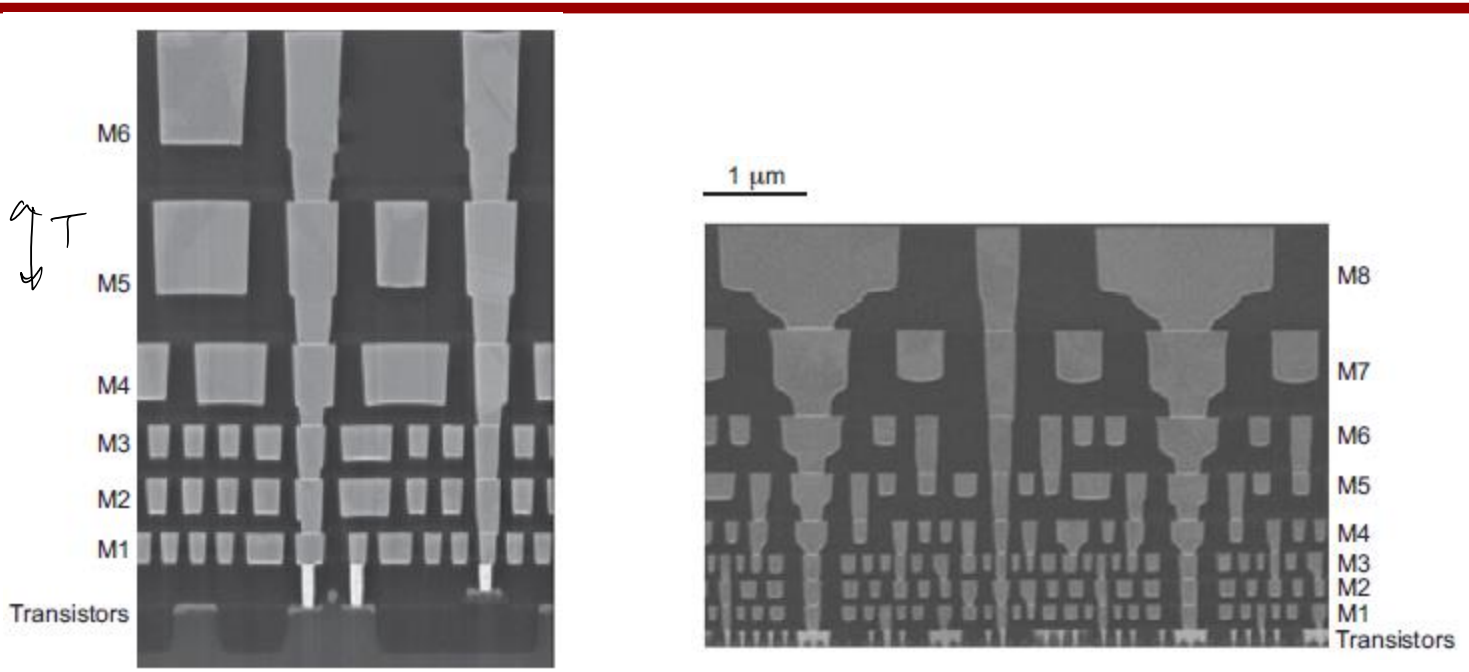
ELEC 402
Interconnects/Trends
Lecture 16

Reza Molavi
Dept. of ECE
University of British Columbia
reza@ece.ubc.ca

Slides Courtesy : Dr. H. Djahanshahi (Microsemi)

Interconnect is everything!



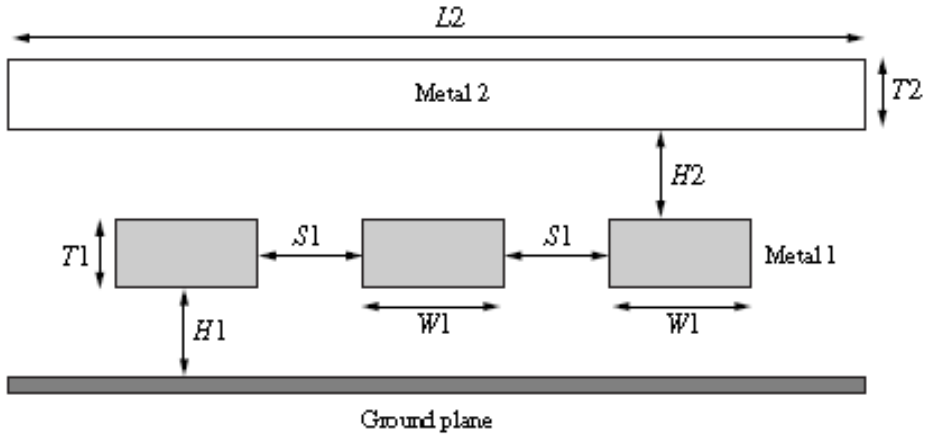
SEM image of wire cross-sections in Intel's 90 nm (left) and 45 nm processes (right)

