Lecture 13:
Interconnects in CMOS Technology

Mark McDermott
Electrical and Computer Engineering
The University of Texas at Austin

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

HVH	VHV
M1: Horizontal	M1: Vertical
M2: Vertical	M2: Horizontal
M3: Horizontal	M3: Vertical
M4: Vertical	M4: Horizontal
:	:

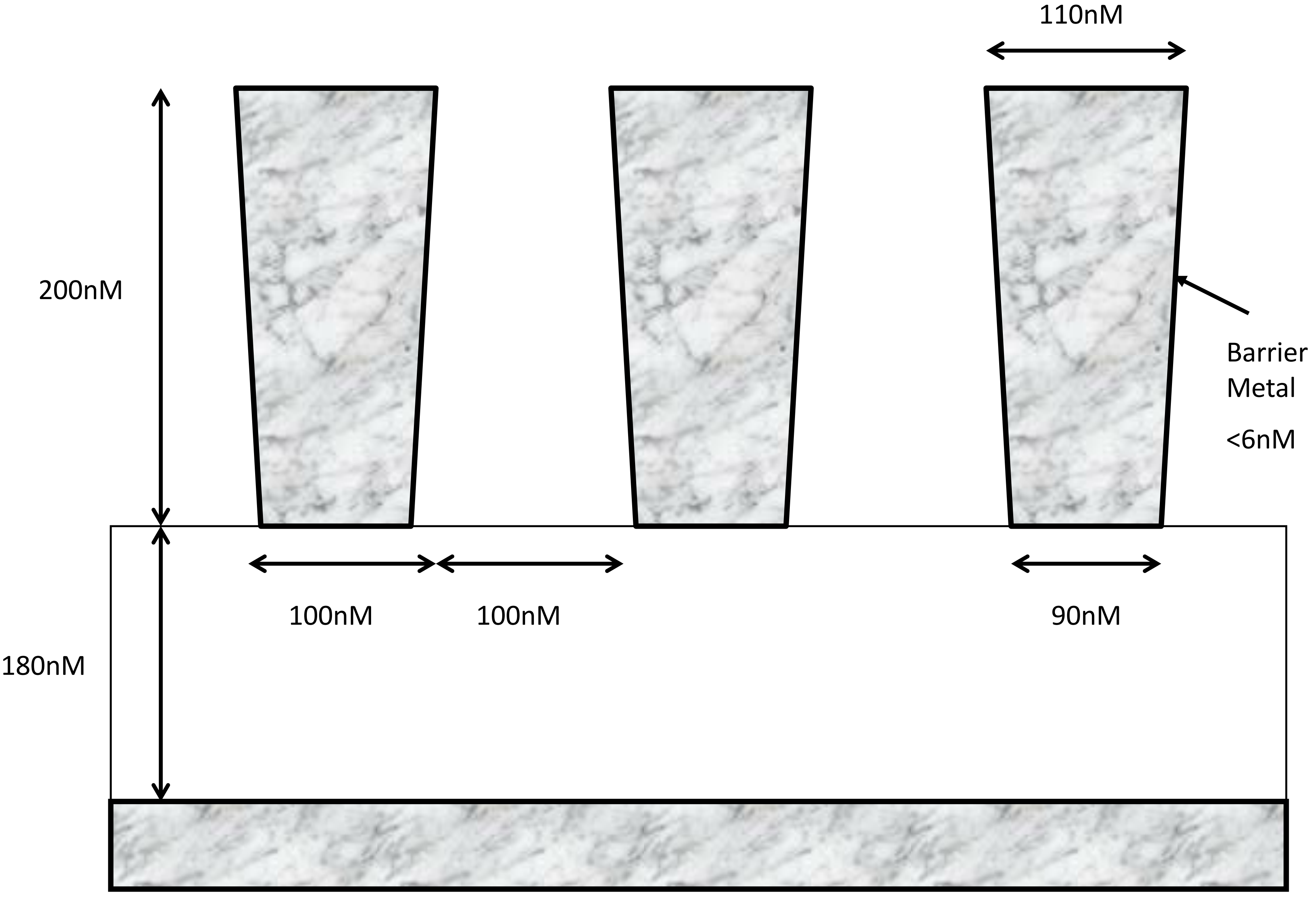
Choice of Metals

- **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

35%-40% improvement

Metal Layer Cross Section 65nM



45nm Interconnect

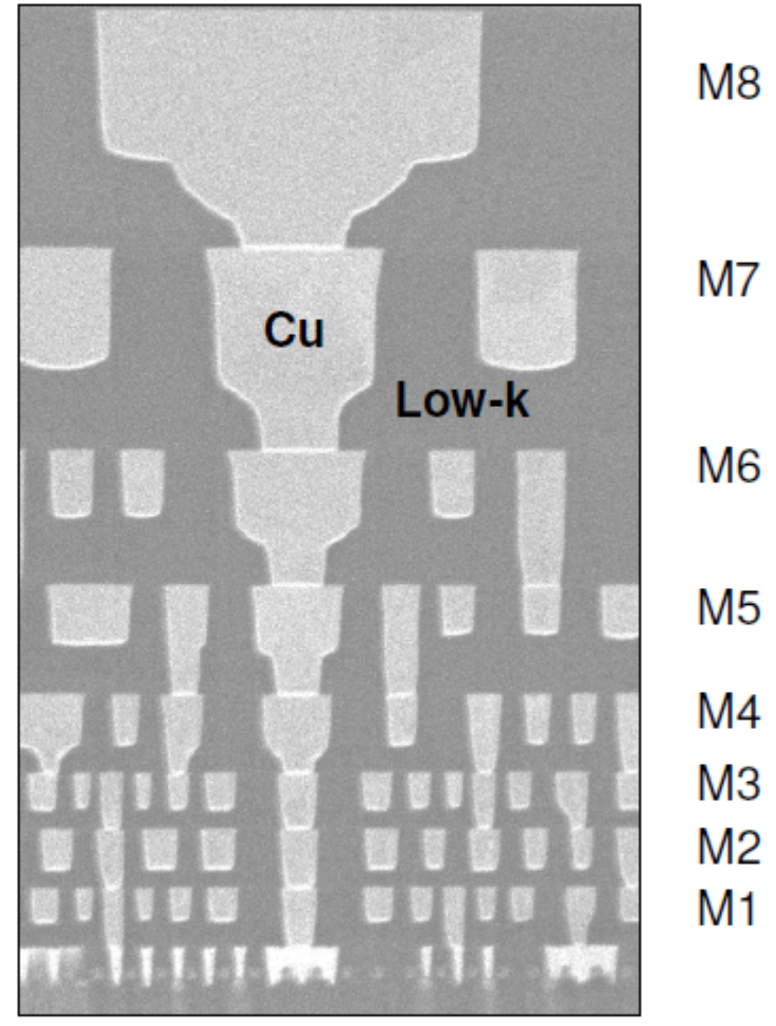
Loose pitch + thick metal on upper layers:

High speed global wires

Low resistance power grid

Tight pitch on lower layers:

