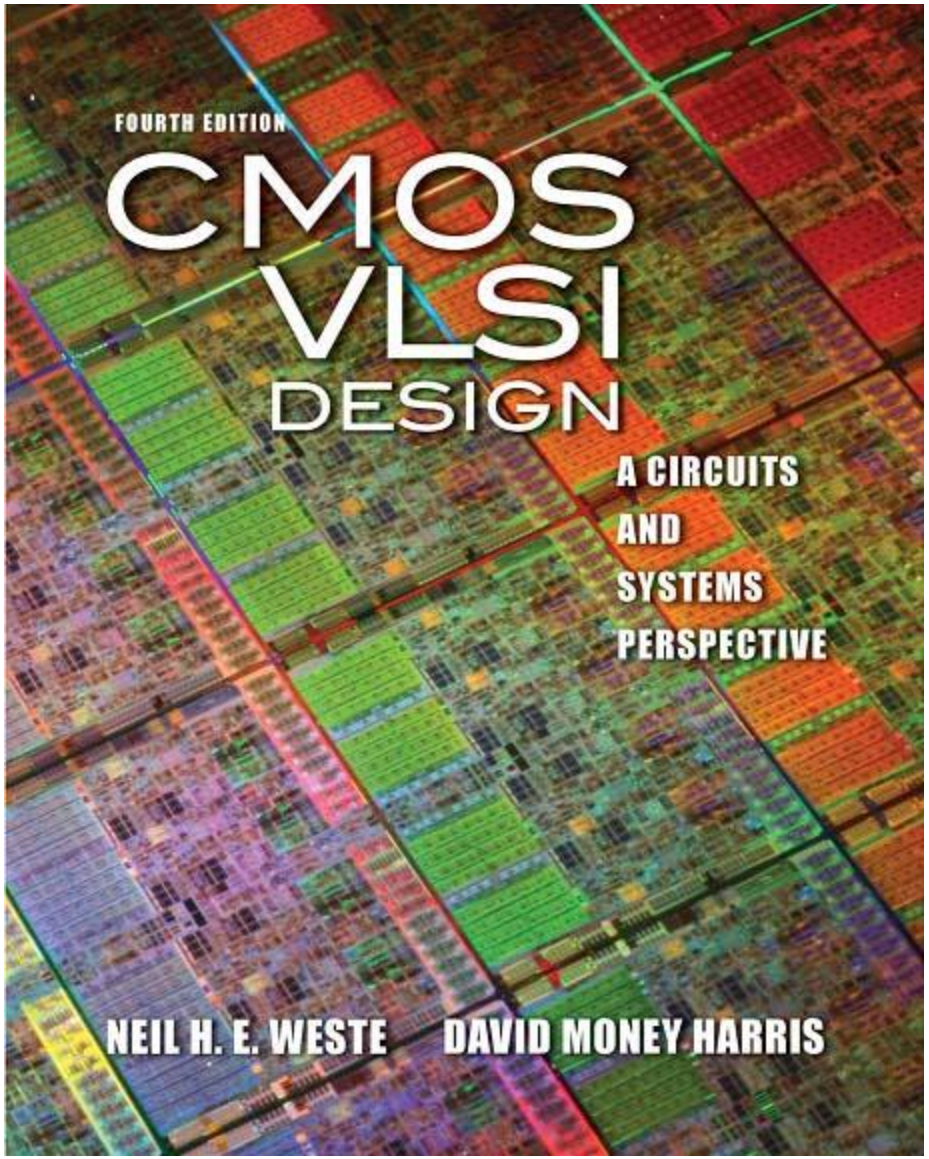


Wires

Dr. Bassam Jamil

Adopted from slides of the textbook

Outline

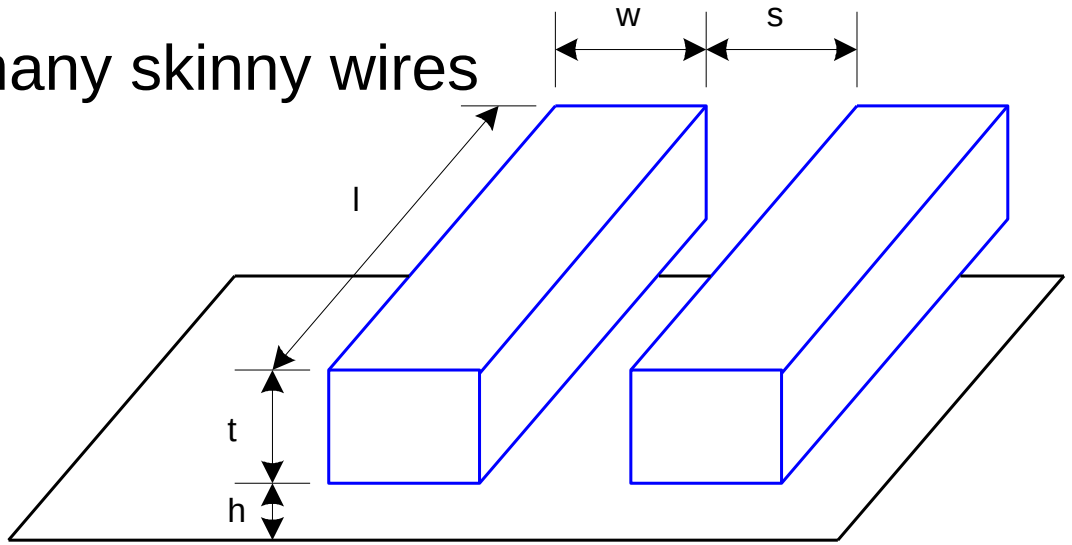
- Introduction
- Interconnect Modeling
  - Wire Resistance
  - Wire Capacitance
- Interconnect Impact
  - Wire RC Delay
  - Crosstalk
  - Wire Engineering
- Repeaters