Chip design is a 3D environment creating several parasitics elements

Todays technologies may have more than 10 metal layers with different thickness and propertie

They control all the important electrical connections on chip

Wire Dimensions



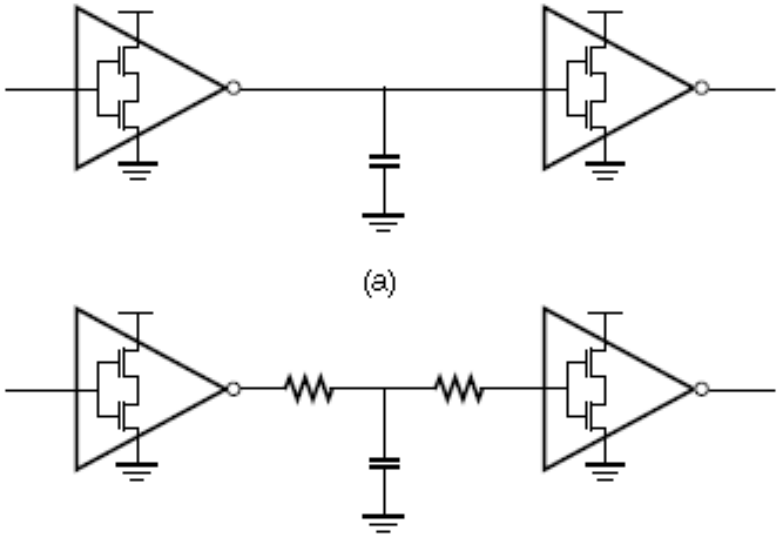
Use has no control over vertical dimensions (metal-metal separation (H), and each Layer thickness (T).

However, proper design of lateral dimensions (S1, and W1) is a full-time job

Sometimes these structures have to run several mm's alongside one another

Signal integrity team in each company plays an important role

Interconnect RC delay



Resistance of an interconnect is given by: $R = \frac{\rho L}{A} = \frac{\rho L}{TW}$

This term (called sheet resistance) is an important one determine $R_{sq} = \frac{\rho}{T}$

The quality of interconnect (note the similarity this term to R_{eqn} and R_{eqp} in transistors)

DSM Sheet Resistance

Sheet resistance of top metal layers are much smaller due to a couple of reasons

-They normally carry much higher current (tope level connection between blocks, and power Grid)

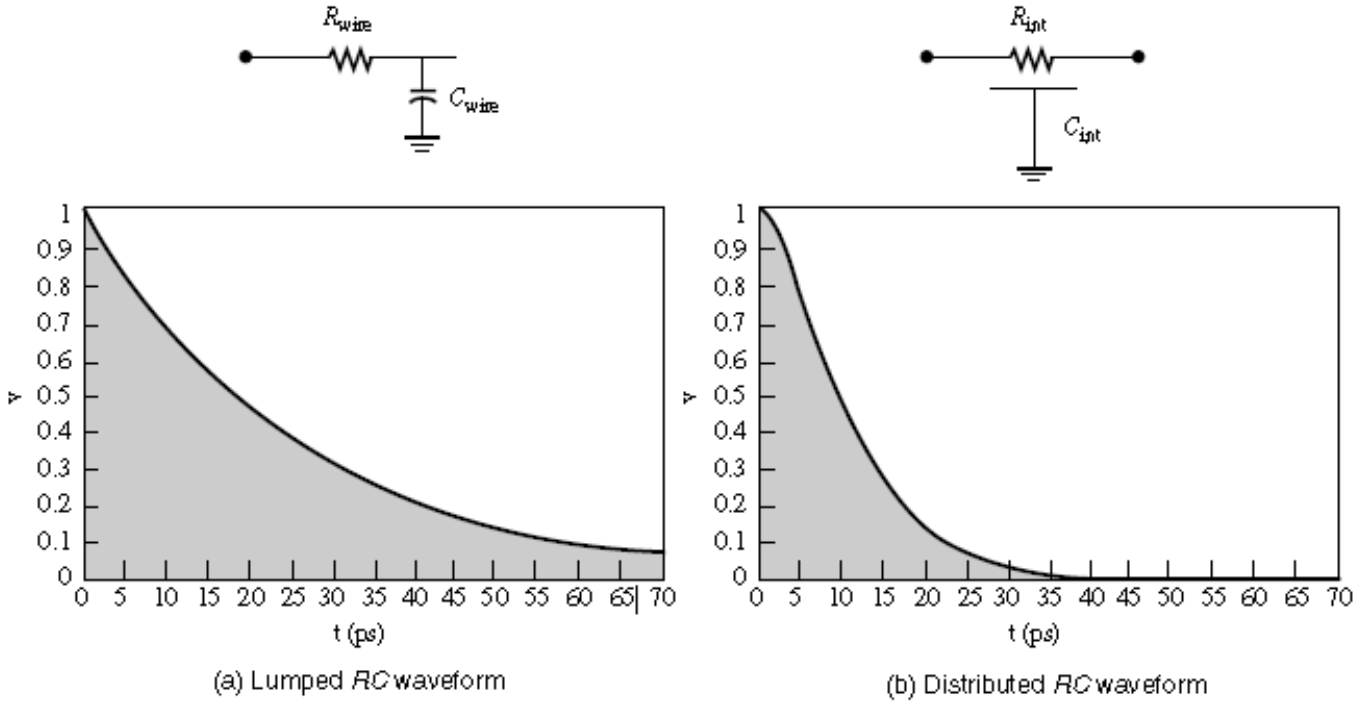
- Implementation of high-quality passive devices such as inductors