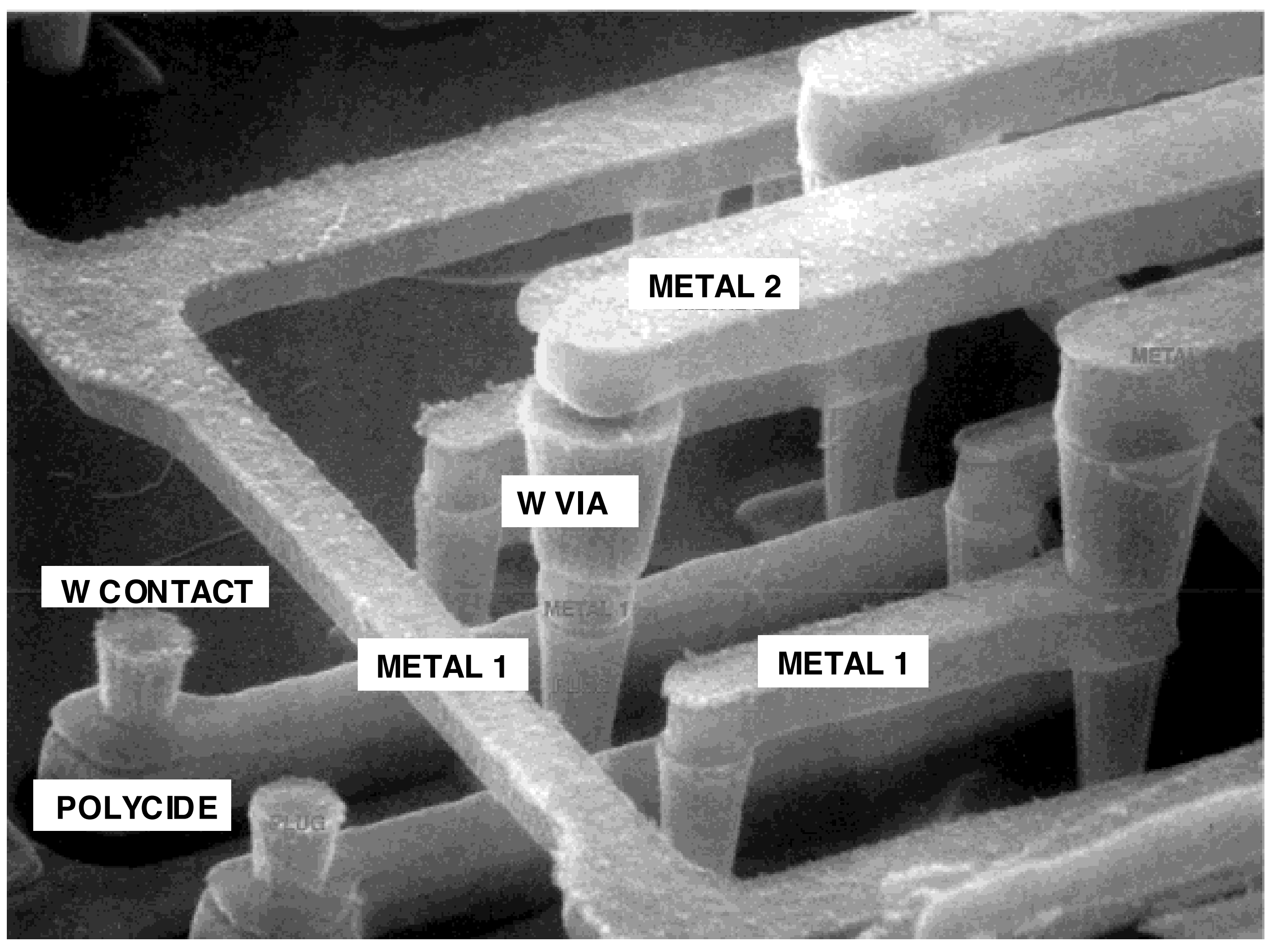
Maximum density for local interconnects



M8
M7
M6
M5
M4
M3
M2
M1

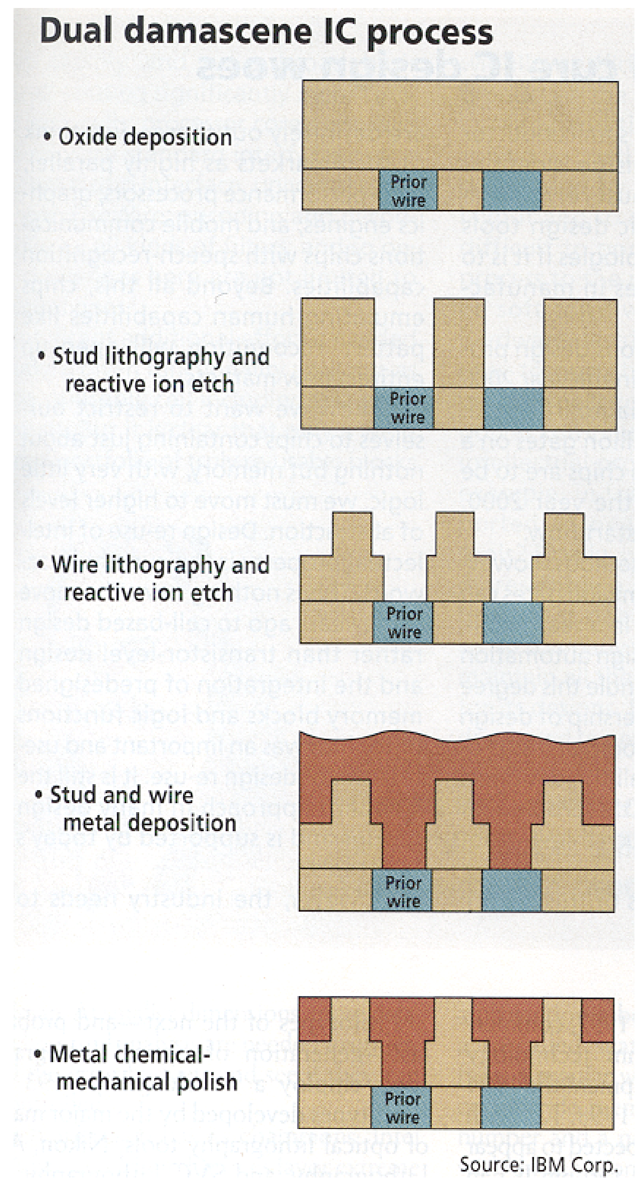
Source: Mark Bohr, Intel Corporation

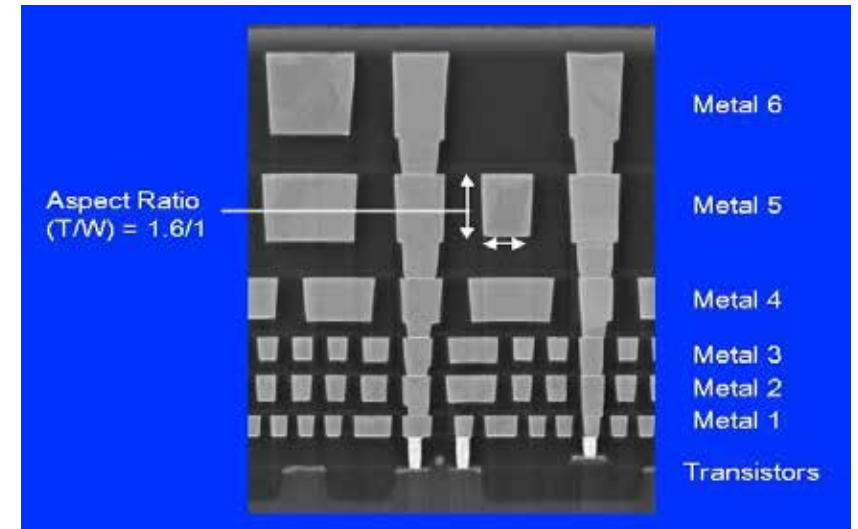
SEM MICRO-GRAPH (ILM DIELECTRIC REMOVED)