Introduction

- Chips are mostly made of wires called *interconnect*
  - In stick diagram, wires set size
  - Transistors are little things under the wires
  - Many layers of wires
- Wires are as important as transistors
  - Speed
  - Power
  - Noise
- Alternating layers run orthogonally

Wire Geometry

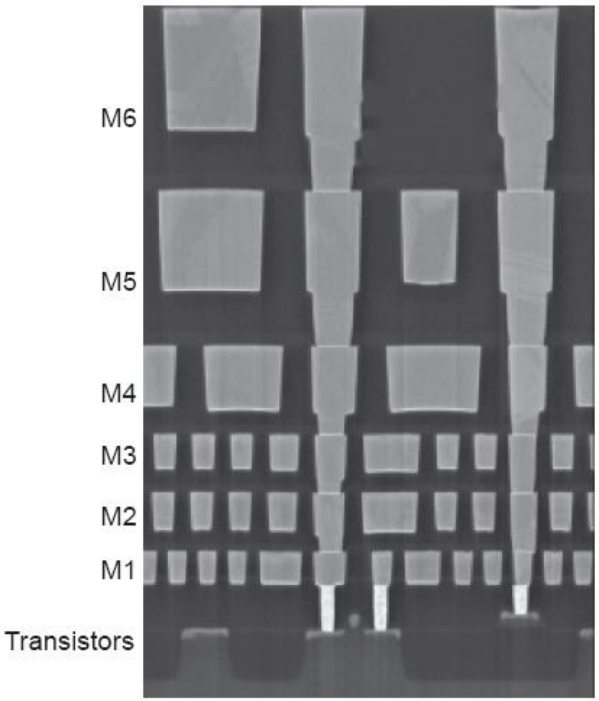
- Pitch = w + s
- Aspect ratio: AR = t/w
  - Old processes had AR << 1
  - Modern processes have AR ≈ 2
    - Pack in many skinny wires



Layer Stack

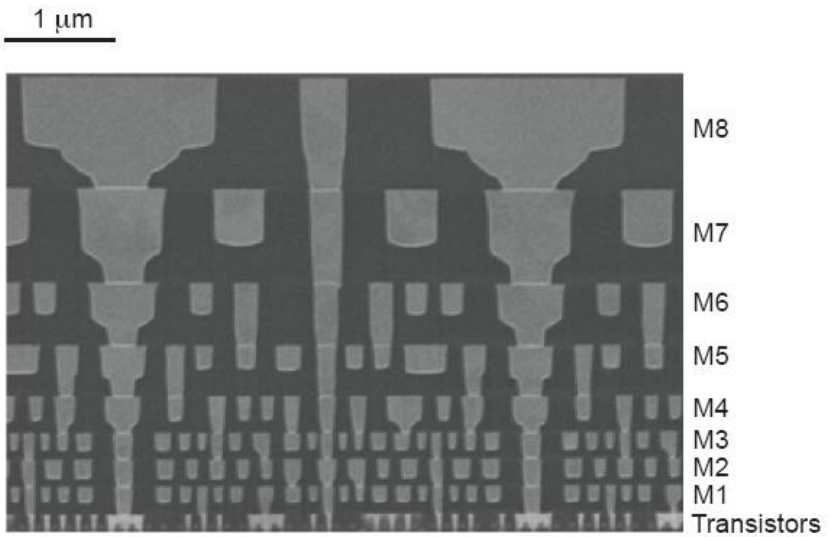
- Example of old process: AMI 0.6 μm process has 3 metal layers
  - M1 for within-cell routing
  - M2 for vertical routing between cells
  - M3 for horizontal routing between cells
- Modern processes use 6-10+ metal layers
  - M1: thin, narrow (< 3λ)
    - High density cells
  - Mid layers
    - Thicker and wider, (density vs. speed)
  - Top layers: thickest
    - For V_{DD}, GND, clk

Example



Intel 90 nm Stack

[Thompson02]

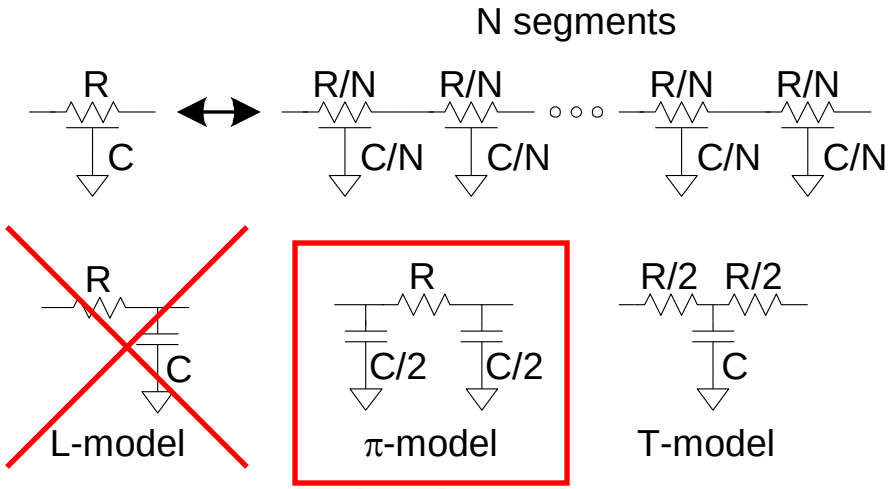


Intel 45 nm Stack

[Moon08]

