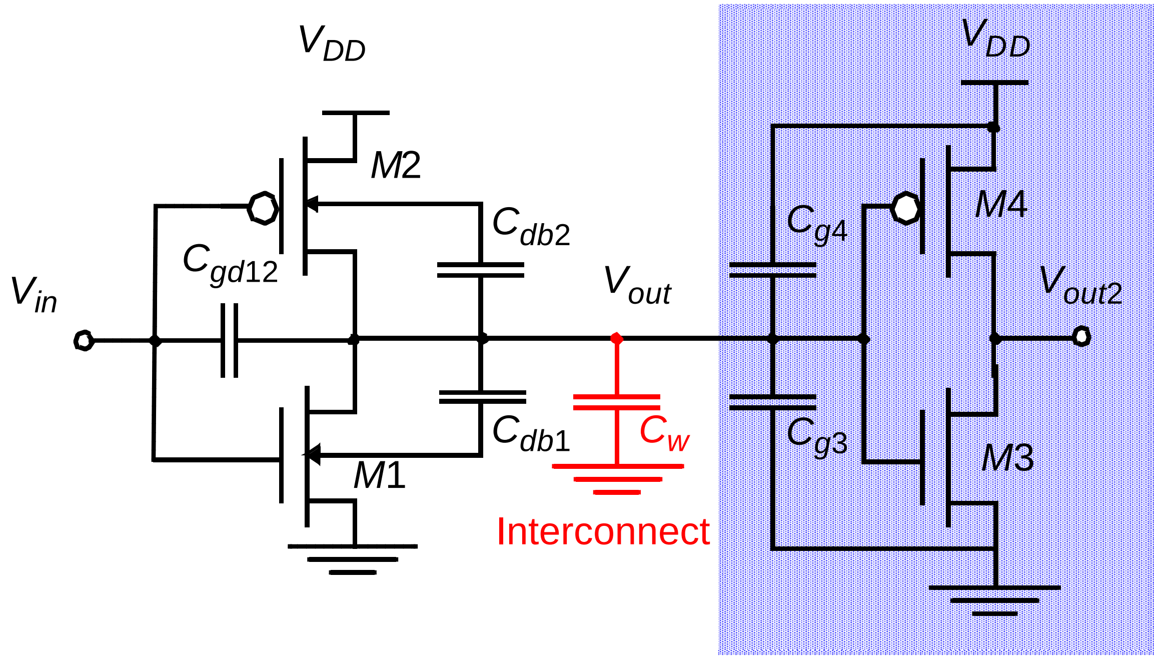
More imp. For higher level wires that are far away from substrates

(from [Bakoglu89])

Accurate Interconnect Capacitance Extraction

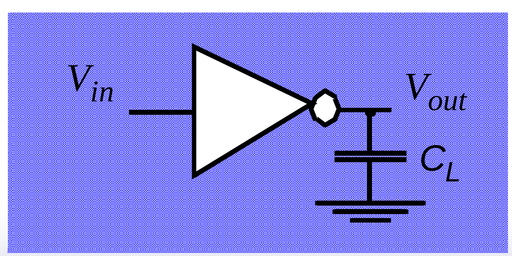
- Requires full 3-D electrostatic simulation
- FASTCAP (from MIT) is a widely used tool in academia
  - See K. Nabors and J. White, "FastCap: a multipole accelerated 3-D capacitance extraction program," IEEE Trans. Computer-Aided Des. Integr. Circuits Syst., vol. 10, no. 11, pp. 1447–1459, Nov. 1991.
- Can be time consuming for complex interconnect topologies