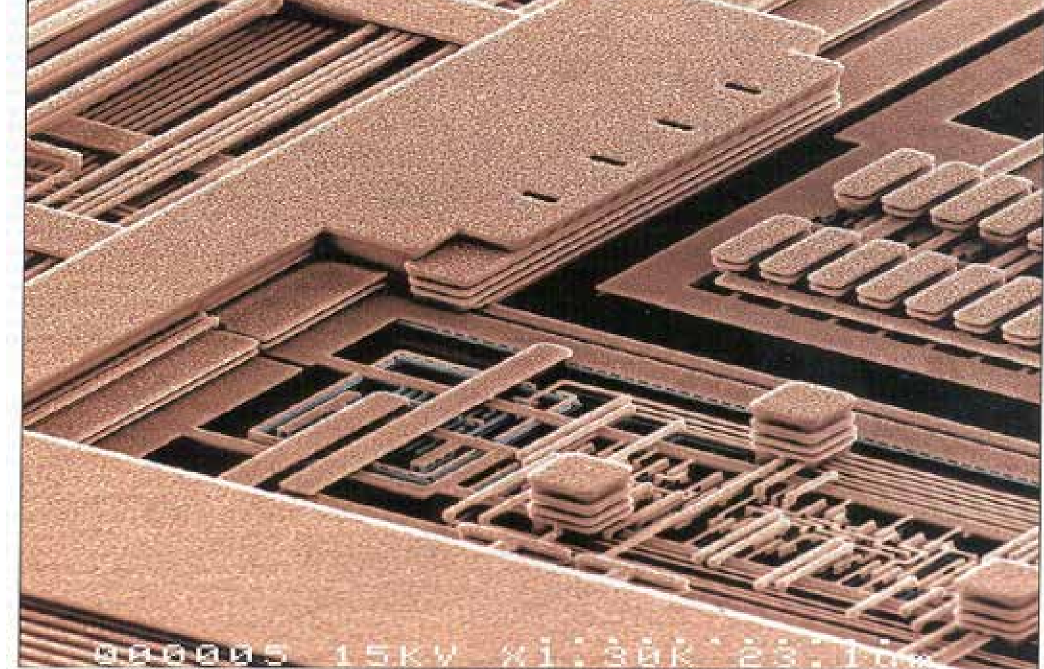


Courtesy: IBM

Advanced Metallization

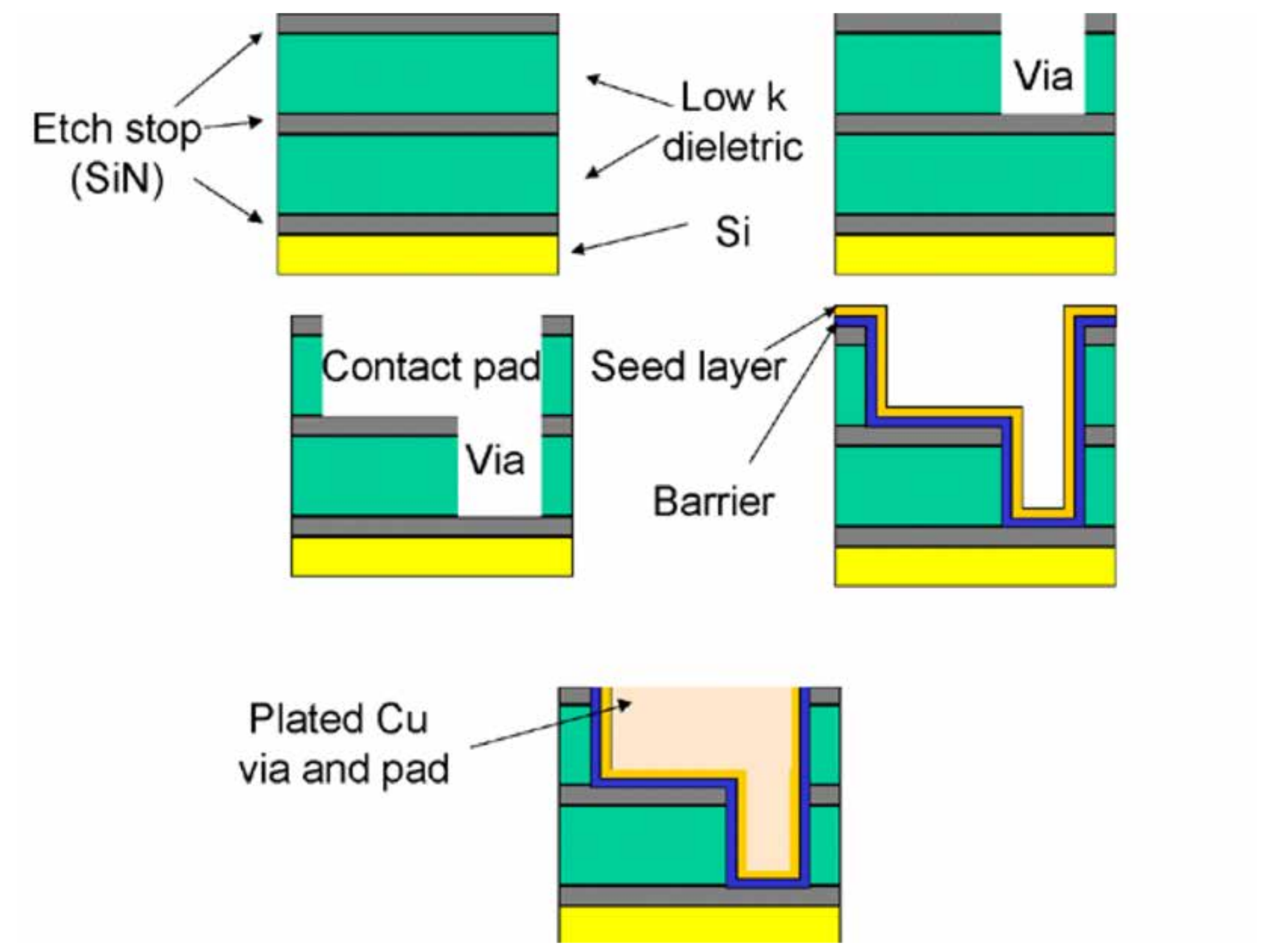






© Digital Integrated Circuits 2nd

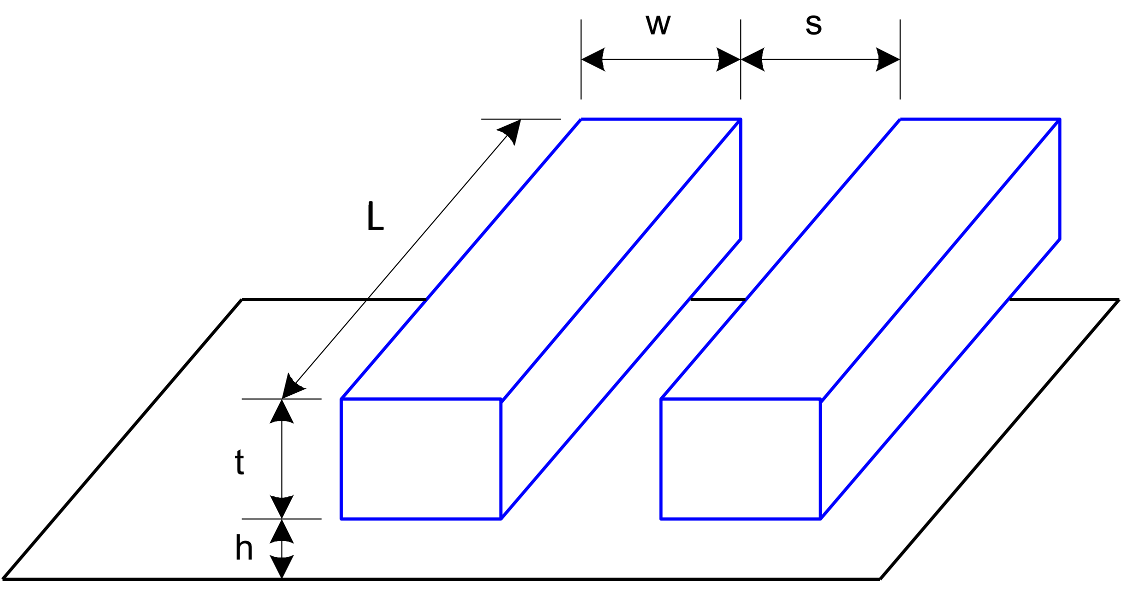
Interconnect Process Dual Damascene



•C.-K. Hu and J.M.E. Harper, *Mater. Chem. Phys.,* 52 (1998), p. 5.

