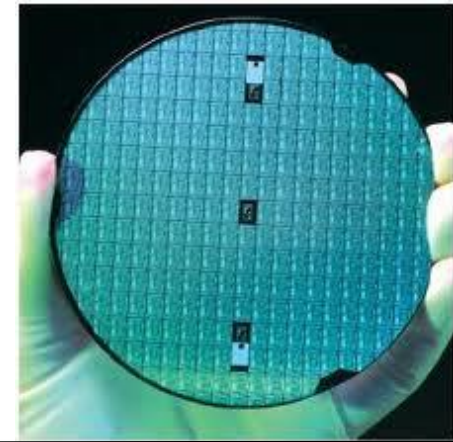
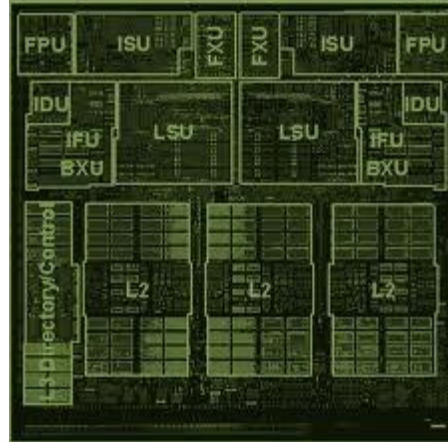
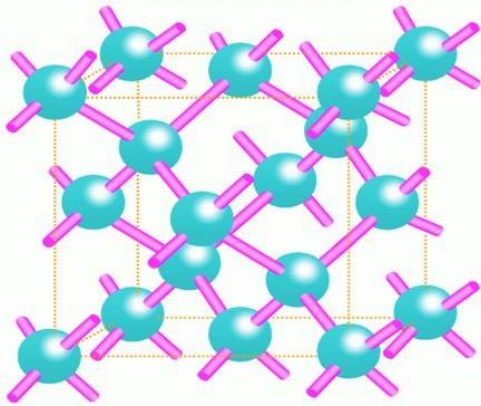


Structure of silicon crystal



ECE 225

Lecture 6

VLSI Interconnects-II

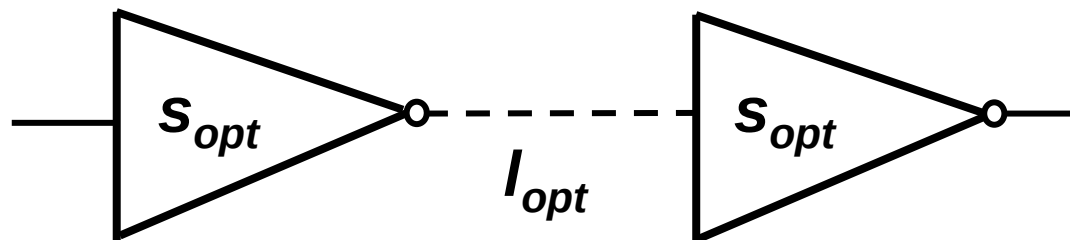
Prof. Kaustav Banerjee
Electrical and Computer Engineering
University of California, Santa Barbara