Lumped Element Models

- Wires are a distributed system
  - Approximate with lumped element models



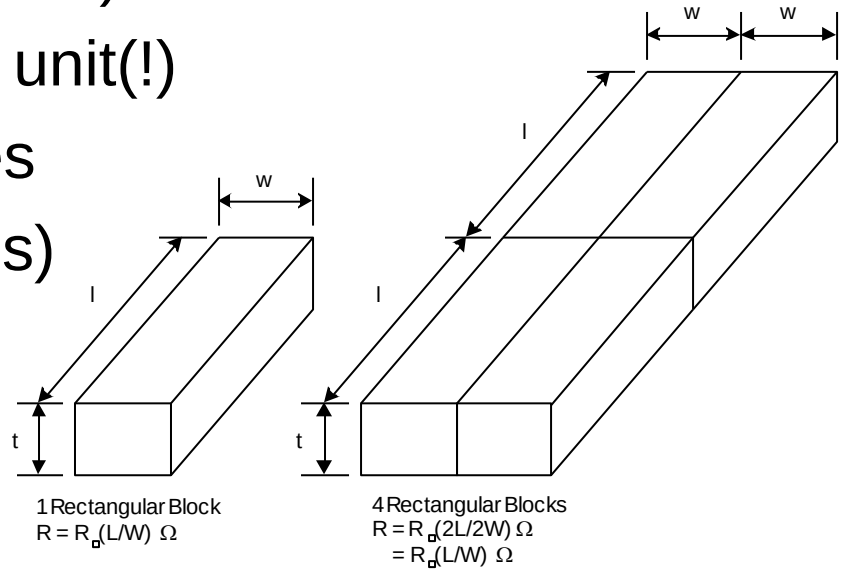
- 3-segment π-model is accurate to 3% in simulation
- L-model needs 100 segments for same accuracy!
- Use single segment π-model for Elmore delay

Wire Resistance

- ρ = *resistivity* (Ω*m)

$$R = \frac{\rho}{t}\frac{l}{w} = R_{\square}\frac{l}{w}$$

- $R_{\square}$ = *sheet resistance* ($\Omega/\square$)
  - $\square$ is a dimensionless unit(!)
- Count number of squares
  - $R = R_{\square} * (\#\ \text{of squares})$



Choice of Metals

- Until 180 nm generation, most wires were **aluminum**
- Contemporary processes normally use **copper**
  - Cu atoms diffuse into silicon and damage FETs
  - Must be surrounded by a diffusion barrier

Metal	**Bulk resistivity ($\mu\Omega \cdot$ cm)**
Silver (Ag)	1.6
Copper (Cu)	1.7
Gold (Au)	2.2
Aluminum (Al)	2.8
Tungsten (W)	5.3
Titanium (Ti)	43.0

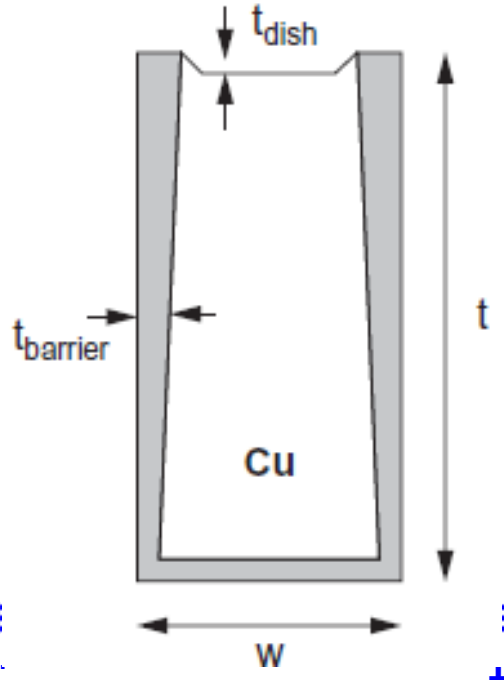
Contacts Resistance

- Contacts and vias also have 2-20 Ω
- Use many contacts for lower R
  - Many small contacts for current crowding around periphery

Copper Issues

- Copper must be surrounded by a lower-conductivity diffusion barrier which effectively reduces the wire cross-sectional area and increases resistance.
- Copper is also prone to *dishing* during polishing
- Effective resistance is higher

$$R = \frac{\rho}{(t - t_{dish} - t_{barrier})} \frac{l}{(w - 2t_{barrier})}$$



Cu Wire Example

- Compute the sheet resistance of a 0.22 μm thick Cu wire in a 65 nm process. The resistivity of thin film Cu is 2.2 x 10-8 Ω•m. Ignore dishing.