WIRE GEOMETRY

- **PITCH = width + space**
- **Height (h) = distance to top/bottom routes**
- **ASPECT RATIO (AR) = thickness / width**
  - Deep submicron processes have AR < 2 to maintain sheet resistances at a reasonable level
  - Coupling to neighboring routes dominates



$R=\rho L/tw$

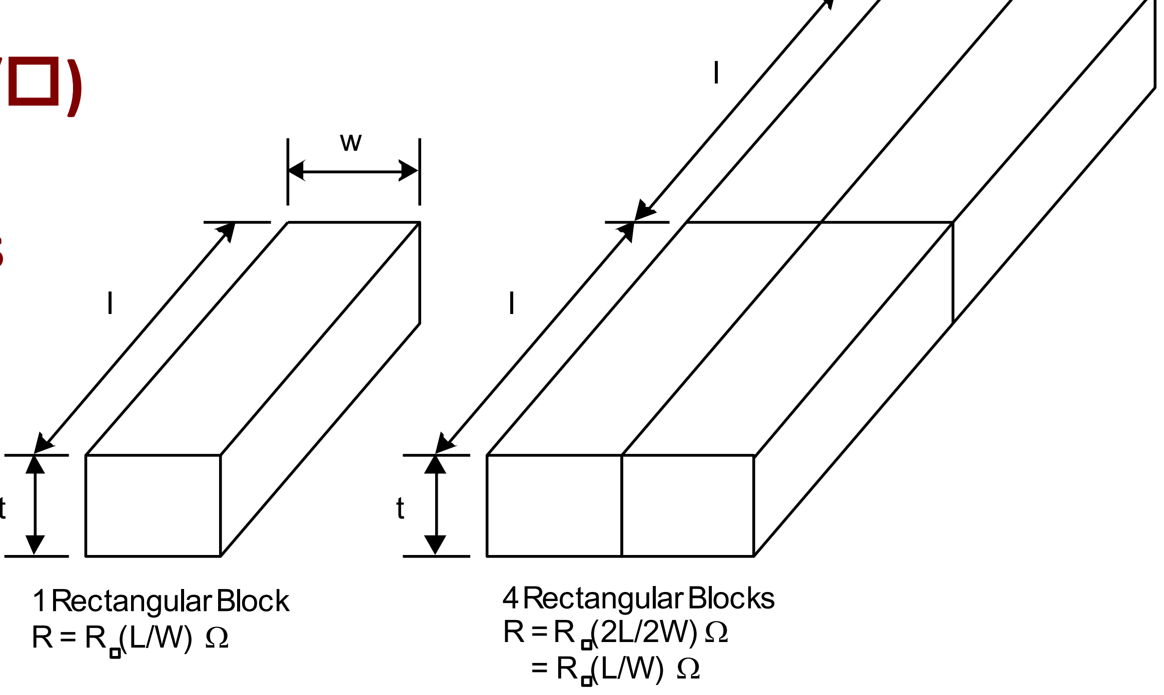
Wire Resistance

- ρ = ***resistivity*** **(Ω*m)**

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

where ρ=resistivity

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



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

CONTRIBUTION OF WIRES (they are not free)

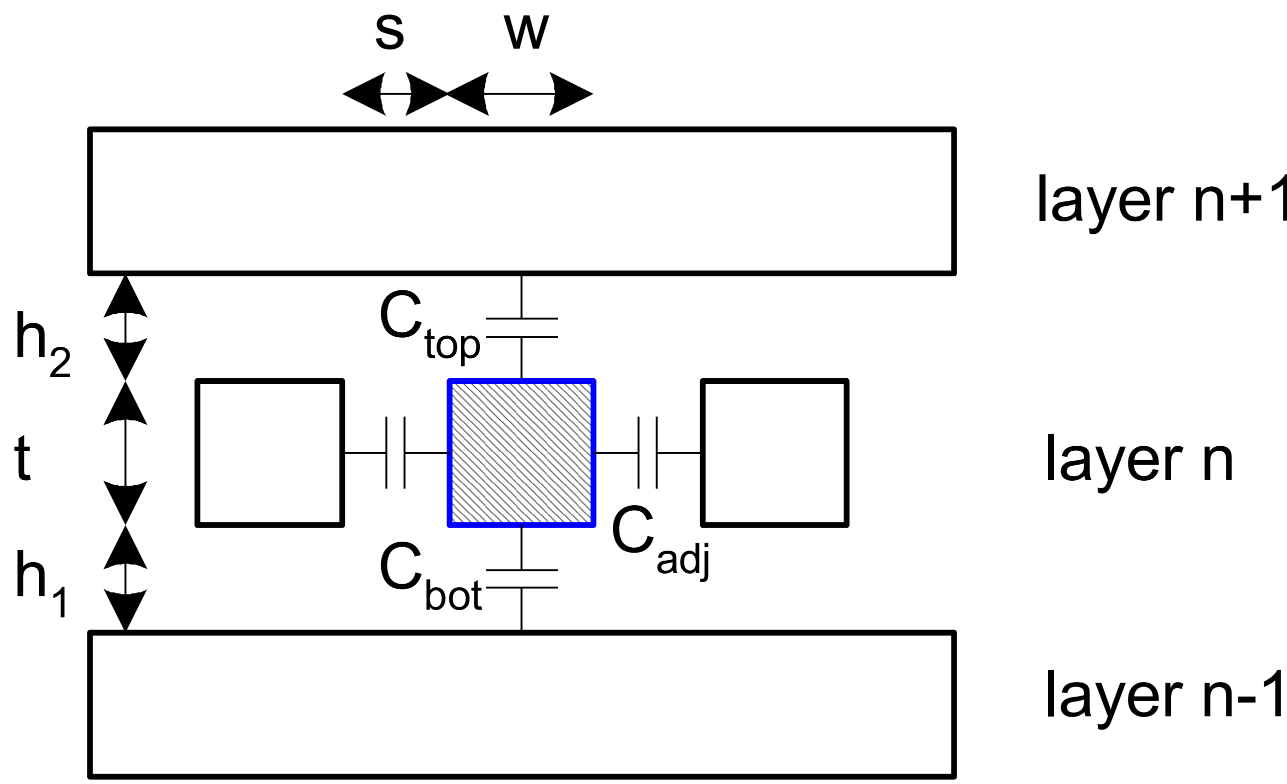
- **DELAY**
  - Function of R, L and C, VIA resistance, local .vs. global
- **POWER**
  - Charging/discharging of C is a function of CV2F
- **NOISE**
  - Attackers/victims impact on functionality and delay
- **POWER SUPPLY IR DROPS and GROUND BOUNCE**
  - Affects delay leading to timing failures
- **RELIABILITY**
  - Electro-migration, self heat, maximum current
- **COST**
  - Number of layers .vs. area/performance targets, yield