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

Repeater Insertion

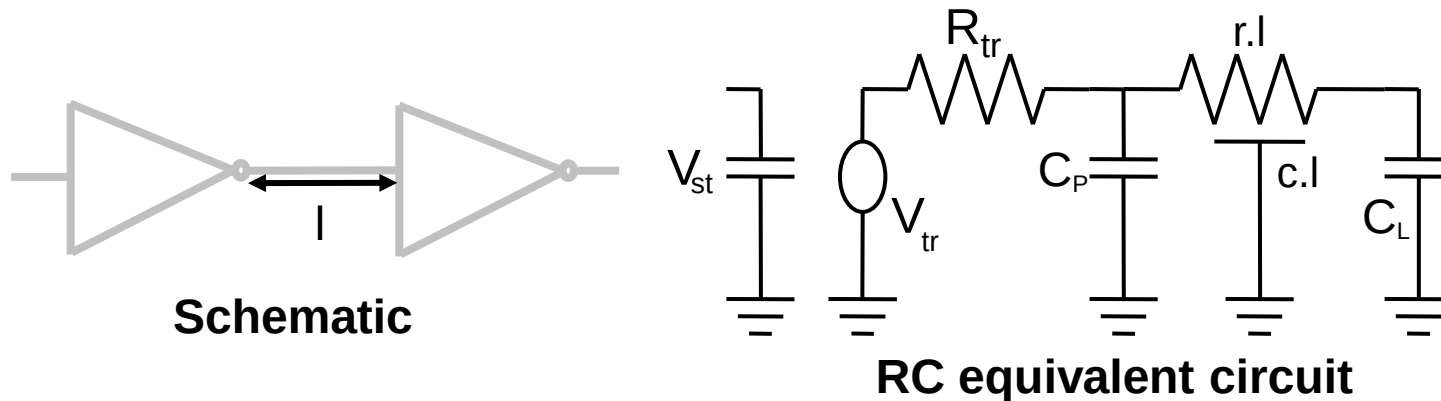
- Delay of wire increases quadratically with its length, L (delay = RC ...if unbuffered)
- **Repeater** (inverter) or **buffer** insertion results in lower delay increase....(delay = $t_{inv} + RC/2$...if one inverter is inserted at $L/2$)
- Repeater size and inter-repeater separation can be optimized to give minimal delay



(from [Bakoglu89])

Driver-Interconnect-Load Equivalent Circuit

- Interconnect segment of length l between two inverters



- V_{st} is the voltage at the input capacitance that controls the voltage source V_{tr}
- R_{tr} is the driver transistor resistance, C_p is output parasitic capacitance
- C_L is the load capacitance of next stage, r and c are interconnect resistance and capacitance per unit length respectively

Interconnect Scaling

Interconnect Scaling Scenarios

- ❑ Performance of a technology can be summarized by the time delay, t_d :

$$t_d = R_{tr} (C_P + C_L) + (R_{tr} c + r C_L) l + \frac{1}{2} r c l^2$$

intrinsic stage delay + load delay + interconnect delay

- ❑ If s is a scaling parameter: ratio of the feature size at a newer technology node to the feature size at an older node, then $s < 1$
- ❑ Two simple scaling scenarios
 - ❑ scale metal pitch at constant thickness
 - ❑ scale metal pitch and the metal thickness

Interconnect Scaling

Implications of Technology Scaling on Technology Performance

- ❑ Scaling metal pitch at constant thickness
 - ❑ load delay, line delay, and current density will scale as $1/s$, $1/s^2$, and $1/s$ respectively.
- ❑ Scaling metal pitch and the metal thickness
 - ❑ load delay, line delay, and current density will scale as $1/s^2$, $1/s^2$, and $1/s^2$ respectively.
- ❑ Hence, interconnect delay begin to dominate technology performance and current density increases with scaling

Interconnect Technology Scaling Trends

- ❑ Interconnect scaling requirements drive several technology enhancements:
 - ❑ use of low-dielectric constant (**low-k**) materials lowers the interconnect capacitance per unit length (c) and therefore reduces the delay time
 - ❑ lower c helps in minimizing **cross-talk** noise
 - ❑ lower c also helps to reduce the **dynamic power** dissipation during switching of gates
 - ❑ lower interconnect resistance and higher current density requirements have driven the use of **new metallization**, such as Cu

VLSI Interconnects

- ❑ Scaling of interconnects makes them slower, unlike scaling of devices
- ❑ **Local interconnects:** metal levels M1-M2, very densely packed (congested), follow transistor scaling, short distances (few μm)
- ❑ **Global interconnects:** top 3-4 metal levels (M7-M10) fewer such wires, more room for wide wires (not scaled aggressively), long distances (100s of μm or few mms)
- ❑ As devices switch faster, signal propagation delay on interconnects becomes appreciable fraction of intrinsic gate delay (at local/semi-global level) or much larger than gate delay (at global level)