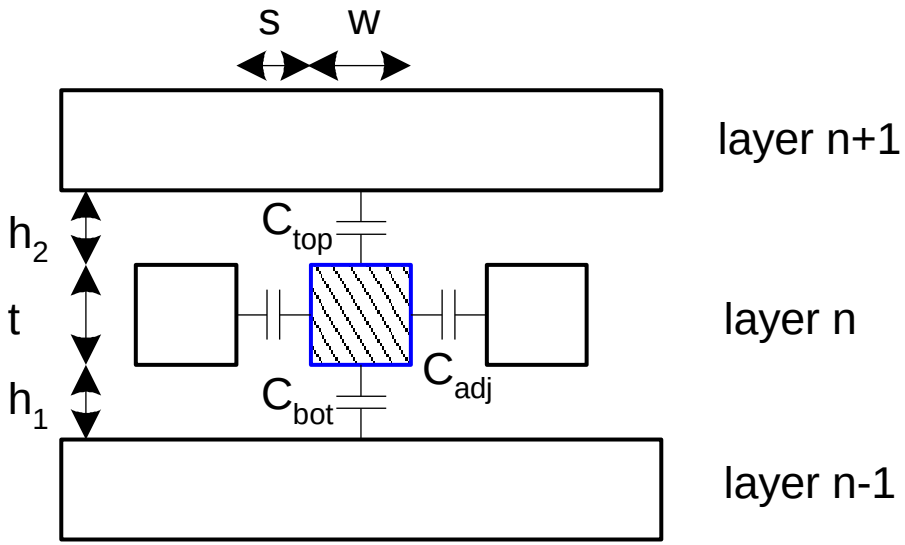
$$R_{\square} = \frac{2.2\times10^{-8}\ \Omega\cdot\text{m}}{0.22\times10^{-6}\ \text{m}} = 0.10\ \Omega/\square$$

- Find the total resistance if the wire is 0.125 μm wide and 1 mm long. Ignore the barrier layer.

$$R = \left(0.10\ \Omega/\square\right)\frac{1000\ \mu\text{m}}{0.125\ \mu\text{m}} = 800\ \Omega$$

Wire Capacitance

- Wire has capacitance per unit length
  - To neighbors
  - To layers above and below
- $C_{total} = C_{top} + C_{bot} + 2C_{adj}$



Capacitance Trends

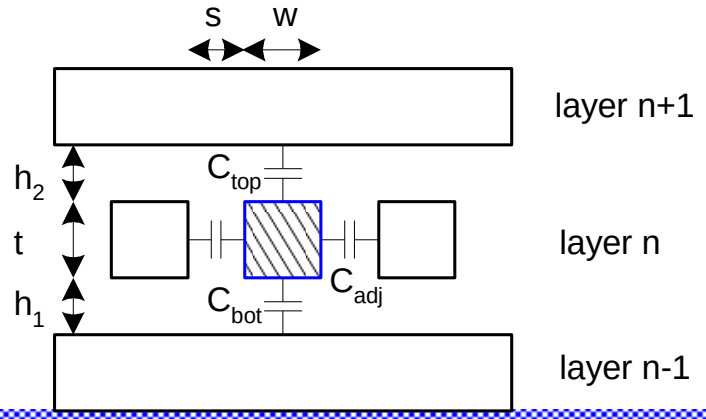
- Parallel plate equation: $C = \varepsilon_{ox} A/d$
  - Wires are not parallel plates, but obey trends
  - Increasing area (W, t) increases capacitance
  - Increasing distance (s, h) decreases capacitance
- Dielectric constant
  - $\varepsilon_{ox} = k\varepsilon_0$
    - $\varepsilon_0 = 8.85 \times 10^{-14}$ F/cm
    - $k = 3.9$ for SiO_2
- Processes are starting to use low-k dielectrics
  - $k \approx 3$ (or less) as dielectrics use air pockets

Capacitance Formula

- Capacitance of a line without neighbors can be approximated as

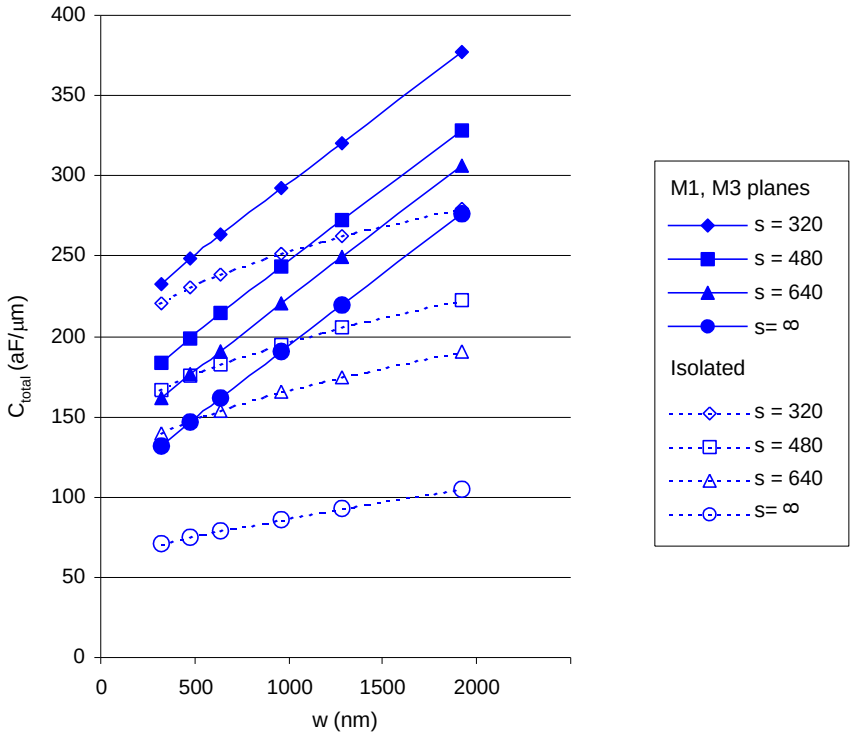
$$C_{tot} = \varepsilon_{ox} l \left[\frac{w}{h} + 0.77 + 1.06 \left(\frac{w}{h} \right)^{0.25} + 1.06 \left(\frac{t}{h} \right)^{0.5} \right]$$