Example: Compute the sheet resistance of a 0.22 µm thick Cu wire in a 45 nm process. Find the Total resistance if the wire is 0.125 µm wide and 1 mm long. Ignore dishing

$$R_{\square} = \frac{2.2 * 10^{-8} \Omega . m}{0.22 * 10^{-6}\, m} = \ 0.1\ \Omega/\square$$

$$R = 0.1\Omega/\square \frac{1000\ \mu m}{0.125\ \mu m} = \ 800\ \Omega$$

Elmore Delay - Review



(a) Lumped *RC* waveform

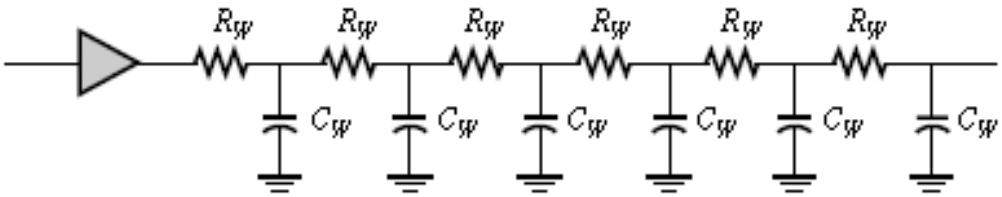
(b) Distributed *RC* waveform

When plotting the output voltage, the area under the curve is called the first moment of the curve

For simple lumped RC this moment is equal to the time constant

For distributed RC, we can still refer to this area as the time constant of the network

RC delay in long wires



Let's say the interconnect is very long length of L and it is divided into N smaller piece of Length ΔL each (n* ΔL= L)

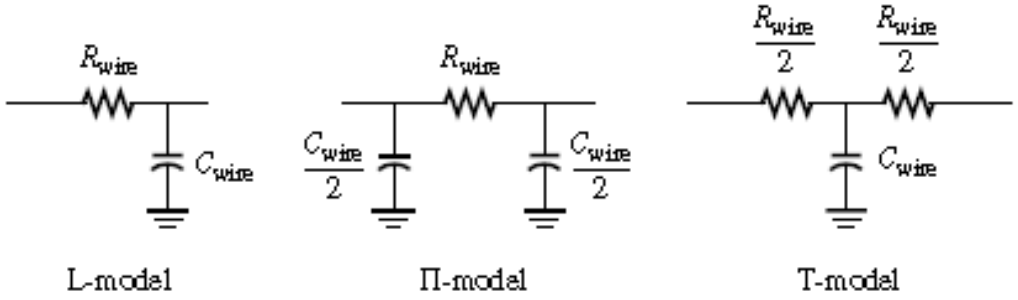
$$\tau_{\text{Elmore}} = (r\Delta L)(c\Delta L) + 2(r\Delta L)(c\Delta L) + \cdots + n(r\Delta L)(c\Delta L)$$

