Introduction to CMOS VLSI Design

Lecture 6: Wires

Outline

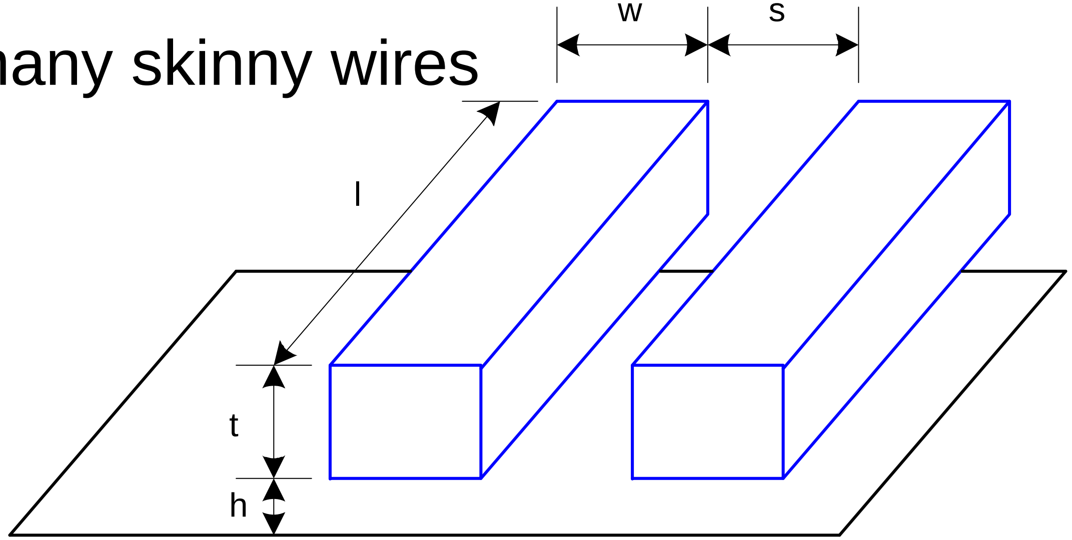
- Introduction
- Wire Resistance
- Wire Capacitance
- Wire RC Delay
- Crosstalk
- Wire Engineering
- Repeaters

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

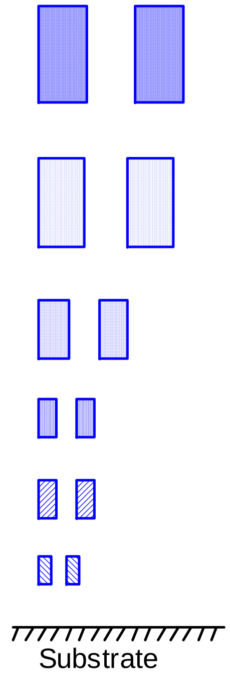
- Pitch = w + s
- Aspect ratio: AR = t/w
  - Old processes had AR << 1
  - Modern processes have AR ≈ 2
    - Pack in many skinny wires



Layer Stack

- AMI 0.6 μm process has 3 metal layers
- Modern processes use 6-10+ metal layers
- Example:
  - Intel 180 nm process
- M1: thin, narrow (< 3λ)
  - High density cells
- M2-M4: thicker
  - For longer wires
- M5-M6: thickest
  - For V_{DD}, GND, clk

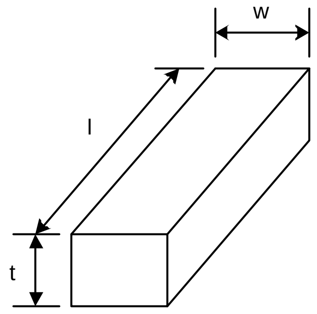
Layer	**T** (nm)	**W** (nm)	**S** (nm)	**AR**
6	1720	860	860	2.0
	1000			
5	1600	800	800	2.0
	1000			
4	1080	540	540	2.0
	700			
3	700	320	320	2.2
	700			
2	700	320	320	2.2
	700			
1	480	250	250	1.9
	800			



Wire Resistance

- ρ = *resistivity* ($\underline{\Omega\text{*m}}$)

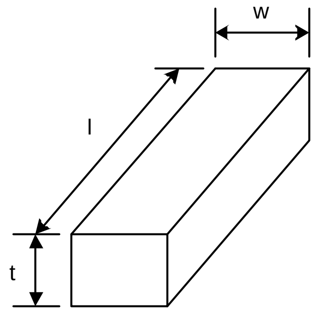
$$R =$$



Wire Resistance

- ρ = *resistivity* (Ω*m)