Contact/VIA Resistance

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

Wire Capacitance

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

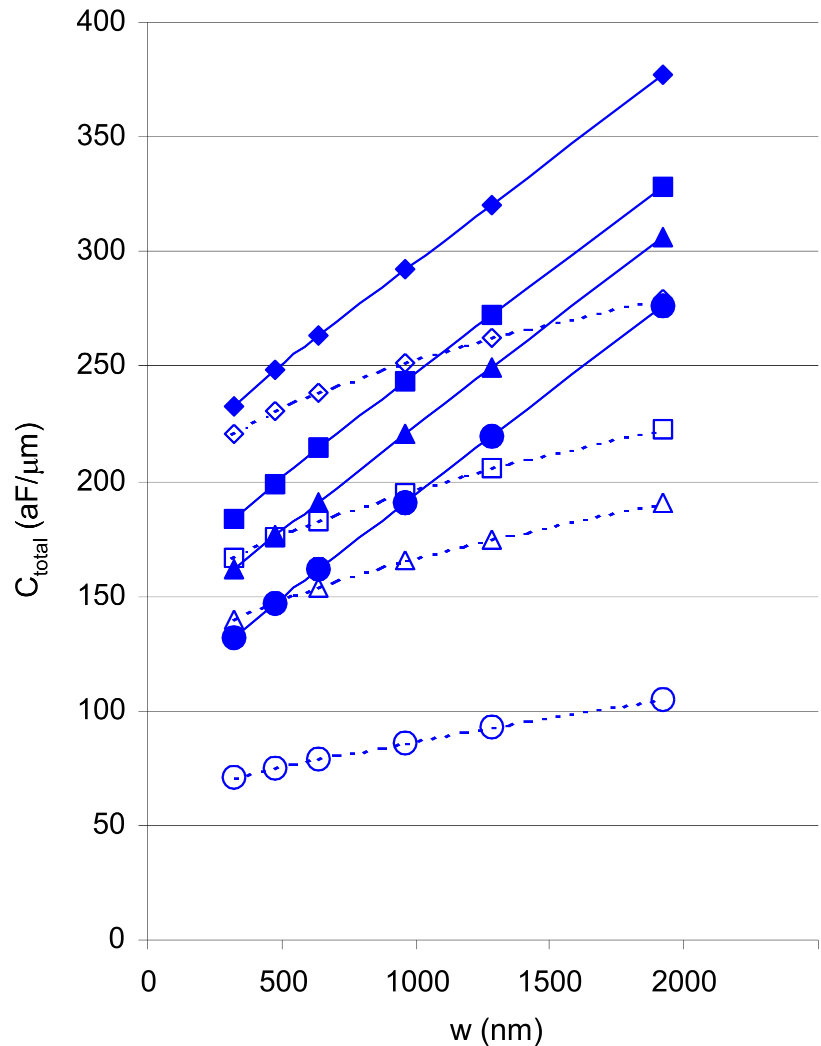


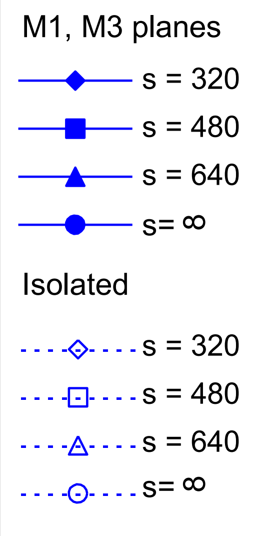
Capacitance Trends

- Parallel plate equation: $C = \varepsilon A/d$
  - 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



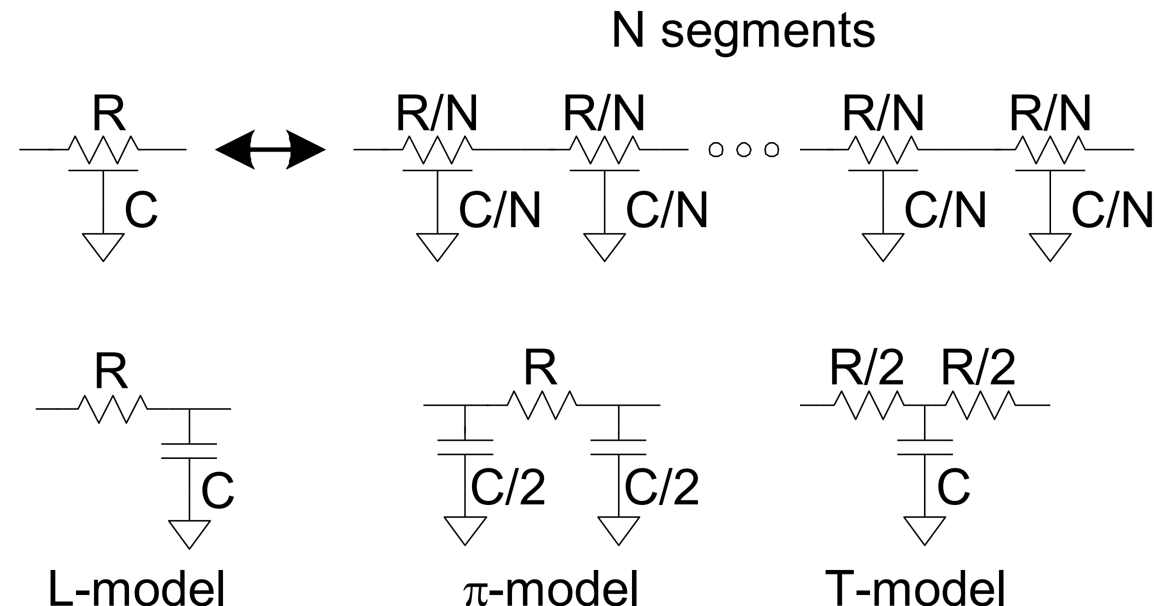


Diffusion & Polysilicon

- **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

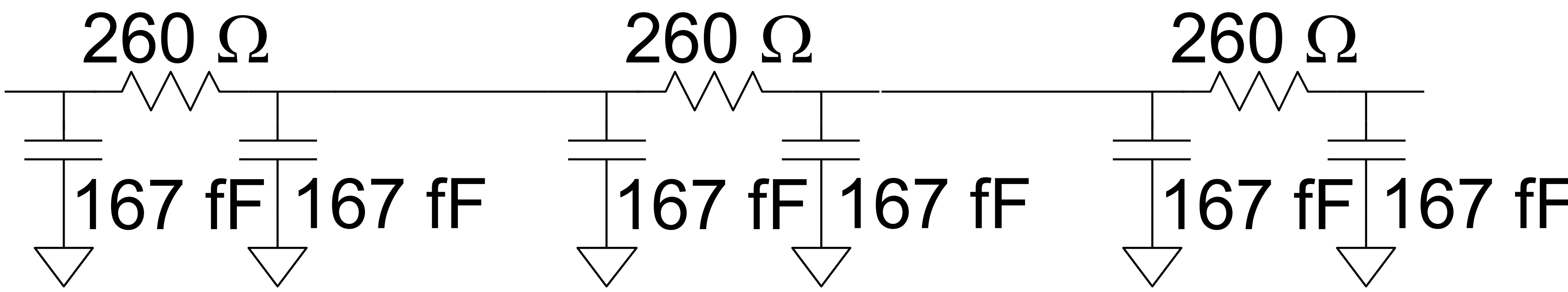
- **Wires are a distributed system**
  - **Approximate with lumped element models**



- **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**

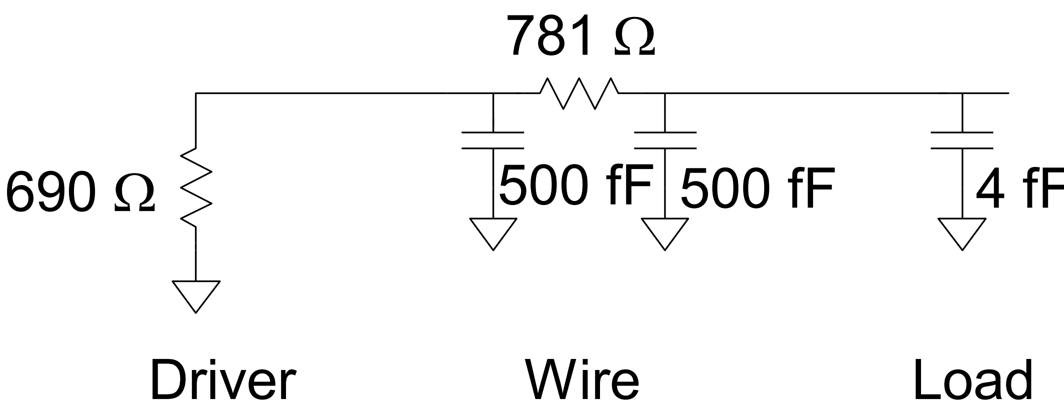
Example

- **Metal2 wire in 180 nm process**
  - 5 mm long
  - 0.32 μm wide
- **Construct a 3-segment π-model**
  - $R_{\square}$ = 0.05 Ω/☐ => R = 781 Ω
  - $C_{permicron}$ = 0.2 fF/μm => 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} = 1.1 ns