$$= (\Delta L)^2 rc(1 + 2 + \cdots + n)$$

$$= (\Delta L)^2 rc(n)(n+1)/2$$

$$\approx (\Delta L)^2 rcn^2/2 = L^2 rc/2 = R_{\text{wire}} C_{\text{wire}}/2$$

Equivalent Model

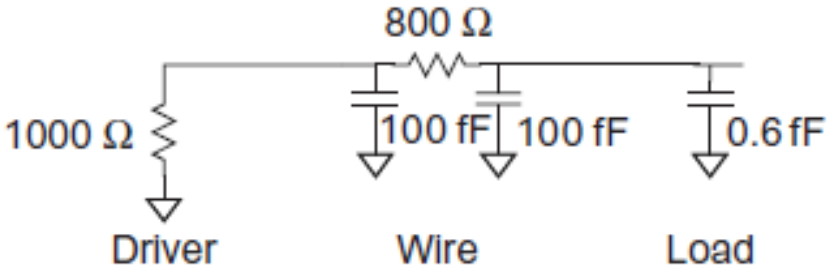


Both T and pi models represent the same delay as the Elmore of long channel

The pi one is less computing-expensive (one less node in the equivalent model)

Interconnect Example

A 10x unit-sized inverter drives a 2x inverter at the end of the 1 mm wire on a 45 nm CMOS chip. Suppose that wire capacitance is 0.2 fF/μm and that unit-sized nMOS transistor has R = 10 kΩ and C = 0.1 fF. Estimate the propagation delay using the Elmore delay model; neglect diffusion capacitance.

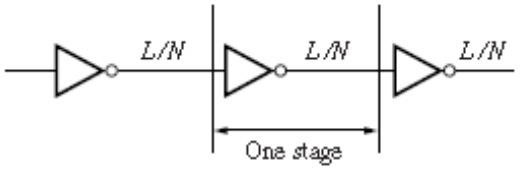


Solution:

The driver has a resistance of 1 kΩ.The receiver has a 2-unit nMOS transistor and a 4-unit pMOS transistor, for a capacitance of 0.6 fF. The wire capacitance is 200 fF.

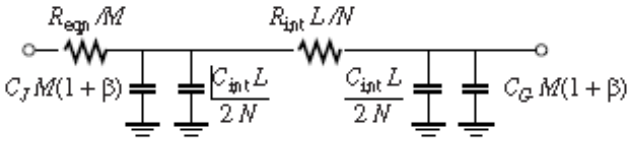
Above figure shows an equivalent circuit for the system using a single-segment π-model. The Elmore delay is t_{pd} *= (1000 Ω)(100 fF) + (1000 Ω + 800 Ω)(100 fF + 0.6 fF) = 281 ps. The capacitance of the long wire dominates the delay; the capacitance of the 2x inverter is negligible in comparison.*

Buffer (repeater) Insertion



(a) Wire of length L with N buffers inserted

(b) One segment of buffer and interconnected

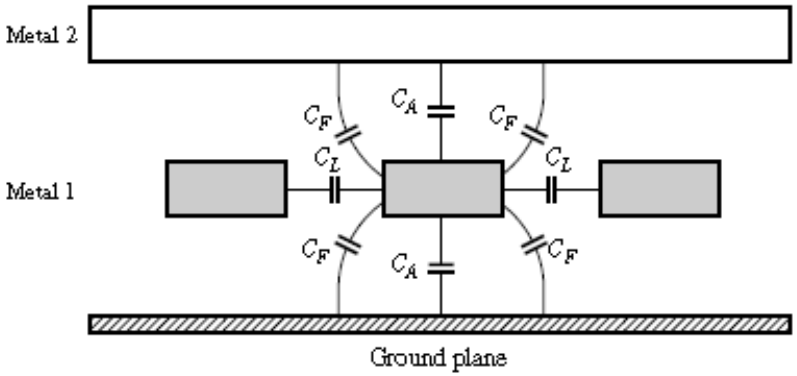


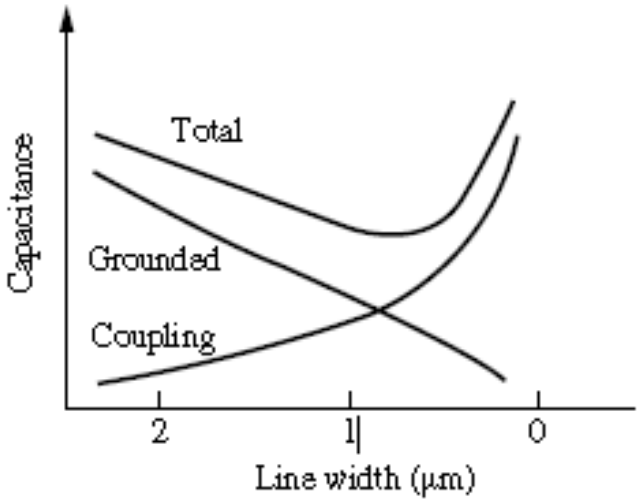
To avoid the exponentially growing delay of long interconnects we can insert equally-sized Buffers along the path (note the difference with the logical effort approach)