Local Interconnects

- ❑ Delay is small (but increasing) fraction of intrinsic gate delay
 - R-C models suffice
 - Capacitance more important than resistance
- ❑ Nanometer scale effects – scattering at surfaces and grain boundaries makes resistivity increase with scaling dimensions
- ❑ Small cross-section, high current density, electromigration reliability big problem
 - Fundamental limitation of material
 - Carbon nanomaterials can circumvent this limitation

Global Interconnects

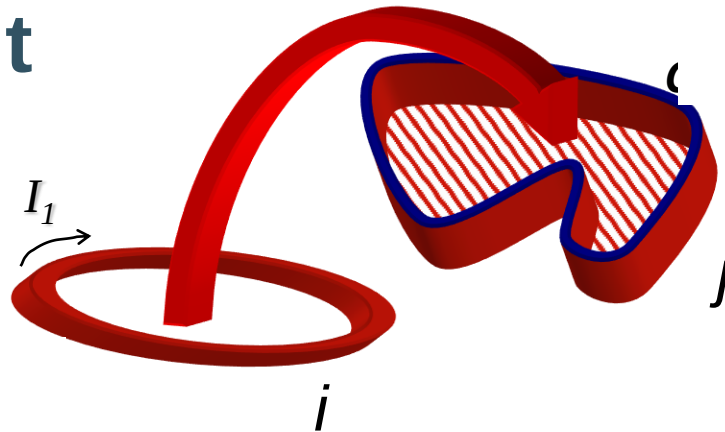
- ❑ Delay is a major challenge – limits speed of critical signal transmission like clock
- ❑ Simple R-C models inaccurate
- ❑ Beyond 100um lengths, few GHz frequencies, R becomes comparable to ωL
- ❑ Electromagnetic interactions beyond immediate neighbors become important
- ❑ Impedance (R, L) frequency dependent (skin and proximity effects): no formulae; field solvers or heuristic calculations

Global Interconnects

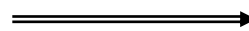
- ❑ Thousands of large repeaters on-chip to control delay: power dissipation (clock network ~ 60-70% of total chip power)
- ❑ **Power-optimal repeater insertion** to save power on non-critical signal lines
- ❑ **3-D ICs**: short (~50um) vertical interconnects replace long (few mm) horizontal interconnects
- ❑ Carbon nanomaterials (CNTs, GNRs) based interconnects expected to have better performance and power dissipation than Cu

Definition of Inductance (L)

L_{ij} = geometrical constant of proportionality between the magnetic flux through victim loop j and the current running on aggressor loop i causing it



$$L_{ij} = \frac{\Psi_{i \rightarrow j}}{I_i} = \frac{1}{I_i} \int_{S_j} \mathbf{B}^{(i)} \cdot d\mathbf{S}_j$$



$$L_{ij} = \frac{\mu_0}{4\pi} \oint_{C_j} \oint_{C_i} \frac{d\mathbf{\ell}_j \cdot d\mathbf{\ell}_i}{|\mathbf{r}_j - \mathbf{r}_i|}$$

On-Chip Inductance

- ❑ L starts to play a role at high GHz frequencies
- ❑ Effect increases for Cu: lower resistivity
- ❑ Can increase interconnect delay
- ❑ Can cause overshoot and undershoot effects
- ❑ Can cause switching noise: $L di/dt$ voltage drops

$$Z = R + j\omega L$$

inductance of a wire surrounded by a uniform dielectric:

$$c \ell = \epsilon \mu$$

c = cap. per unit length

l = inductance per unit length

From Maxwell's Laws:

$$v = \frac{1}{\sqrt{\ell c}} = \frac{1}{\sqrt{\epsilon \mu}} = \frac{c_0}{\sqrt{\epsilon_r \mu_r}}$$