Crosstalk

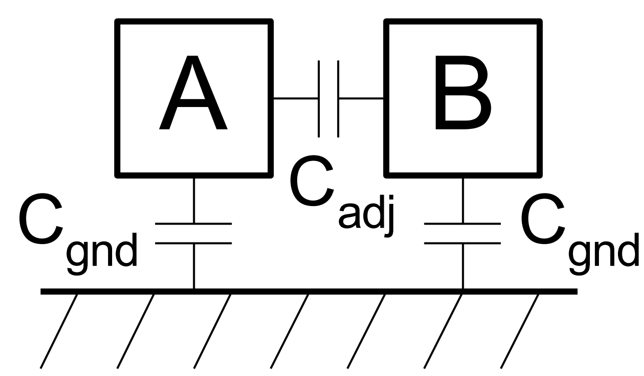
- **A capacitor does not like to change its voltage instantaneously.**

- **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 non-switching 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 Coupling Factor (MCF)***

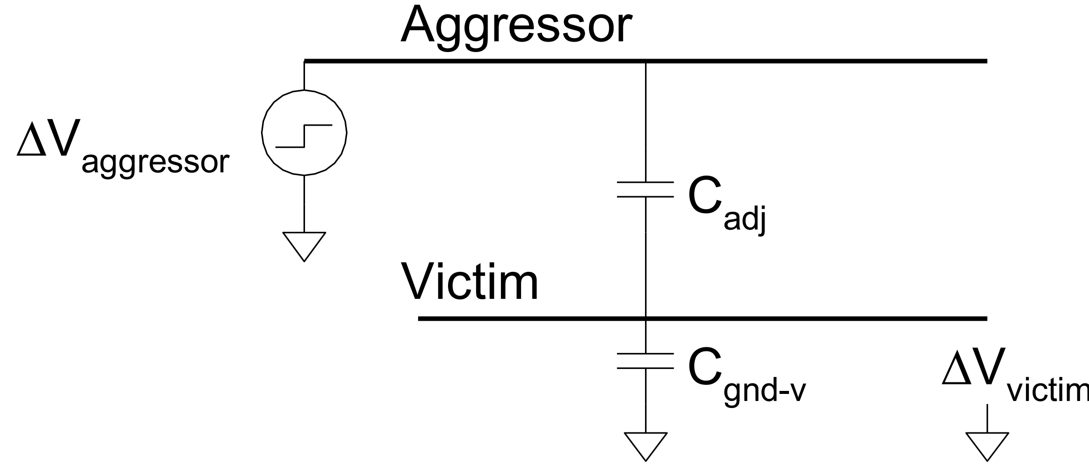


B	DV	$C_{eff(A)}$	MCF
Constant	V_{DD}	$C_{gnd} + C_{adj}$	1
Switching with A	0	C_{gnd}	0
Switching opposite A	$2V_{DD}$	$C_{gnd} + 2\ C_{adj}$	2

Crosstalk Noise

- **Crosstalk causes [functional/voltage] noise on nonswitching wires**
- **If victim is floating:**
  - **model as capacitive voltage divider**

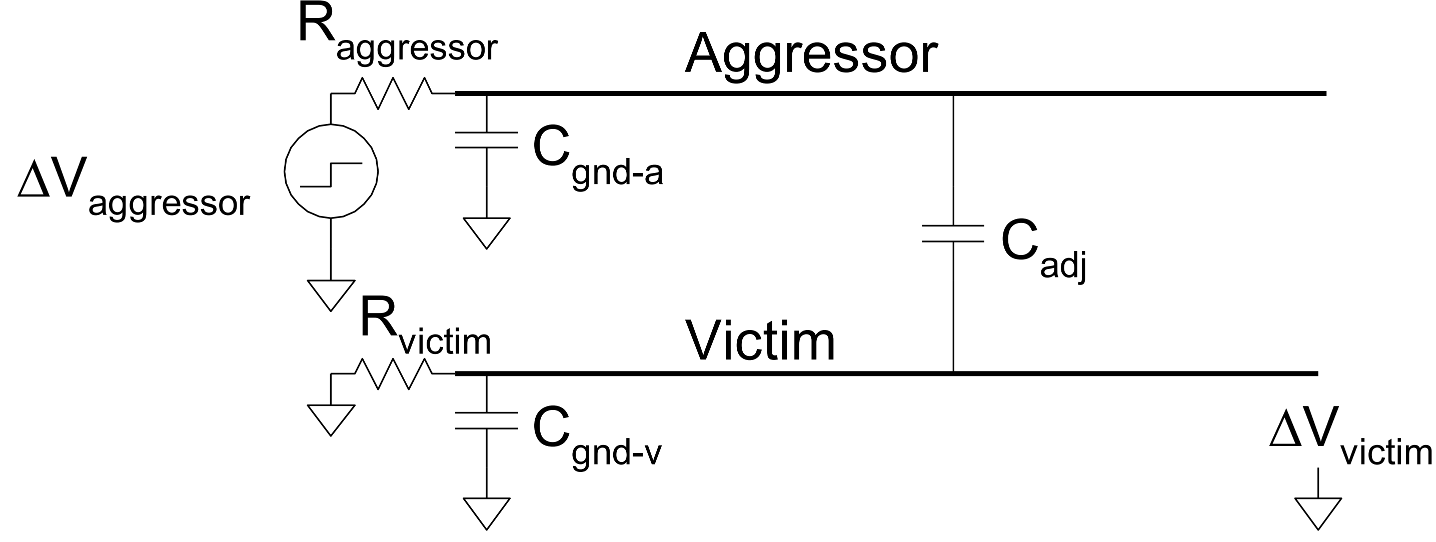
$$\Delta V_{victim} = \frac{C_{adj}}{C_{gnd-v} + C_{adj}} \Delta V_{aggressor}$$



Driven Victims

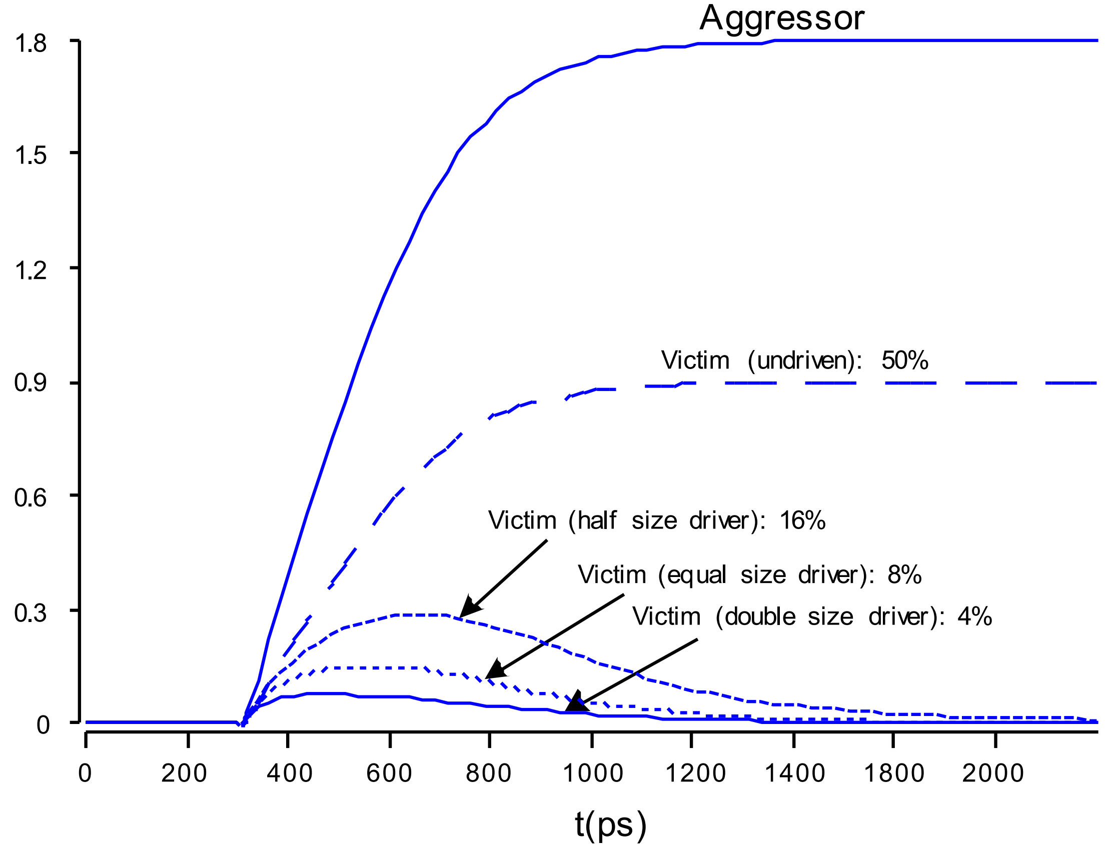
- **Usually victim is driven by a gate that fights noise**
  - **Noise depends on relative resistances**
  - **Victim driver is in linear region, aggressor in saturation**
  - **If sizes are same, $R_{aggressor}$ = 2-4 x R_{victim}**