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



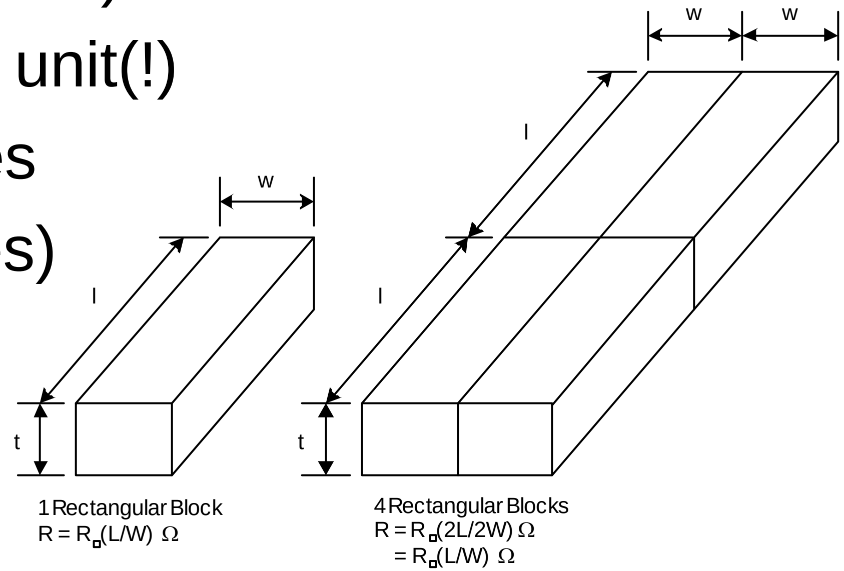
Wire Resistance

- 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
  - $R = R_{\square}$ * (# of squares)

Identical Resistance



Choice of Metals

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

Metal	**Bulk resistivity (μΩ*cm)**
Silver (Ag)	1.6
Copper (Cu)	1.7
Gold (Au)	2.2
Aluminum (Al)	2.8
Tungsten (W)	5.3
Molybdenum (Mo)	5.3

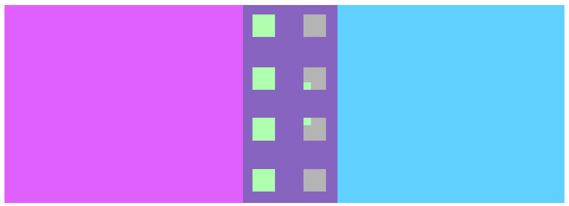
Sheet Resistance

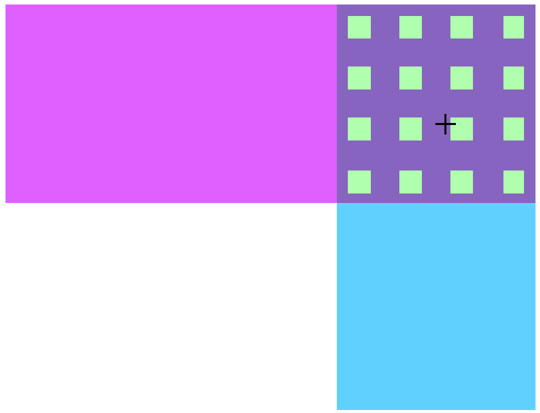
- Typical sheet resistances in 180 nm process

Layer	**Sheet Resistance (Ω/□)**
Diffusion (silicided)	3-10
Diffusion (no silicide)	50-200
Polysilicon (silicided)	3-10
Polysilicon (no silicide)	50-400
Metal1	0.08
Metal2	0.05
Metal3	0.05
Metal4	0.03
Metal5	0.02
Metal6	0.02

Contacts Resistance

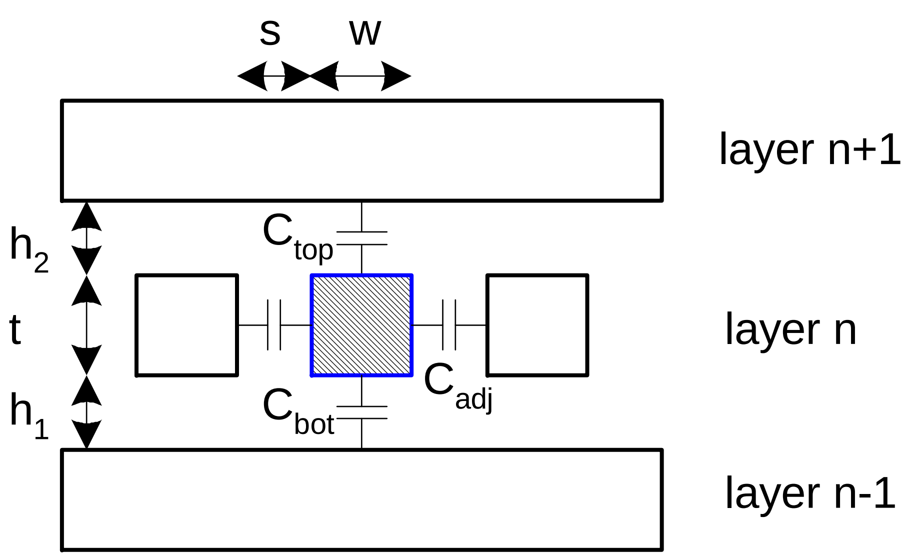
- Contacts and vias also have 2-20 Ω
- Use many contacts for lower R
  - Many small contacts for current crowding around periphery