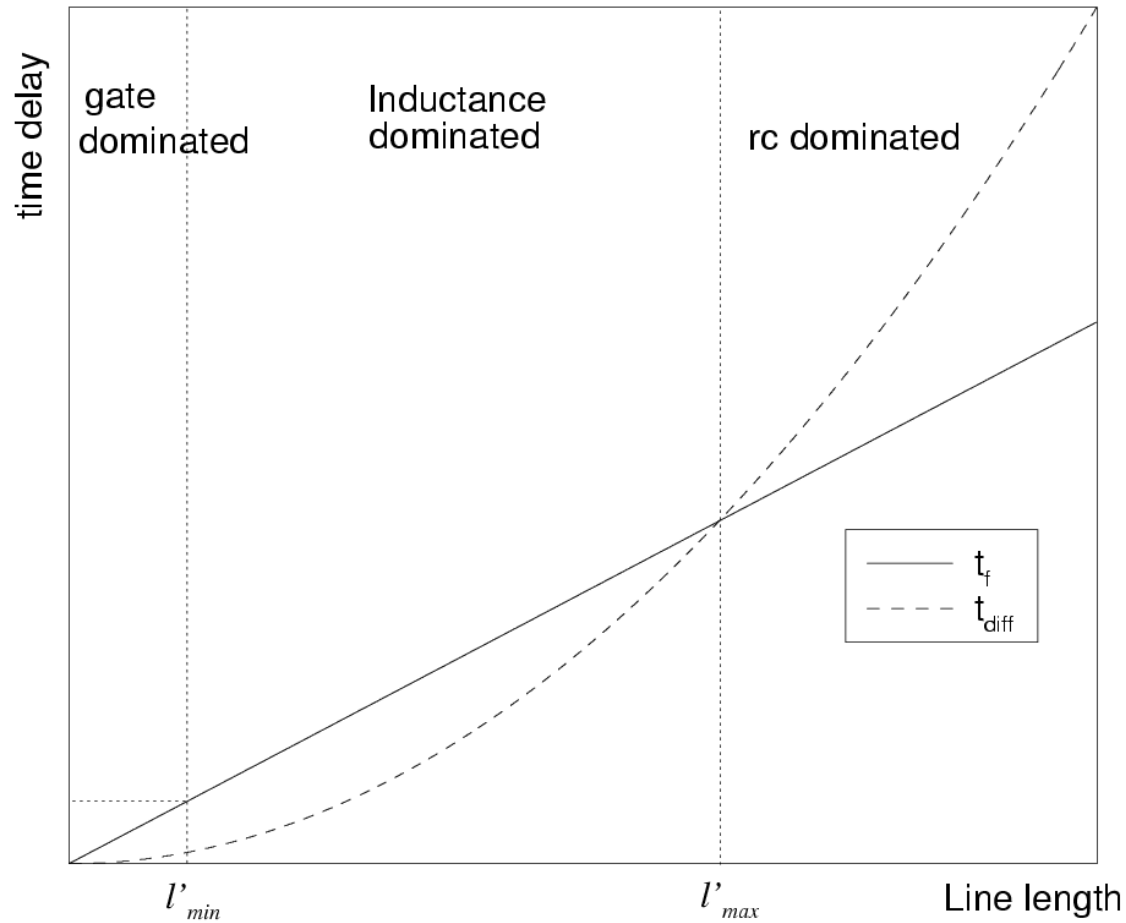
v = propagation speed of EM wave

c₀ = speed of light in vacuum = 30 cm/ns

Accurate Interconnect Inductance Extraction

- ❑ Requires full 3-D electromagnetic simulation
- ❑ FASTHENRY (from MIT) is a widely used tool in academia
 - See M. Kamon, M. J. Tsuk, and J. K. White, “FASTHENRY: A multipole-accelerated 3-D inductance extraction program,” IEEE Trans. Microwave Theory Techn., vol. 42, pp. 1750–1758, Sept. 1994.
 - Can be time consuming for complex interconnect topologies
 - Has limitations at very high frequencies