Capacitances of Driver-Wire-Load

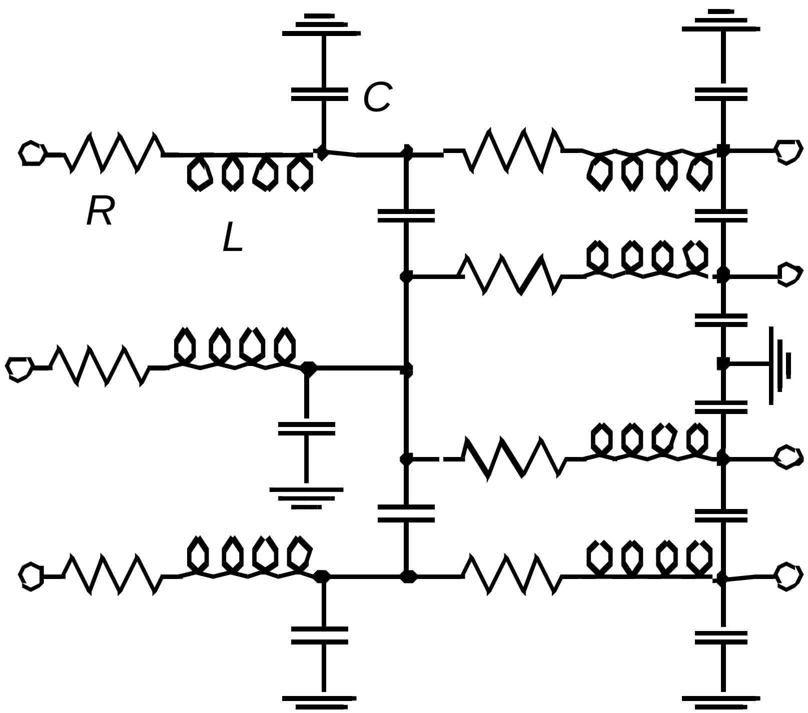


Fanout

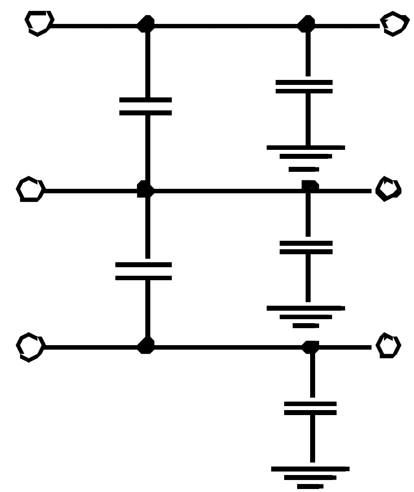
Simplified Model



Wire Models



All-inclusive model

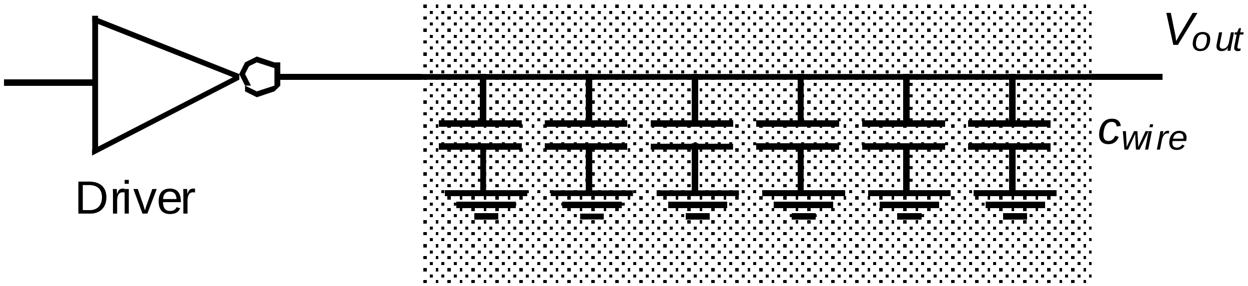


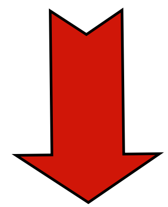
Capacitance-only

Kaustav Banerjee

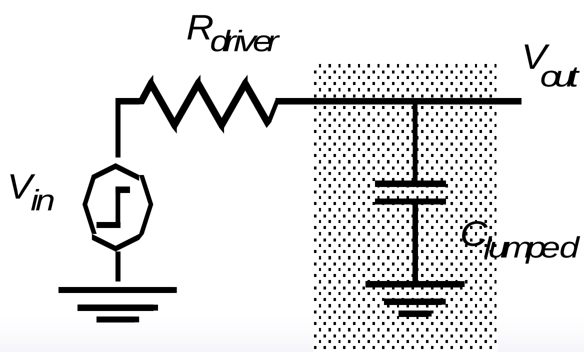
The Lumped Capacitor Model

Assuming no voltage drops along the wire….
i.e., wire is equipotential





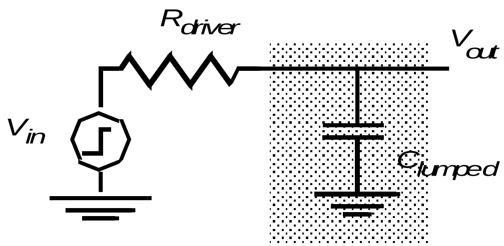
If r is small and switching frequencies are not too high…



***Advantage:* Operation can be described by ordinary differential equation**

The Lumped Capacitor Model

Driver modeled as a voltage source and a source resistance (R_{driver}), $C_{lumped} = L \times c_{wire}$



Only impact on performance is due to the loading effect of the capacitor on the driving gate

Operation of a simple RC circuit:

$$C_{lumped}\frac{dV_{out}}{dt} + \frac{V_{out} - V_{in}}{R_{driver}} = 0$$