Wire Capacitance (R&C결합)

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

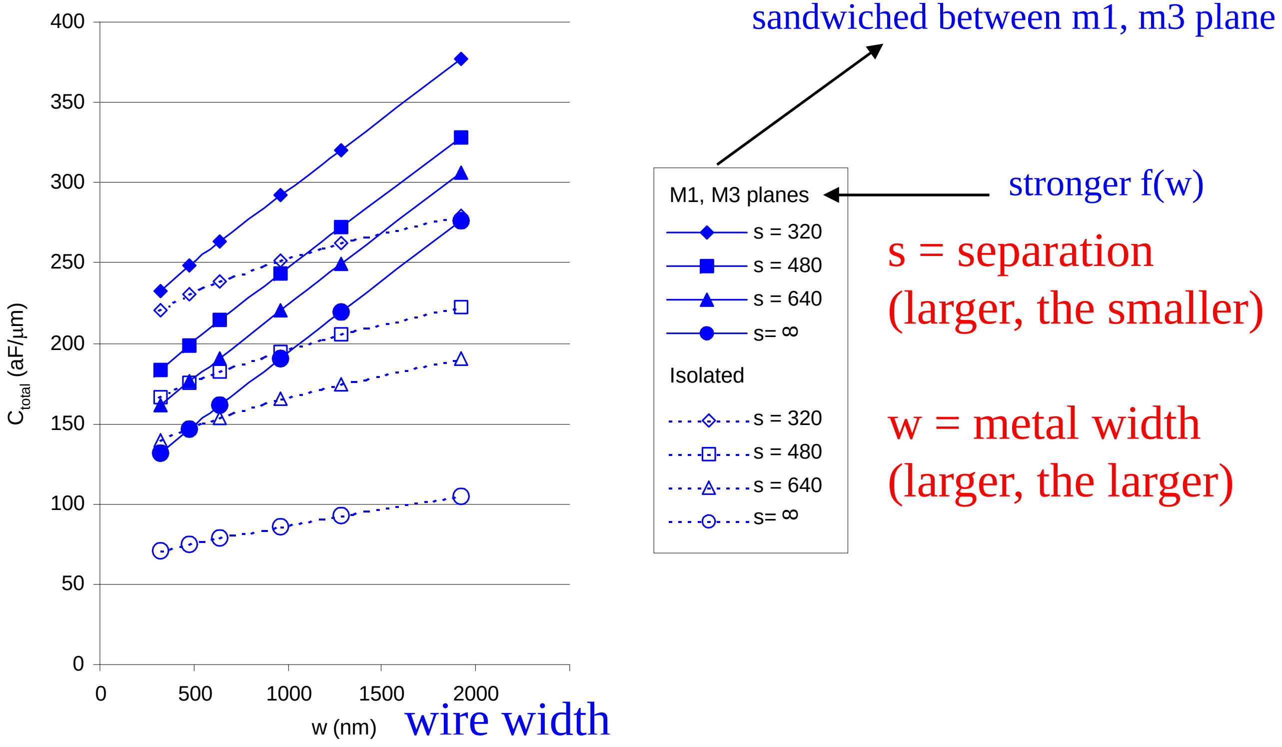


Capacitance Trends

- Parallel plate equation: $C = \varepsilon A/d$
  - Wires are not parallel plates, but obey trends
  - Increasing area (W, t) increases capacitance
  - Increasing distance (s, h) decreases capacitance
- Dielectric constant
  - $\varepsilon = k\varepsilon_0$
- $\varepsilon_0 = 8.85 \times 10^{-14}$ 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