Wire Length Dependency of L Effects



$$l'_{min} = \frac{T_{rise} v}{2}$$

$$l'_{max} = \frac{2Z_0}{r}$$

Suaya et al., Advanced Metallization Conf. 2006

RLC Delay: Design Rules of Thumb

Wires in RLC domain:

$$t_{50\%} = \sqrt{LC} + \frac{\tau_{FO4}}{2} + 0.7Z_0C_{load} \quad \text{if}$$

$$\frac{rL}{Z_0} < 2\ln\left(\frac{4Z_0}{R_{driver} + Z_0}\right), \quad R_{driver} < 3Z_0$$

$$\text{where } Z_0 = \sqrt{l/c}$$

Also see: A. Deutsch, et al., “When are transmission-line effects important for on-chip interconnections?,” IEEE Trans. Microwave Theory Tech., vol. 45, pp. 1836–1846, Oct. 1997.

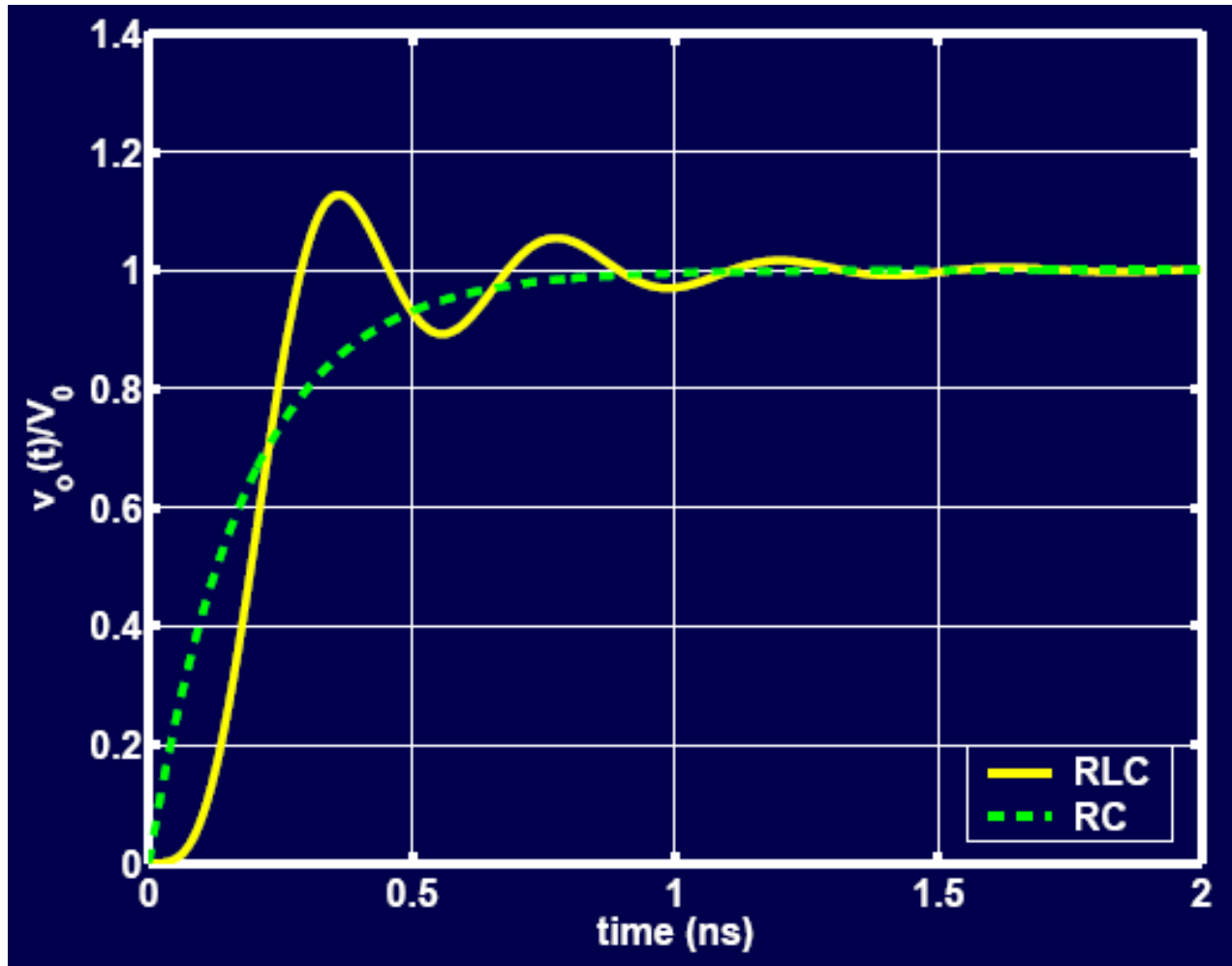
On-Chip Inductance Effects