Let N be the number of inverters and M the optimal size of the buffer

The approach is again to minimize the total delay using these variable, exteremum problem

Capacitance of interconnects





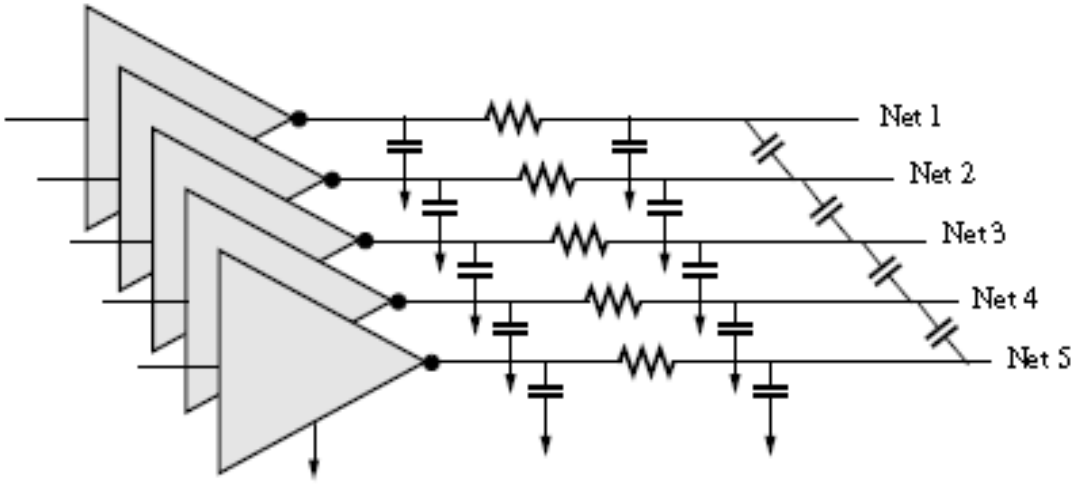
Due to scaling effects the ground capacitance is becoming less important

However, area and fringe capacitance effects are more pronounces in DSM technologies

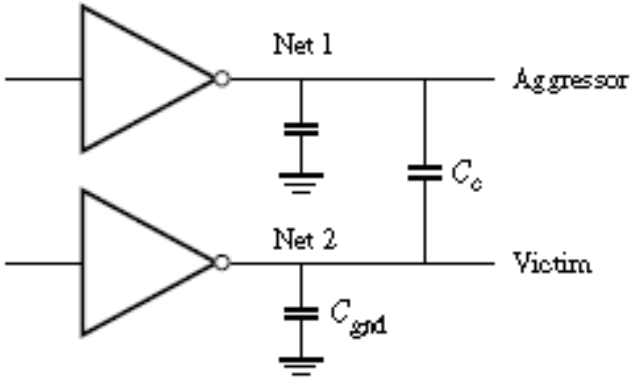
(these days more than 70% of total cap is due to coupling to adjacent nodes)

Capacitance Summary
(optional details in App slides)

For upper layers, when closely spaced the capacitance is maximum and is roughly 0.2fF/um, for other situations we can assume 0.1fF/um (lower metal layers or upper Metal layers when loosely coupled)



Coupling effect on Delay



When a line is switching (victim), the presence of another line (aggressor) in the vicinity

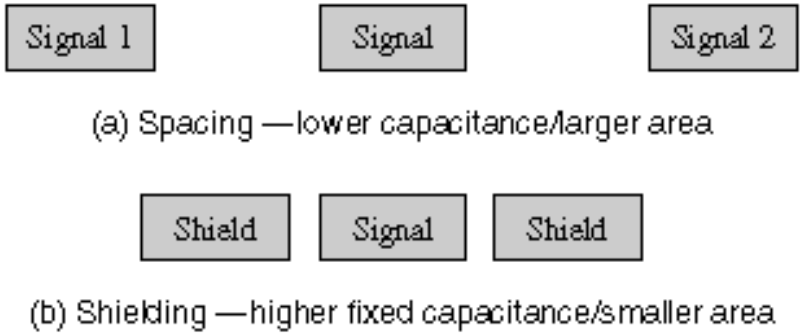
Can cause change to the total capacitance seen by the victim node