- This empirical formula is accurate to 6% for AR < 3.3



M2 Capacitance Data

- Typical dense wires have ~ 0.2 fF/μm
  - Compare to 1-2 fF/μm for gate capacitance



Diffusion & Polysilicon

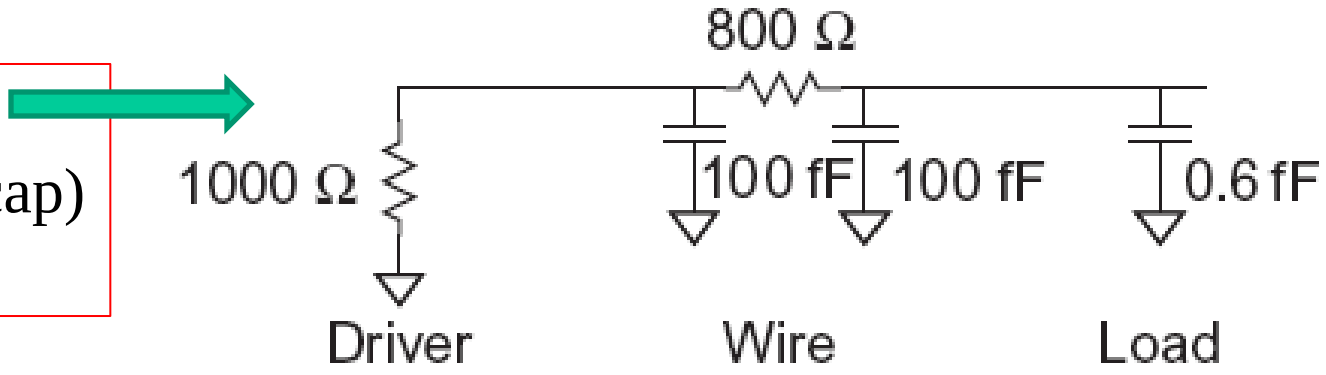
- Diffusion capacitance is very high (1-2 fF/μm)
  - Comparable to gate capacitance
  - Diffusion also has high resistance
  - Avoid using diffusion *runners* for wires!
- Polysilicon has lower C but high R
  - Use for transistor gates
  - Occasionally for very short wires between gates

Wire RC Delay Example

- Estimate the delay of a 10x inverter driving a 2x inverter at the end of the 1 mm wire. Assume :
  - Wire capacitance is 0.2 fF/μm and resistance of the Cu Wire example (in earlier slide)
  - A unit-sized inverter has R = 10 KΩ and C = 0.1 fF.

t_{pd} = (1000 Ω)(100 fF) + (1000 + 800 Ω)(100 + 0.6 fF) = 281 ps

Intrinsic delay (due to diffusion cap) is ignored.



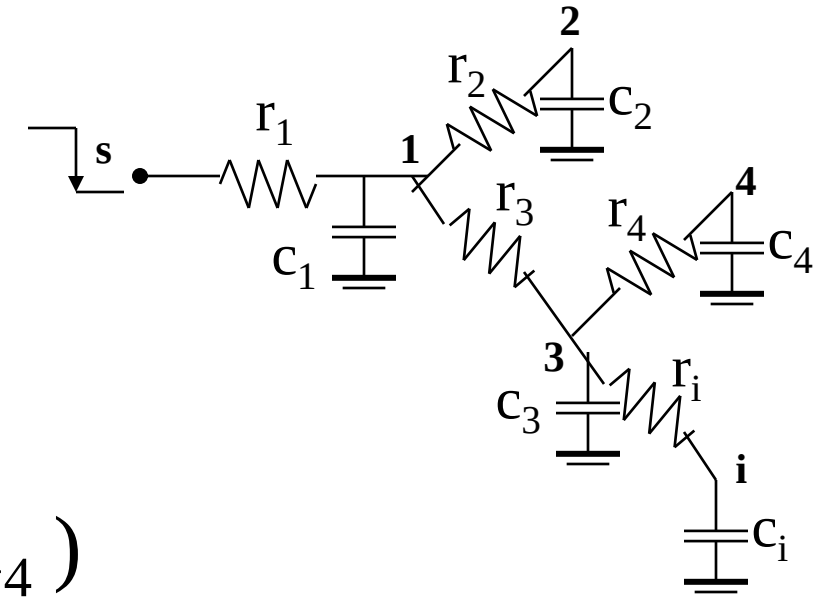
Wire Delay Example (2)

Find the RC flight time per mm^2 for a wire using the parameters from Example 6.3. Express the result in FO4/mm^2, if the FO4 inverter delay is 15 ps. What is the flight time to cross a 10 mm die?

SOLUTION: $R = 800\ \Omega/mm$. $C = 0.2$ pF/mm. The flight time is $RC/2 = 80$ ps/mm^2, or 5.3 FO4/mm^2. The flight time across a 10 mm die is thus 530 FO4, which is dozens of clock cycles.