If a step input is applied from 0 to V:

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

where $\tau = R_{driver} C_{lumped}$ *: time-constant of network*

The time to reach 50% point can be calculated as:

$$t = \ln(2)\tau = 0.69\tau$$

If $R_{driver} = 10K\Omega$ *and* $C_{lumped} = 11\,pF$:

$t = 76\,ns$ (*a very large number for high-performance digital circuits*)

The Lumped RC-Model

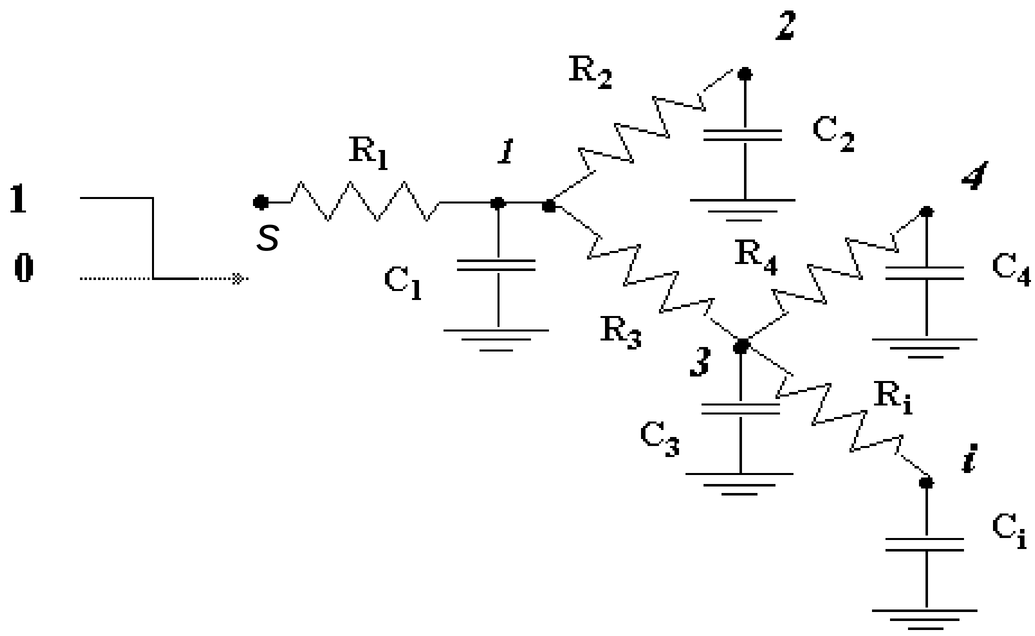
- On-chip metal wires at the semi-global and global tiers can be several millimeters long and can have significant resistance
- The equipotential assumption is no longer valid
- Need an RC model……
- A simple approach: lump wire resistance of each segment into a single R and the global capacitance into a single C
  - This is called a lumped RC model….pessimistic and inaccurate for long wires
  - Should use a distributed RC model for such long wires

Then what is the use of the lumped RC model?

Lumped RC-Networks…..