M2 Capacitance Data

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



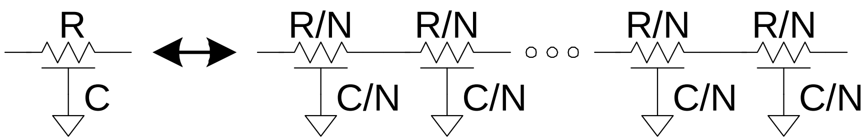
Diffusion & Polysilicon

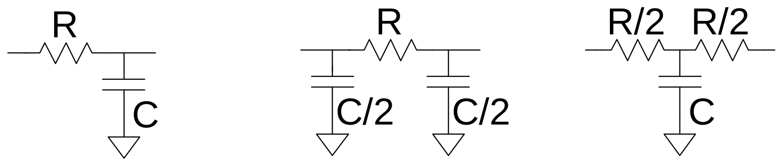
- Diffusion capacitance is very high (about 2 fF/μm)
  - Comparable to gate capacitance
  - Diffusion also has high resistance
  - Avoid using diffusion *runners* for wires!
- Polysilicon has lower C but high R
  - Use for transistor gates
  - Occasionally for very short wires between gates

Lumped Element Models

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

N segments





L-model π-model T-model

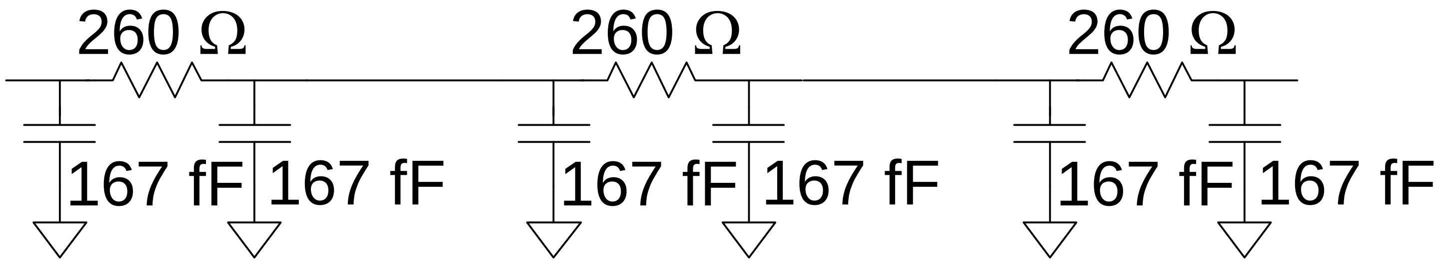
- 3-segment π-model is accurate to 3% in simulation
- L-model needs 100 segments for same accuracy!
- Use single segment π-model for Elmore delay

Example

- Metal2 wire in 180 nm process
  - 5 mm long
  - 0.32 μm wide
- Construct a 3-segment π-model
  - $R_{\square}$ =
  - $C_{\text{permicron}}$ =

Example

- Metal2 wire in 180 nm process
  - 5 mm long
  - 0.32 μm wide
- Construct a 3-segment π-model 5mm/0.32um*sheet R
  - $R_{\square} = 0.05\ \Omega/\square$ => $R = 781\ \Omega$
  - $C_{permicron} = 0.2$ fF/μm *5mm => C = 1 pF



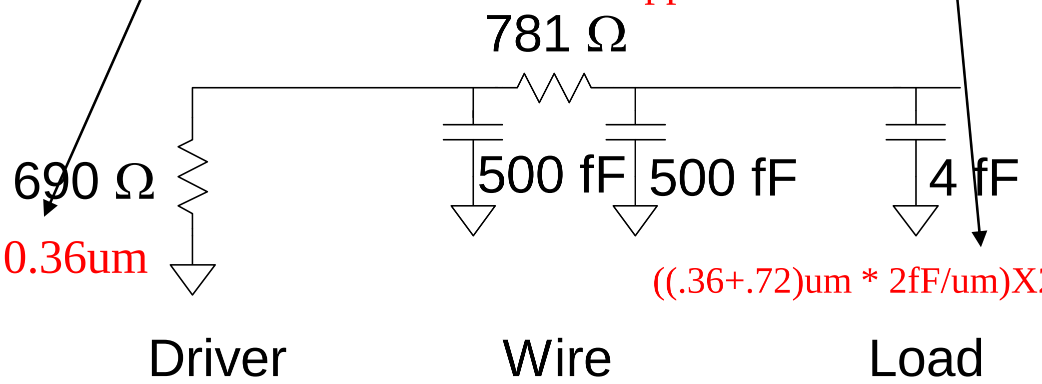
Wire RC Delay

- Estimate the delay of a 10x inverter driving a 2x inverter at the end of the 5mm wire from the previous example.
  - R = 2.5 kΩ*μm for gates
  - Unit inverter: 0.36 μm nMOS, 0.72 μm pMOS
  - t_{pd} =

Wire RC Delay

- Estimate the delay of a 10x inverter driving a 2x inverter at the end of the 5mm wire from the previous example.
  - R = 2.5 kΩ*μm for gates (10X gate)
  - Unit inverter: 0.36 μm nMOS, 0.72 μm pMOS

Phi model from pp. 18



2.5Kohm*um / 0.36um

((.36+.72)um * 2fF/um)X2(2x inv.)