$$\Delta V_{victim} = \frac{C_{adj}}{C_{gnd-v} + C_{adj}} \frac{1}{1+k} \Delta V_{aggressor} \qquad k = \frac{\tau_{aggressor}}{\tau_{victim}} = \frac{R_{aggressor}\left(C_{gnd-a} + C_{adj}\right)}{R_{victim}\left(C_{gnd-v} + C_{adj}\right)}$$



Coupling Waveforms

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

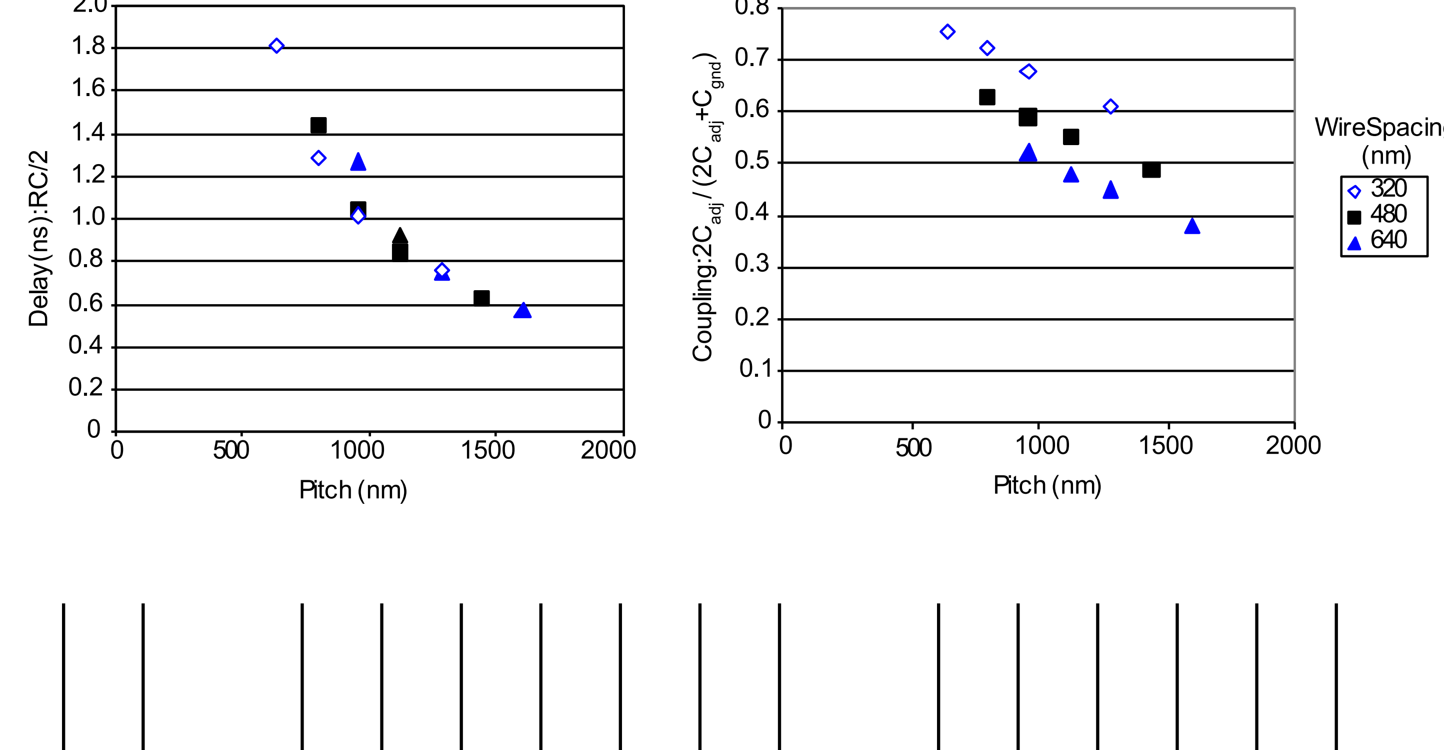


Noise Implications

- Do we care 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:
  - Width
  - Spacing
  - Layer
  - Shielding