K. Banerjee and A. Mehrotra, IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, Vol. 21, No. 8, pp. 904-915, August 2002.

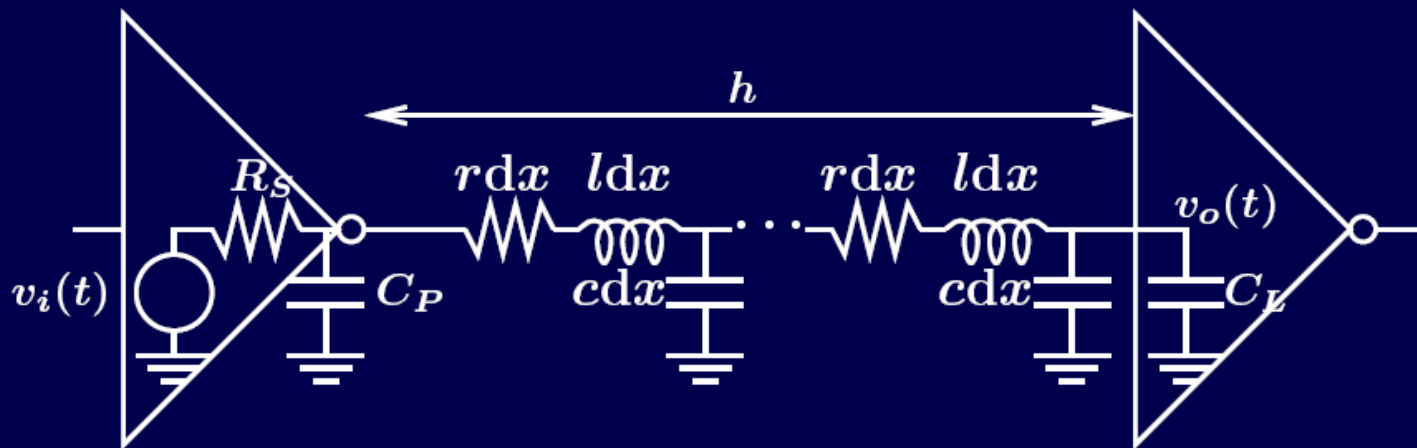
On-Chip Inductance Effects

- *Increasing clock speeds and decreasing rise times cause inductive effects*
- *Low resistance wires more susceptible to inductive effects*
 - *Global wires that are usually wide*
 - *introduction of Cu: lower resistivity*
- *Line inductance depends on current return path*
 - *strongly dependent on circuit topology*
 - *difficult to extract accurately*
 - *large variations*

Step Response of an RLC Line



Time-Domain Response of RLC Segment



$$H(s) = \frac{1}{[1 + sR_S(C_P + C_L)] \cosh(\theta) + \left[\frac{R_S}{Z_0} + sC_L Z_0 + s^2 R_S C_P C_L Z_0 \right] \sinh(\theta)}$$