Lumped RC Network/Tree

- $R_{44} = R_1 + R_3 + R_4$
- $R_{i4} = R_1 + R_3$
- $R_{i2} = R_1$
- $\tau_{Di} = r_1 (c_1 + c_2) + (r_1 + r_3)(c_3 + c_4) + (r_1 + r_3 + r_i) c_i$



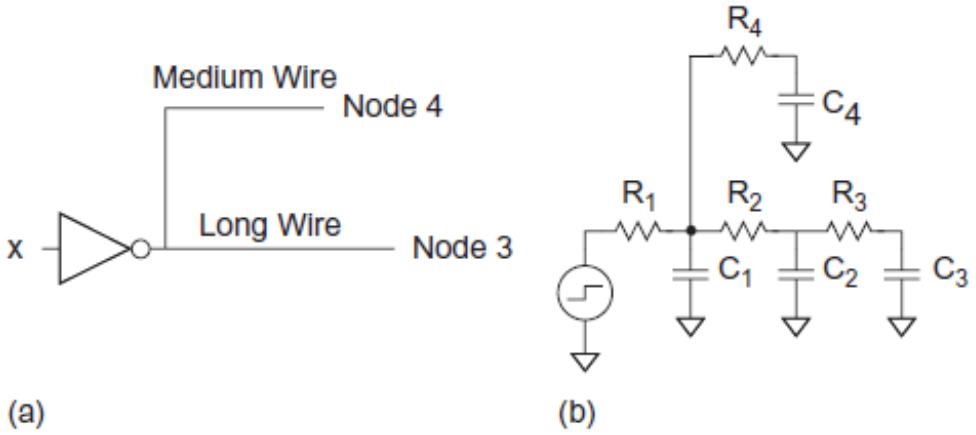
Wire Delay Example (3)

Example 6.5

Figure 6.15 models a gate driving wires to two destinations. The gate is represented as a voltage source with effective resistance R_1. The two receivers are located at nodes 3 and 4. The wire to node 3 is long enough that it is represented with a pair of π-segments, while the wire to node 4 is represented with a single segment. Find the Elmore delay from input x to each receiver.

SOLUTION: The Elmore delays are

$$T_{D_3} = R_1C_1 + (R_1 + R_2)C_2 + (R_1 + R_2 + R_3)C_3 + R_1C_4$$
$$T_{D_4} = R_1C_1 + R_1(C_2 + C_3) + (R_1 + R_4)C_4 \qquad (6.18)$$



Wire Energy

- Estimate the energy per unit length to send a bit of information (one rising and one falling transition) in a CMOS process.

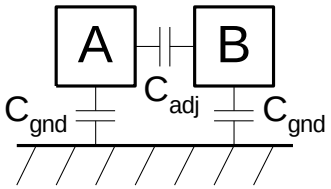
- $E = E_{C(rising)} + E_{C(falling)} +$
  $= \frac{1}{2} C V^2 + \frac{1}{2} C V^2$
- $E = (0.2 \text{ pF/mm})(1.0 \text{ V})^2$ $= 0.2$ pJ/bit/mm
  $= 0.2$ mW/Gbps/mm

Crosstalk

- A capacitor does not like to change its voltage instantaneously.
- A wire has high capacitance to its neighbor.
  - When the neighbor switches from 1-> 0 or 0->1, the wire tends to switch too.
  - Called capacitive *coupling* or *crosstalk*.
- Crosstalk effects
  - Noise on nonswitching wires
  - Increased delay on switching wires

Crosstalk Delay

- Assume layers above and below on average are quiet
  - Second terminal of capacitor can be ignored
  - Model as $C_{gnd} = C_{top} + C_{bot}$
- Effective C_{adj} depends on behavior of neighbors
  - *Miller effect*

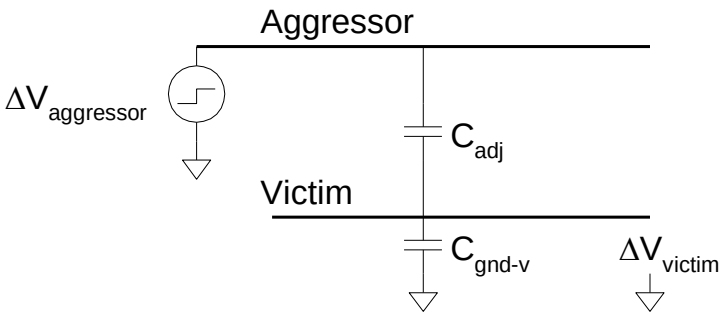


B	**ΔV**	**$C_{eff(A)}$**	**MCF**
Constant	V_{DD}	$C_{gnd} + C_{adj}$	1
Switching with A	0	C_{gnd}	0
Switching opposite A	$2V_{DD}$	$C_{gnd} + 2\ C_{adj}$	2

Crosstalk Noise

- Crosstalk causes noise on nonswitching wires
- If victim is floating:
  - model as capacitive voltage divider

$$\Delta V_{victim} = \frac{C_{adj}}{C_{gnd-v} + C_{adj}} \Delta V_{aggressor}$$

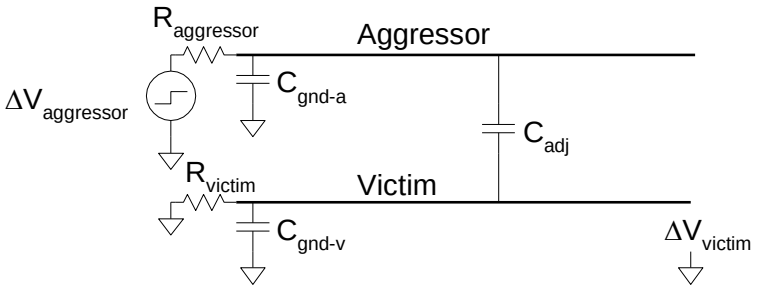


Driven Victims

- Usually victim is driven by a gate that fights noise
  - Noise depends on relative resistances
  - Victim driver is in linear region, agg. in saturation
  - If sizes are same, $R_{aggressor}$ = 2-4 x R_{victim}