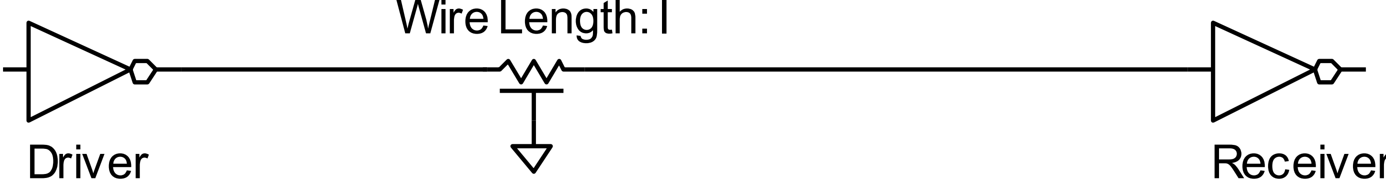
vdd a_0 a_1 gnd a_2 a_3 vdd

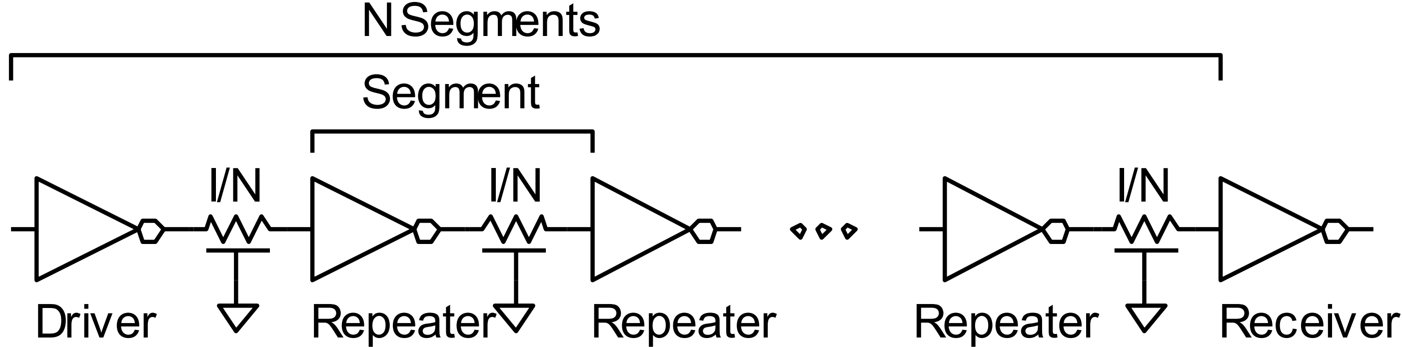
vdd a_0 gnd a_1 vdd a_2 gnd

a_0 b_0 a_1 b_1 a_2 b_2

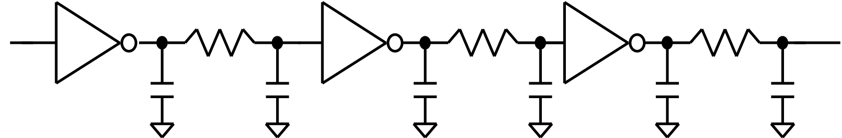
Repeaters

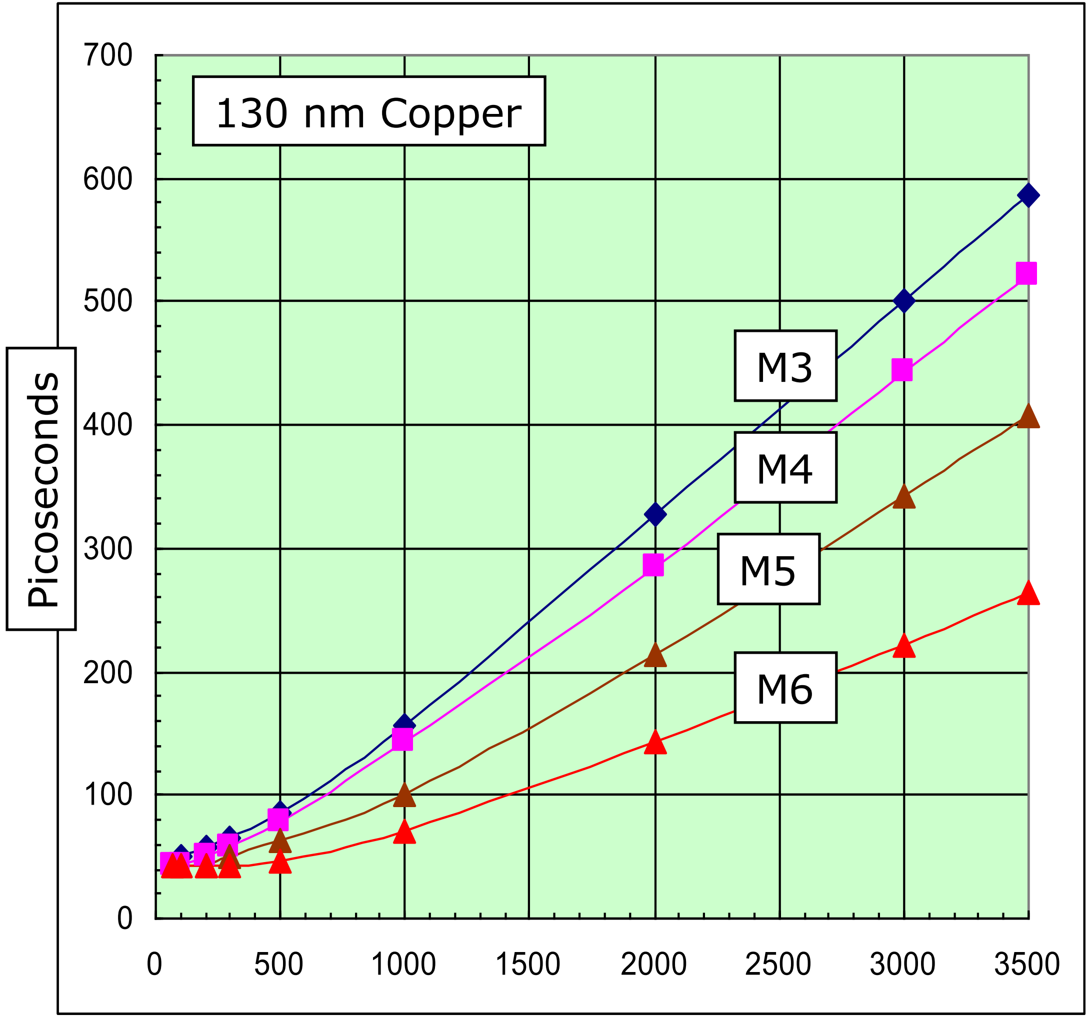
- R and C are proportional to *L*
- RC delay is proportional to L^2
  - Unacceptably great for long wires
- Break long wires into N shorter segments
  - Drive each one with an inverter or buffer





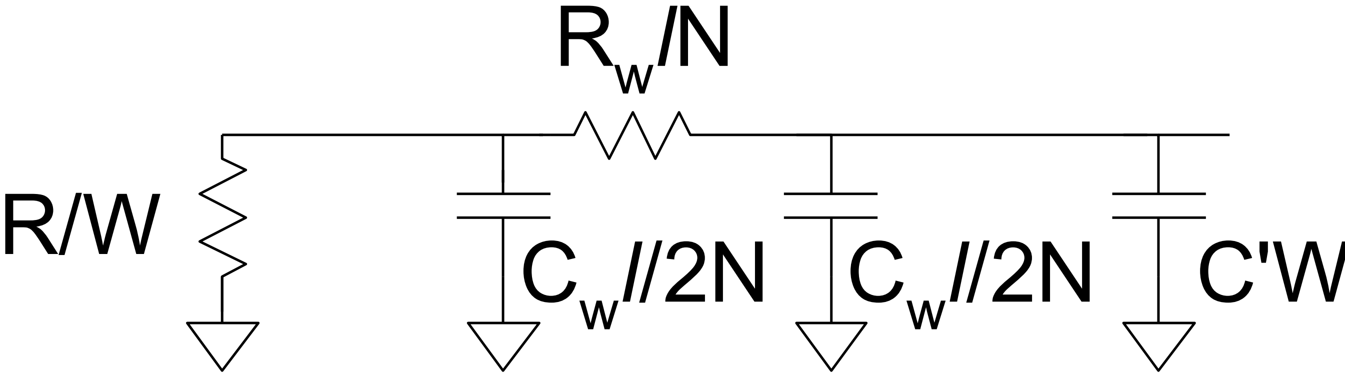
Repeated Interconnect





Repeater Design

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



Repeater Results

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

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

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:

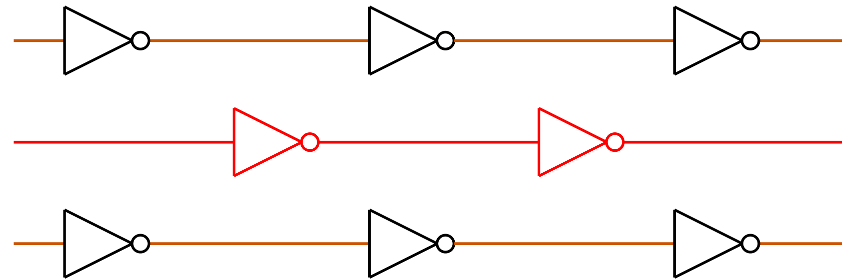
$$\frac{\partial t_{pd}}{\partial N} = RC' - l^2 \frac{R_w C_w}{2N^2} = 0 \Rightarrow N = l\sqrt{\frac{R_w C_w}{2RC'}}$$

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

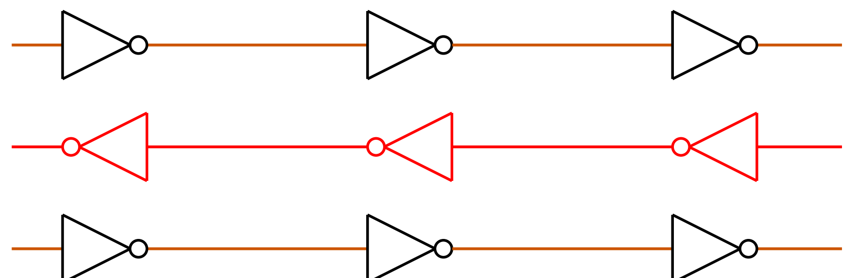
~60-80 ps/mm

in 180 nm process

Repeater Placements



Staggering the inverters



Avoiding the Miller cap by opposite going signals

BACKUP

Layer Stack

- AMI 0.6 μm process has 3 metal layers
- Modern processes use 6-10+ metal layers
- Example:

  Intel 180 nm process

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			

Substrate

- M1: thin, narrow (< 3λ)
  - High density cells
- M2-M4: thicker
  - For longer wires
- M5-M6: thickest
  - For V_{DD}, GND, clk