- The distributed RC model is complex: involves partial differential equations, closed form solution does not exist
- However, the behavior of a distributed RC line can be adequately modeled by a simple RC network
- It is more convenient to reduce any driver-wire-load networks to an equivalent RC network and predict its first order response
- ***However, unlike the single R-C network analyzed earlier that had a single time constant (network pole), deriving the correct waveforms for a complex network becomes intractable***
  - Need to solve a set of ordinary differential equations with many time constants (poles and zeroes)….
  - Or, run time-consuming SPICE simulations for the entire network….

Elmore Delay Comes to the Rescue…..



RC Tree: *unique resistive paths between s and any node i*

*Total resistance along this path is called the **path resistance*** R_{ii}

e.g., path resistance between s and node 4: $R_{44} = R_1 + R_3 + R_4$

Shared path resistance: *resistance shared among the paths from root node s to nodes k and i*

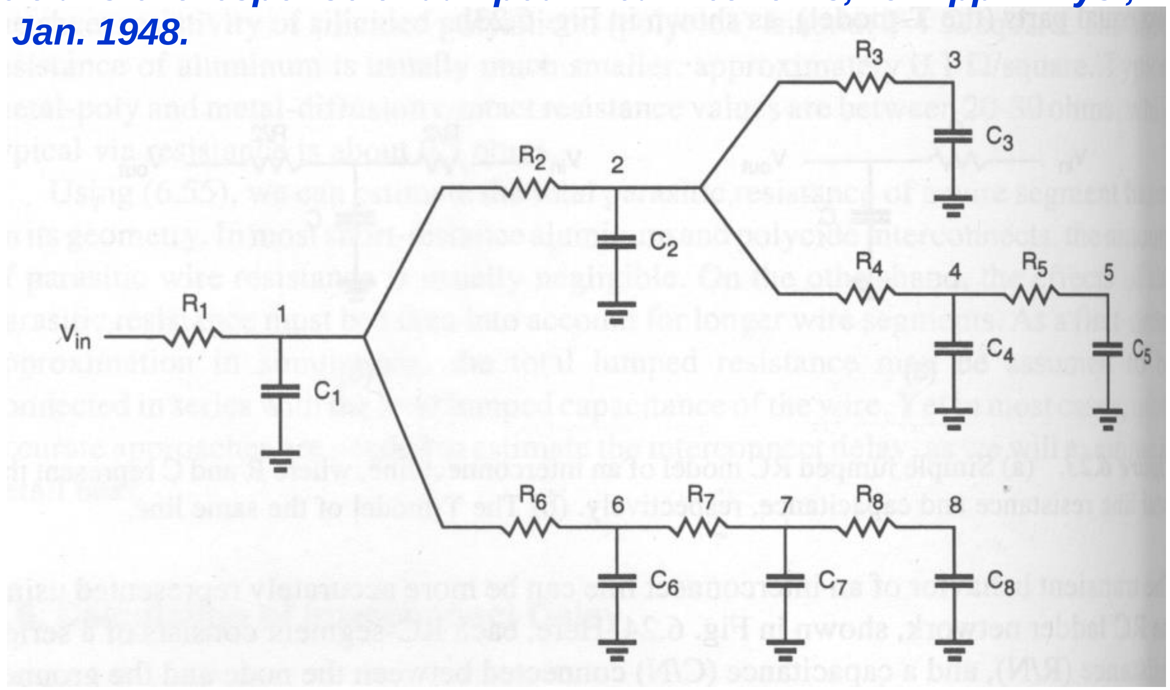
$$R_{ik} = \sum R_j \Rightarrow (R_j \in [path(s \rightarrow i) \cap path(s \rightarrow k)])$$

Example: $R_{i4}=R_1+R_3$

$R_{i2}=R_1$

Elmore Delay-RC Chain (Branched)