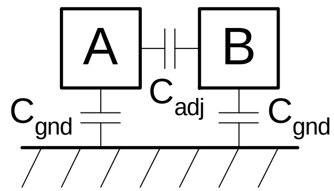
  - t_{pd} = 1.1 ns Elmore delay : 690*500+(690+781)*(500+4)

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*

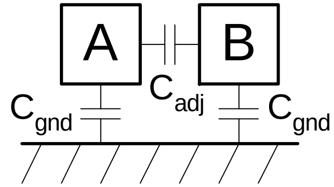


B	ΔV	$C_{eff(A)}$	MCF
Constant			
Switching with A			
Switching opposite A			

Crosstalk Delay

- Assume layers above and below on average are quiet
  - Second terminal of capacitor can be ignored
    Small signal GND in case of Vdd
  - Model as $C_{gnd} = C_{top} + C_{bot}$ (Because above & below @ const. V)
- Effective C_{adj} depends on behavior of neighbors
  - *Miller effect*

(remember 전자회로, think of effective terminal vtg.)



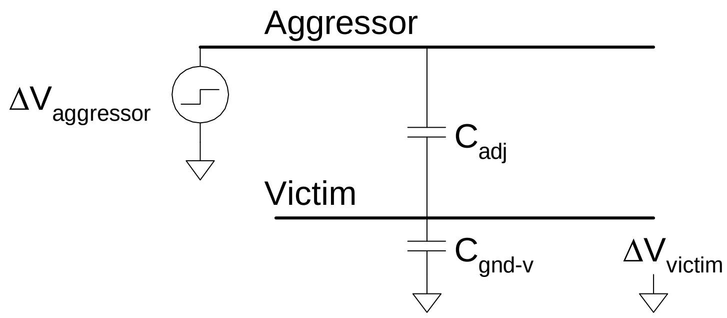
B	ΔV	$C_{eff(A)}$	MCF
Constant	V_{DD}	$C_{gnd} + C_{adj}$	1
Switching with A	0	C_{gnd}	0
Switching opposite A	$2V_{DD}$	$C_{gnd} + 2\,C_{adj}$	2

Crosstalk Noise

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

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

1/Cg / (1/Cg+1/Ca)



Driven Victims (w/o derivation)

(not floating as previous slide)

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

Aggressor dominates & victim in intermediate range

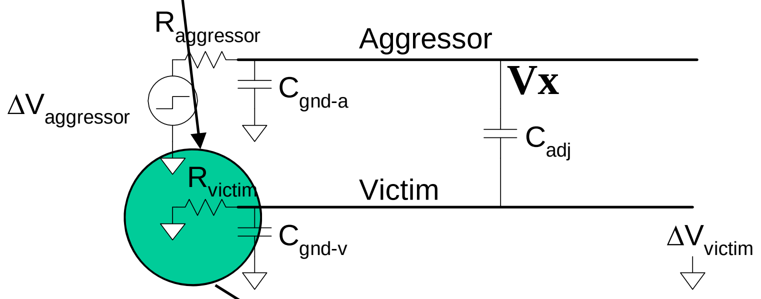
$$\Delta V_{victim} = \frac{C_{adj}}{C_{gnd-v} + C_{adj}} \boxed{\frac{1}{1+k} \Delta V_{aggressor}}$$

Vaggressor to Vx factor
(Detailed derivation, refer to ref. paper)

If C is equal, 1 /2* (1 + Ra/Rv)
Smaller the victim gate (larger Rv),
Larger the damage

Vx to Vvictim factor (as in the previous slide)