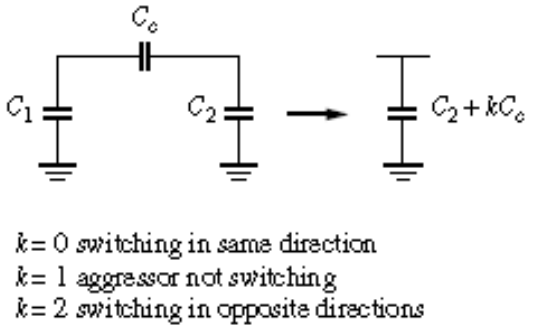
$$C_L = C_{gnd} + C_c \quad \text{aggressor not switching}$$

$$C_L = C_{gnd} \quad \text{both switching together}$$

$$C_L = C_{gnd} + 2C_c \quad \text{both switching opposite}$$

How To Deal with The issue?





Therefore if we have n nodes in the circuit there might be 3^n different scenarios for total cap

- We should space out neighboring signals as much as area allows

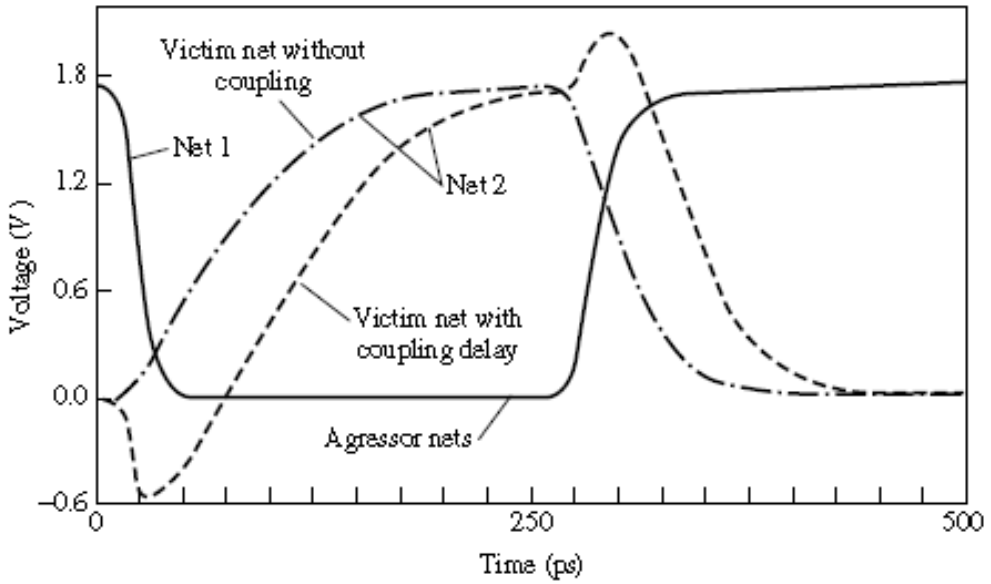
- Using ground shields in between differential signals can potentially reduce the capacitance

By a factor of 2 (increase in total area though)

- Using *low-k* materials

-Using timing tools based on worst-case capacitance (consider worst-case k and ground all nodes

Effect of Signal Feedthrough



An early timing in aggressor can potentially cause wrong direction for initial movement

Of the victim signal

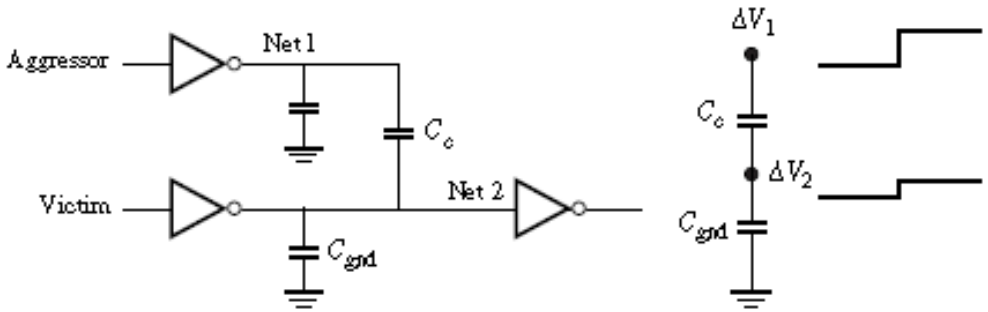
This can be included in worst-case k model (extend k range from -1 to 3 rather than 0 to 2)