W. C. Elmore, "The transient response of damped linear networks," J. Appl. Phys., vol. 19, pp. 55–63, Jan. 1948.



If a step input Vin is applied at t=0

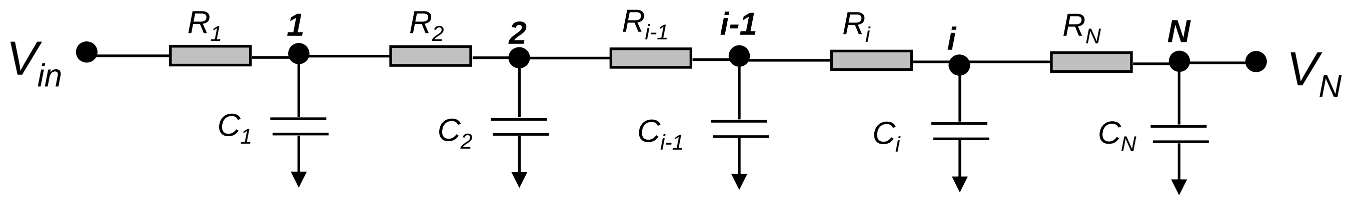
Elmore Delay at node i:

$$\tau_{Di} = \sum_{k=1}^{N} C_k R_{ik}$$

(First-order time constant of the network)

N = # of capacitances in the network

Wire Model



For a chain, number of capacitances = number of nodes, hence,

For a non-branched RC chain: $\tau_{DN} = \sum_{i=1}^{N} C_i R_{ii}$ $\quad t_{Di} = C_1R_1 + C_2 (R_1+R_2)+\ldots\ldots + C_i (R_1+R_2+\ldots R_i)$

If wire modeled by N equal-length segments ($R_{seg}= rL/N$, $C_{seg}= cL/N$)

$$\tau_{DN} = \left(\frac{L}{N}\right)^2 (rc + 2rc + \ldots + Nrc) = (rcL^2)\frac{N(N+1)}{2N^2} = RC\frac{N+1}{2N}$$

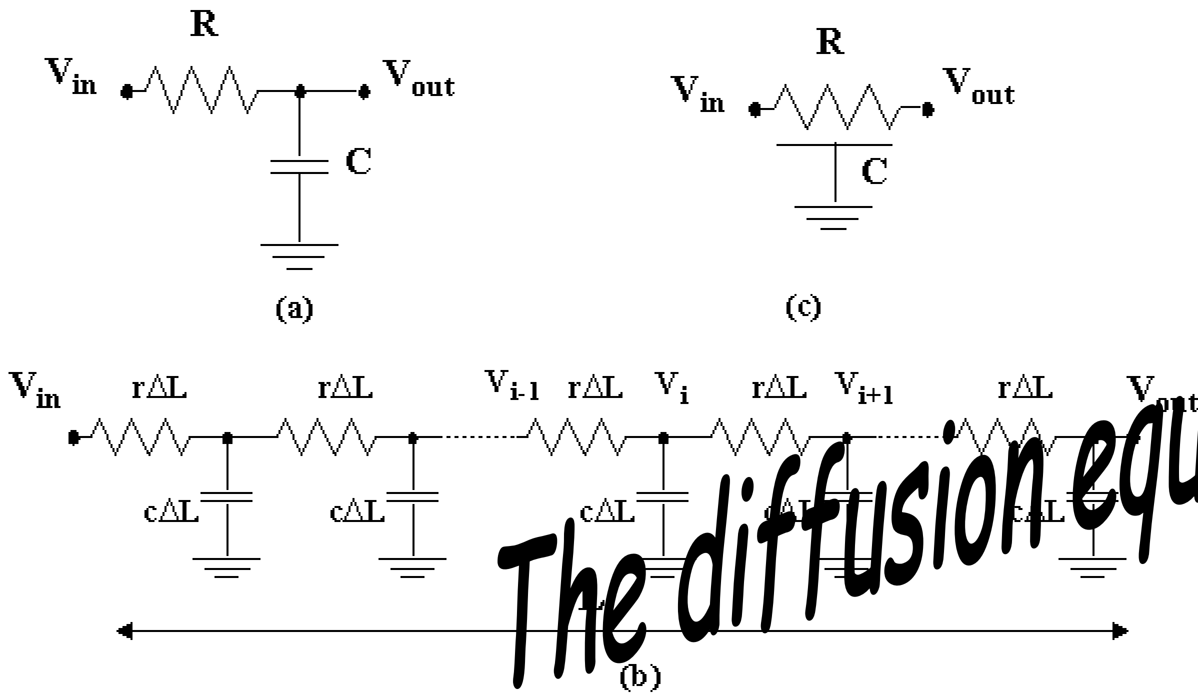
r and c represent the resistance and capacitance per unit length of the wire, respectively…