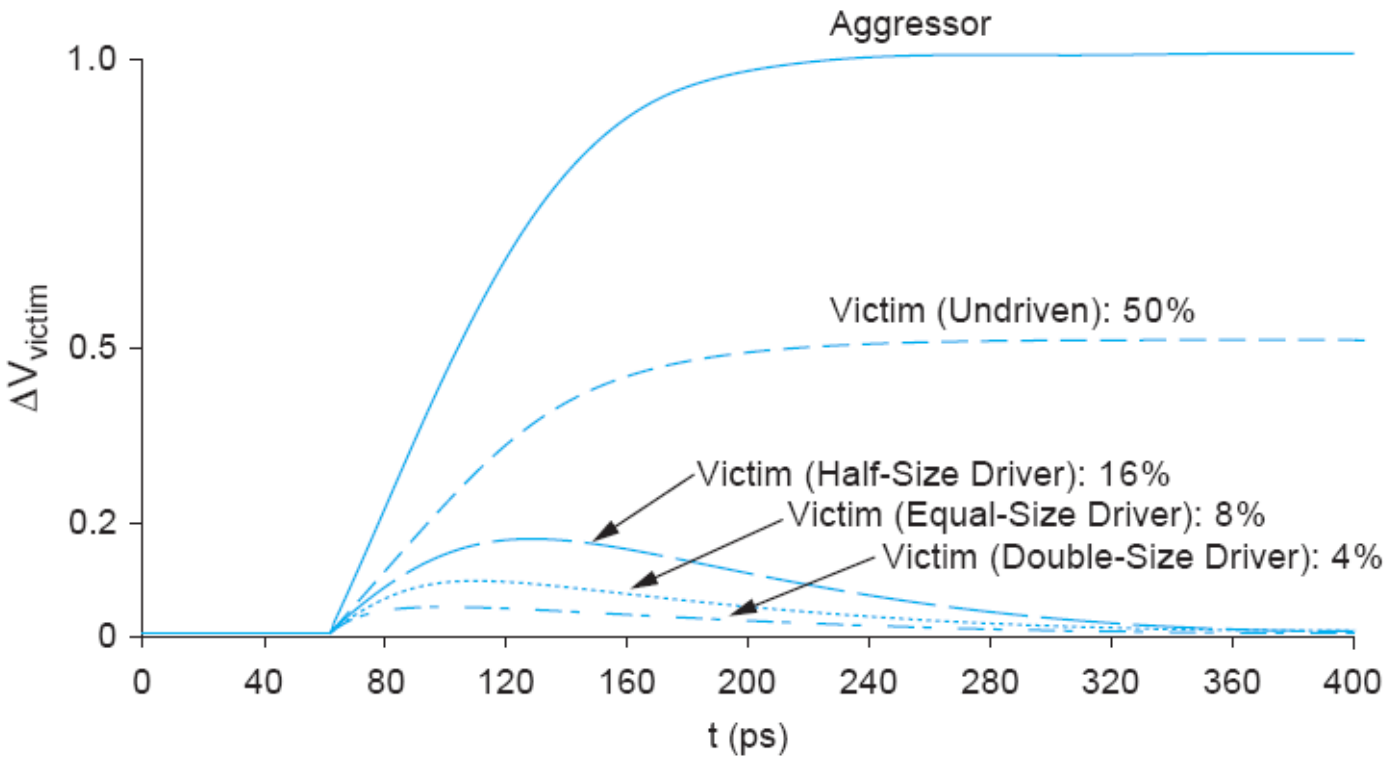
$$\Delta V_{victim} = \frac{C_{adj}}{C_{gnd-v} + C_{adj}} \frac{1}{1+k} \Delta V_{aggressor}$$

$$k = \frac{\tau_{aggressor}}{\tau_{victim}} = \frac{R_{aggressor}\left(C_{gnd-a} + C_{adj}\right)}{R_{victim}\left(C_{gnd-v} + C_{adj}\right)}$$



Coupling Waveforms

- Simulated coupling for $C_{adj} = C_{victim}$

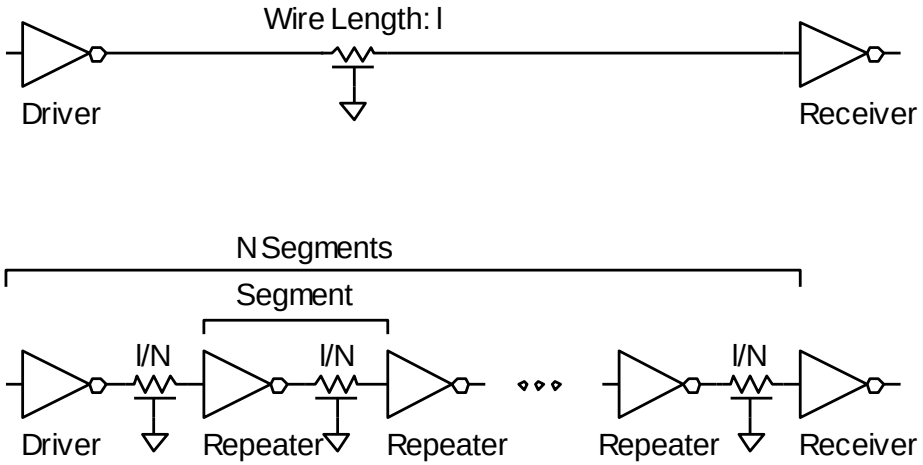


Noise Implications

- *So what* if we have noise?
- If the noise is **less than the noise margin**, nothing happens
- **Static CMOS** logic will eventually settle to correct output even if disturbed by large noise spikes
  - But glitches cause extra delay
  - Also cause extra power from false transitions
- **Dynamic logic never recovers** from glitches
- **Memories** and other sensitive circuits also can produce the wrong answer