Capacitive Noise or Cross-talk

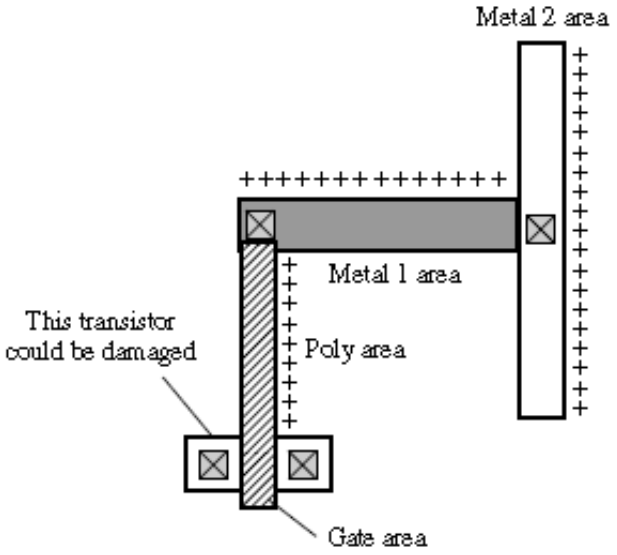
Even when the victim is supposed to be silent, changes in aggressor node my pose transient

Changes in the victim, also called *soft errors* (might be an issue for flip-flops and dynamic logic)

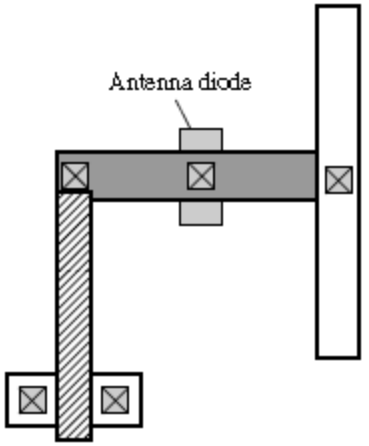
$$\Delta V_2 = \frac{C_c \Delta V_1}{(C_{gnd} + C_c)} \qquad (1$$

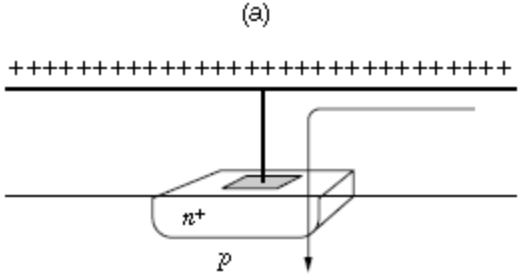


Antenna Effect



$$\text{Antenna Ratio} = \frac{\text{Area}_{\text{metal}}}{\text{Area}_{\text{gate}}}$$





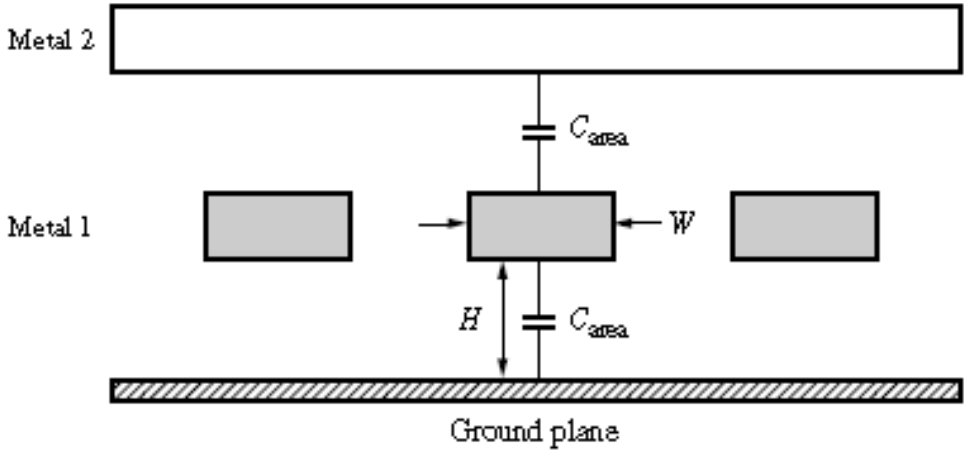
During CMP potentially large amount of charge can build up

One metal layers that may cause damage to gate-oxide

(voltage break-down) if the antenna ratio is too large (is detected by DRC deck)