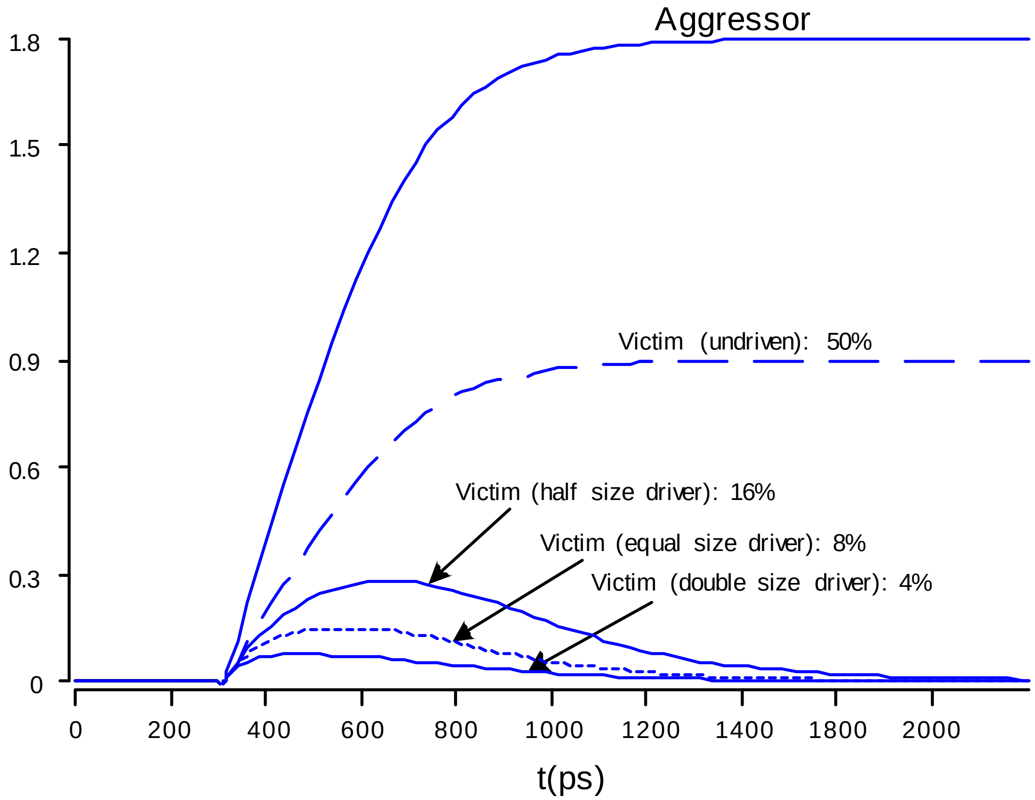
$$k = \frac{\tau_{aggressor}}{\tau_{victim}} = \frac{R_{aggressor}\left(C_{gnd-a} + C_{adj}\right)}{R_{victim}\left(C_{gnd-v} + C_{adj}\right)}$$



Instead of floating victim,
gate driven victim (noise reduced)

Coupling Waveforms

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



Noise Implications

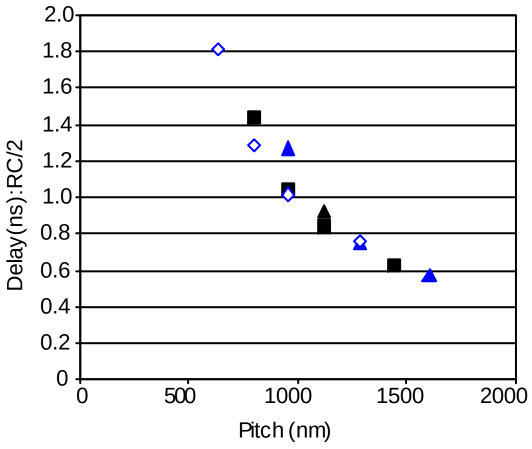
- *So what* if we have noise?
- If the noise is less than the noise margin, nothing happens
- Static CMOS logic will eventually settle to correct output even if disturbed by large noise spikes
  - But glitches cause extra delay
  - Also cause extra power from false transitions
- Dynamic logic never recovers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

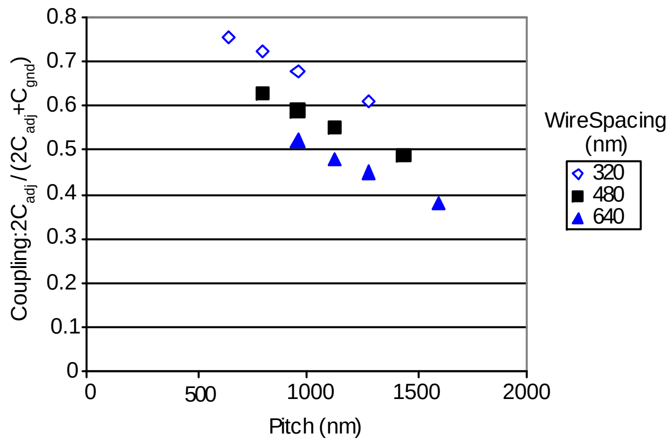
Wire Engineering

- Goal: achieve delay, area, power goals with acceptable noise
- Degrees of freedom:

Wire Engineering

- Goal: achieve delay, area, power goals with acceptable noise
- Degrees of freedom:
  - Width
  - Spacing