Repeaters

- R and C are proportional to
- RC delay is proportional to
  - Unacceptably great for long wires
- Break long wires into N shorter segments
  - Drive each one with an inverter or buffer



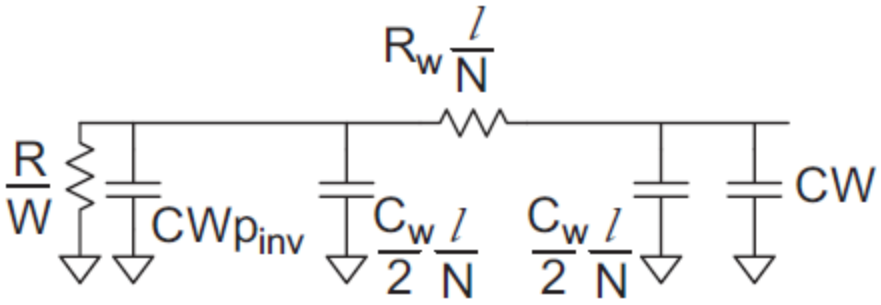
Repeater Design

- How many repeaters should we use?
- How large should each one be?
- Equivalent Circuit
  - Wire length *l/N*
    - Wire Capacitance C_w*l/N, Resistance R_w*l/N
  - Inverter width W (nMOS = W, pMOS = 2W)
    - Gate Capacitance C'*W, Resistance R/W

Repeater Results

$$t_{pd} = N\left[\frac{R}{W}\left(C_w\frac{l}{N} + CW\left(1+p_{\text{inv}}\right)\right) + R_w\frac{l}{N}\left(\frac{C_w}{2}\frac{l}{N} + CW\right)\right]$$

- Differentiate with respect to W and N. Set equal to 0, solve



$$\frac{l}{N} = \sqrt{\frac{2RC\left(1+p_{\text{inv}}\right)}{R_w C_w}}$$

$$\frac{t_{pd}}{l} = \left(2+\sqrt{2\left(1+p_{\text{inv}}\right)}\right)\sqrt{RCR_wC_w} \approx 1.67\sqrt{\text{FO4}\,R_wC_w}$$

$$W = \sqrt{\frac{RC_w}{R_wC}}$$

Repeater Energy

- Energy / length ≈ $1.87C_wV_{DD}^2$
  - 87% premium over unrepeated wires
  - The extra power is consumed in the large repeaters
- If the repeaters are downsized for minimum EDP:
  - Energy premium is only 30%
  - Delay increases by 14% from min delay

Repeater Example

Example 6.10

Compute the delay per mm of a repeated wire in a 65 nm process. Assume the wire is on a middle routing layer and has 2x width, spacing, and height, so its resistance is 200 Ω/mm and capacitance is 0.2 pF/mm. The FO4 inverter delay is 15 ps. Also find the repeater spacing and driver size to achieve this delay and the energy per bit.

SOLUTION: Using EQ (6.29), the delay is

$$\frac{t_{pd}}{l} = \left(2 + \sqrt{2(1+p_{inv})}\right)\sqrt{RCR_wC_w} \approx 1.67\sqrt{\text{FO4}\, R_wC_w}$$

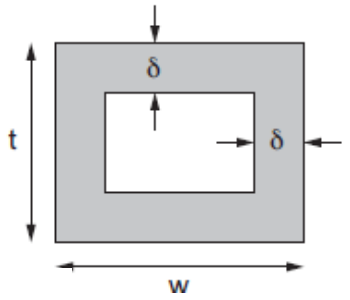
$$t_{pd} = 1.67\sqrt{(15\text{ ps})(200\ \Omega/\text{mm})(0.2\text{ pF/mm})} = 41\text{ ps/mm} \qquad (6.32)$$

This delay is achieved using a spacing of 0.45 mm between repeaters and an nMOS driver width of 18 μm (180x unit size). The energy per bit is 0.4 pJ/mm.

$$\text{Energy / length} \approx 1.87 C_w V_{DD}^2$$

Skin Effect

- At high frequency, current tends to flow near the surface of the conductor, and thus reducing the effective cross-sectional area of the thick conductor and raising resistance.
- Skin depth (δ) is the depth where the current falls off e^{-1} c value. **($e^{-1} = 0.37$)**



r he skin depth for a conductor is

$$\delta = \sqrt{\frac{2\rho}{\omega\mu}} \qquad (6.14)$$

where μ is the magnetic permeability of the dielectric (normally the same as in free space, $4\pi \times 10^{-7}$ H/m). T

Therefore, the frequency associated with the edge can be approximated as

$$\omega = \frac{2\pi}{8.65\, t_{rf}} \qquad (6.15)$$

where t_{rf} is the average 20–80% rise/fall time.

Skin Depth Example

Determine the skin depth for a copper wire in a chip with 20 ps edge rates.

SOLUTION: According to EQ (6.15), the maximum frequency of interest is

$$\omega = \frac{2\pi}{8.65 \times 20 \text{ ps}} = 3.6 \times 10^{10} \text{ rad/s} = 5.8 \text{ GHz}$$