For large values of N:

$$\tau_{DN} = \frac{RC}{2} = \frac{rcL^2}{2}$$

The Distributed RC-line

(can be approximated by the lumped RC network model)

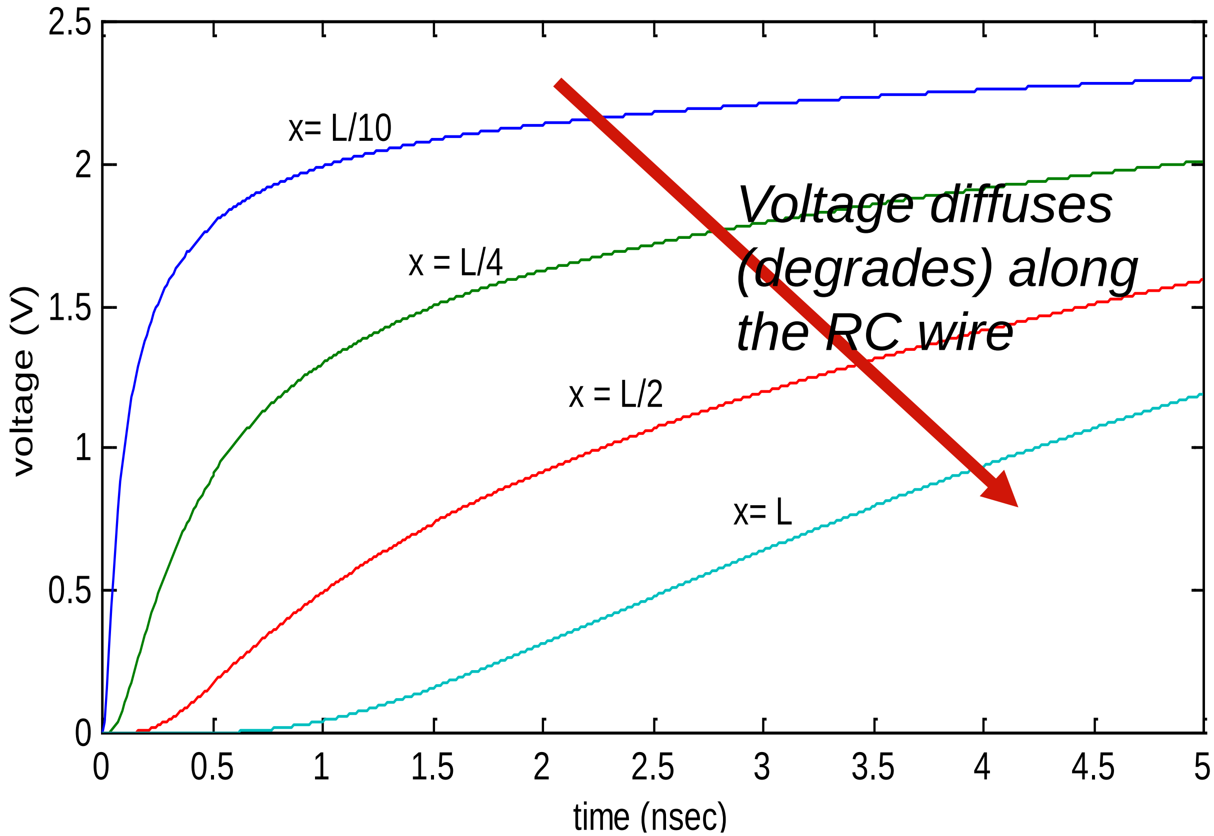


The diffusion equation

$$rc\frac{\partial V}{\partial t} = \frac{\partial^2 V}{\partial x^2}$$

$$\tau(V_{out}) = \frac{rc\ L^2}{2}$$

Step-response of RC wire as a function of time and space



Kaustav Banerjee

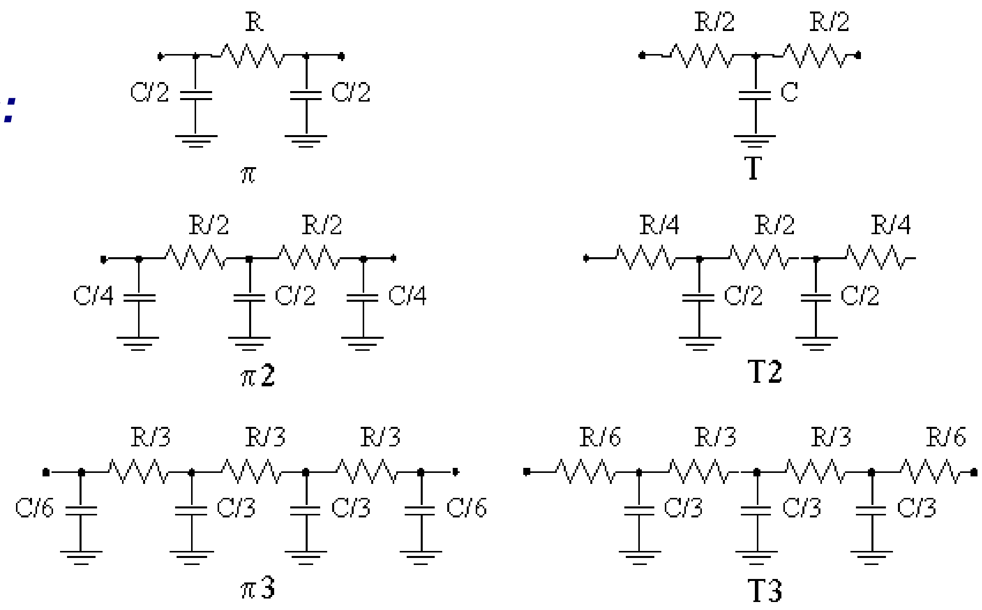
RC-Models

Voltage Range	Lumped RC-network	Distributed RC-network
0→50% (t_p)	0.69 RC	0.38 RC
0→63% (τ)	RC	0.5 RC
10%→90% (t_r)	2.2 RC	0.9 RC

Step Response of Lumped and Distributed RC Networks: Points of Interest.

SPICE Wire (RC) Models:

accuracy of the model determined by number of stages



RC Delay: Design Rules of Thumb