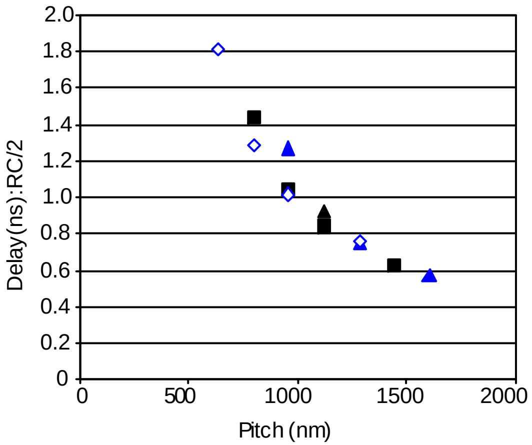


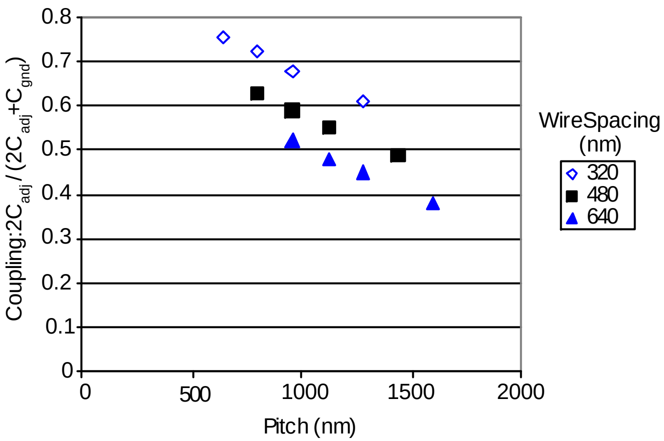


Larger the width, smaller the delay,
Larger the spacing, smaller the coupling

Wire Engineering

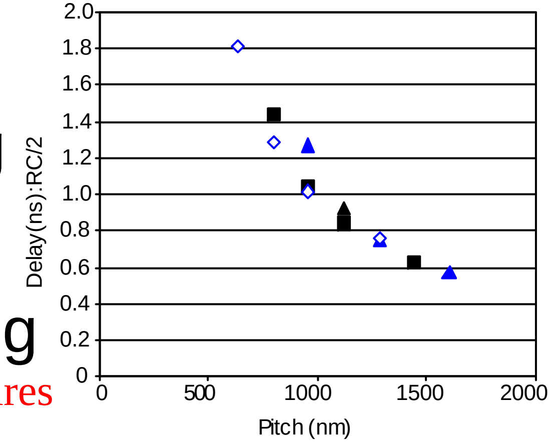
- Goal: achieve delay, area, power goals with acceptable noise
- Degrees of freedom:
  - Width
  - Spacing
  - Layer

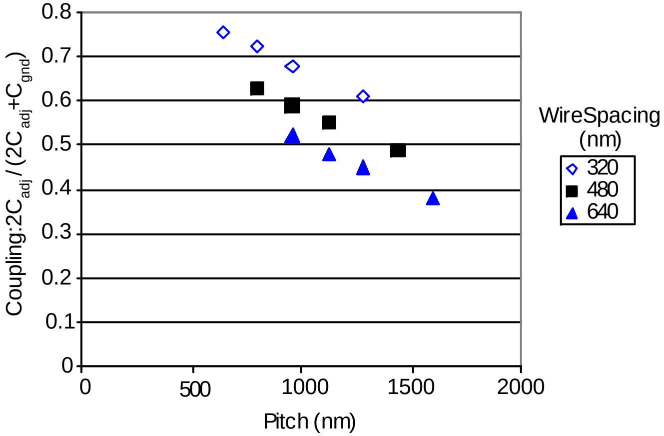




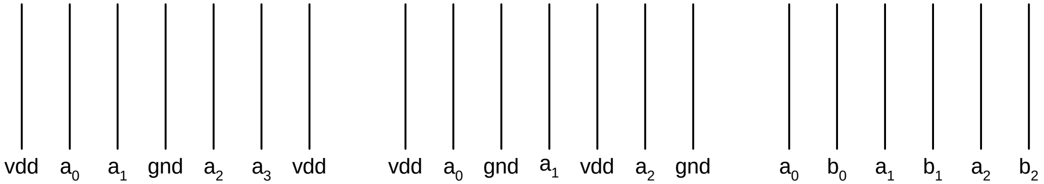
Wire Engineering

- Goal: achieve delay, area, power goals with acceptable noise
- Degrees of freedom:
  - Width
  - Spacing
  - Layer
  - Shielding