Can be avoided by using buffer insertion, metal jumpers or diodes

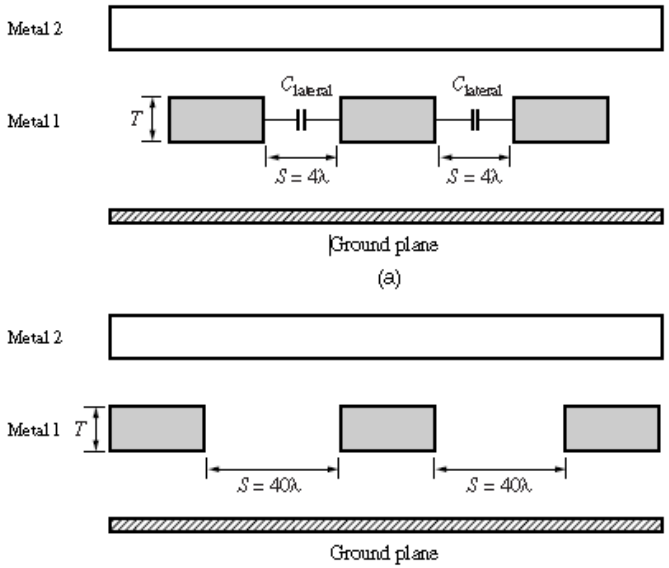
Appendix - The area Capacitances Calculation



$$C_{area} = \varepsilon_{ox}\frac{W}{H} = 4\varepsilon_0\frac{W}{H} = 4(88.5 \times 10^{-4}\ \text{fF}/\mu\text{m})\frac{W}{H} = 0.035\frac{W}{H}\ \text{fF}/\mu\text{m}$$

It is maximum when W is at largest (top metal layers)

Appendix - Lateral Capacitance Calculation

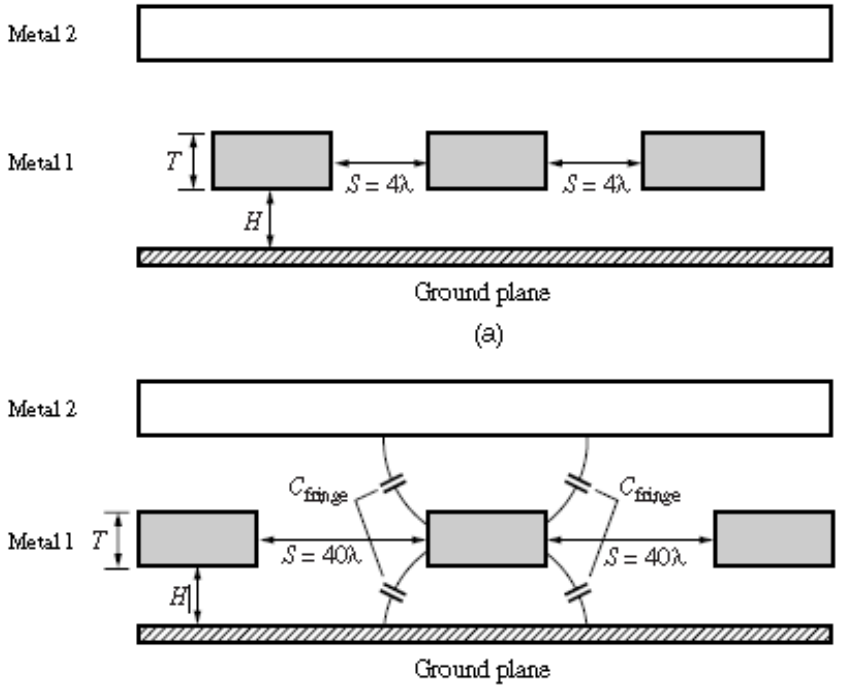


$$C_{\text{lateral}} = \varepsilon_{\text{ox}} \frac{T}{S}$$

$$C_{\text{lateral}} = \varepsilon_{\text{ox}} \frac{T}{S} = 4\varepsilon_0 \frac{T}{S} = 4(88.5 \times 10^{-4}\ \text{fF}/\mu\text{m}) \frac{T}{S} = 0.035 \frac{T}{S}\ \text{fF}/\mu\text{m}$$

For Tightly-coupled lines this might be 10times higher loosely couple ones

Appendix - Fringe Capacitance Calculations



$$C_{\text{fringe}} = \varepsilon_{ox} \ln\left(1 + \frac{T}{H}\right) = 4(88.5 \times 10^{-4}\ \text{fF}/\mu\text{m}) \ln\left(1 + \frac{0.5}{0.5}\right)$$

$$= 0.025\ \text{fF}/\mu\text{m}$$