- rc delays should only be considered when $t_{pRC} >> t_{pgate}$ of the driving gate

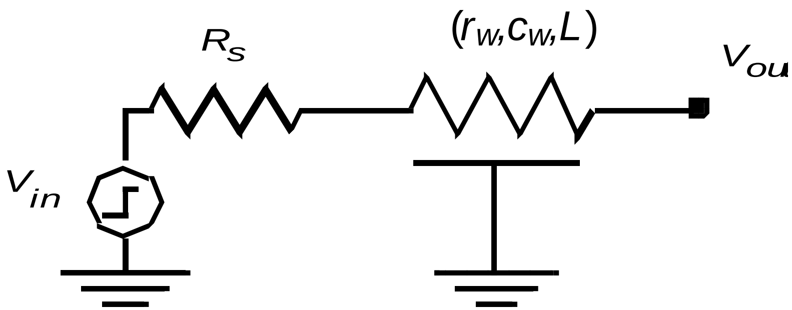
$$L_{crit} >> \sqrt{t_{pgate}/0.38rc}$$

- rc delays should only be considered when the rise (fall) time at the line input is smaller than RC, the rise (fall) time of the line

$$t_{rise} < RC$$

  - when not met, the change in the signal is slower than the propagation delay of the wire

Illustration of Rule: Driving an RC-line



$R_w = r_w L$

$C_w = c_w L$

Total Propagation Delay: $\tau_D = R_s C_w + \frac{R_w C_w}{2} = R_s C_w + 0.5 r_w c_w L^2$

Delay due to wire resistance becomes important when $(R_w C_w/2) \geq R_s C_w$ or when $L \geq 2R_s/r_w$

0 to 50% Response Delay: $t_p = 0.69 R_s C_w + 0.38 R_w C_w$