Shield noise sensitive wires with power lines

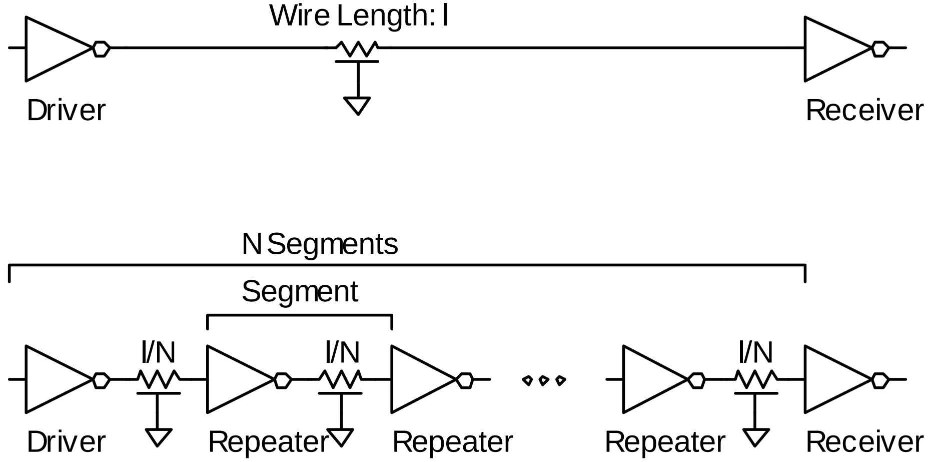


Repeaters

- R and C are proportional to l
- RC delay is proportional to l^2
  - Unacceptably great for long wires

Repeaters

- R and C are proportional to l
- RC delay is proportional to l^2
  - 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* *(total length l divided by N sections*
    - Wire Capaitance C_w*l/N, Resistance R_w*l/N
  - Inverter width W (nMOS = W, pMOS = 2W)
    - Gate Capacitance C'*W, Resistance R/W

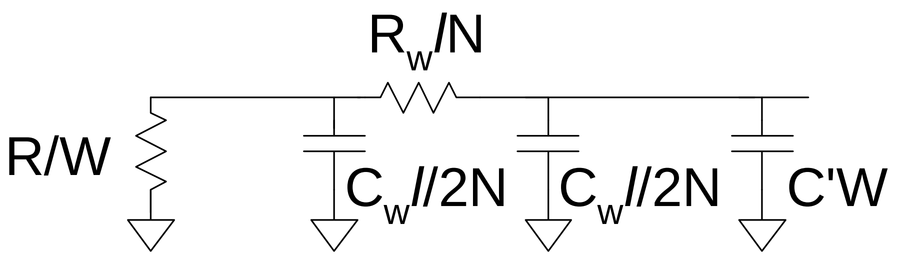
Repeater Design

- How many repeaters should we use?
- How large should each one be?
- Equivalent Circuit
  - Wire length *l*

All quantity per unit length

    - Wire Capacitance $C_w{*}l$, Resistance $R_w{*}l$
  - Inverter width W (nMOS = W, pMOS = 2W)
    - Gate Capacitance C'*W, Resistance R/W

C' → 3C(Cox) 2W PMOS 1W NMOS



Repeater Results

(Derivation @next slide)

- Write equation for Elmore Delay
  - Differentiate with respect to W and N
  - Set equal to 0, solve

$$\frac{l}{N} = \sqrt{\frac{2RC'}{R_w C_w}}$$

$$\frac{t_{pd}}{l} = \left(2+\sqrt{2}\right)\sqrt{RC'R_w C_w}$$

$$W = \sqrt{\frac{RC_w}{R_w C'}}$$

Best delay that can be minimized (optimized) per length

~60-80 ps/mm

in 180 nm process

Derivation

The Elmore delay of each segment is Slide 35의 Schematic 기준

$$t_{pd-seg} = \frac{R}{W}\left(\frac{C_w l}{N} + C'W\right) + \left(\frac{R_w l}{N}\right)\left(\frac{C_w l}{2N} + C'W\right)$$

The total delay is N times greater:

$$t_{pd} = NRC' + L\left(R_w C'W + \frac{RC_w}{W}\right) + L^2 \frac{R_w C_w}{2N}$$

Take the partial derivatives with respect to N and W and set them to 0 to minimize delay:

$$\frac{\partial t_{pd}}{\partial N} = RC' - l^2 \frac{R_w C_w}{2N^2} = 0 \Rightarrow N = l\sqrt{\frac{R_w C_w}{2RC'}}$$

$$\frac{\partial t_{pd}}{\partial W} = l\left(R_w C' - \frac{RC_w}{W^2}\right) = 0 \Rightarrow W = \sqrt{\frac{RC_w}{R_w C'}}$$

Using these gives a delay per unit length of

$$\frac{t_{pd}}{l} = \left(2 + \sqrt{2}\right)\sqrt{RC'R_w C_w}$$