$$\approx \frac{1}{1 + sb_1 + s^2 b_2 + s^3 b_3 + s^4 b_4}$$

$$Z_0 = \sqrt{\frac{r + sl}{sc}} \quad \theta = \sqrt{(r + sl)sc}h$$

Critical Inductance

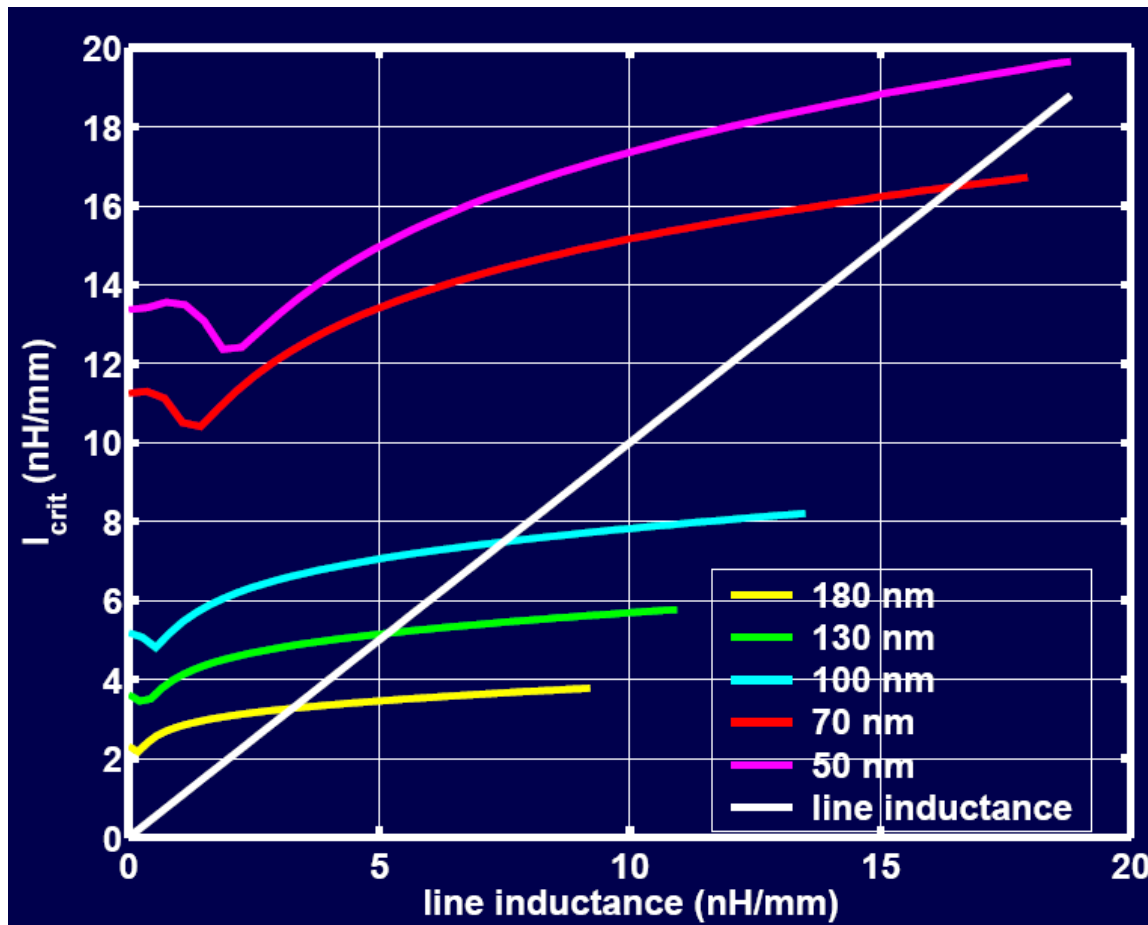
- Second order Padé expansion of $H(s)$

$$H(s) \approx \frac{1}{1 + b_1 s + b_2 s^2}$$

- System overdamped, critically damped or underdamped when $b_1^2 - 4b_2$ is $>, =, < 0$.
- At optimum buffer size and interconnect length, find l_{crit} for which the system is critically damped.
- For $l <, =, > l_{crit}$, the system is overdamped, critically damped or underdamped

Impact of Scaling: Case 1

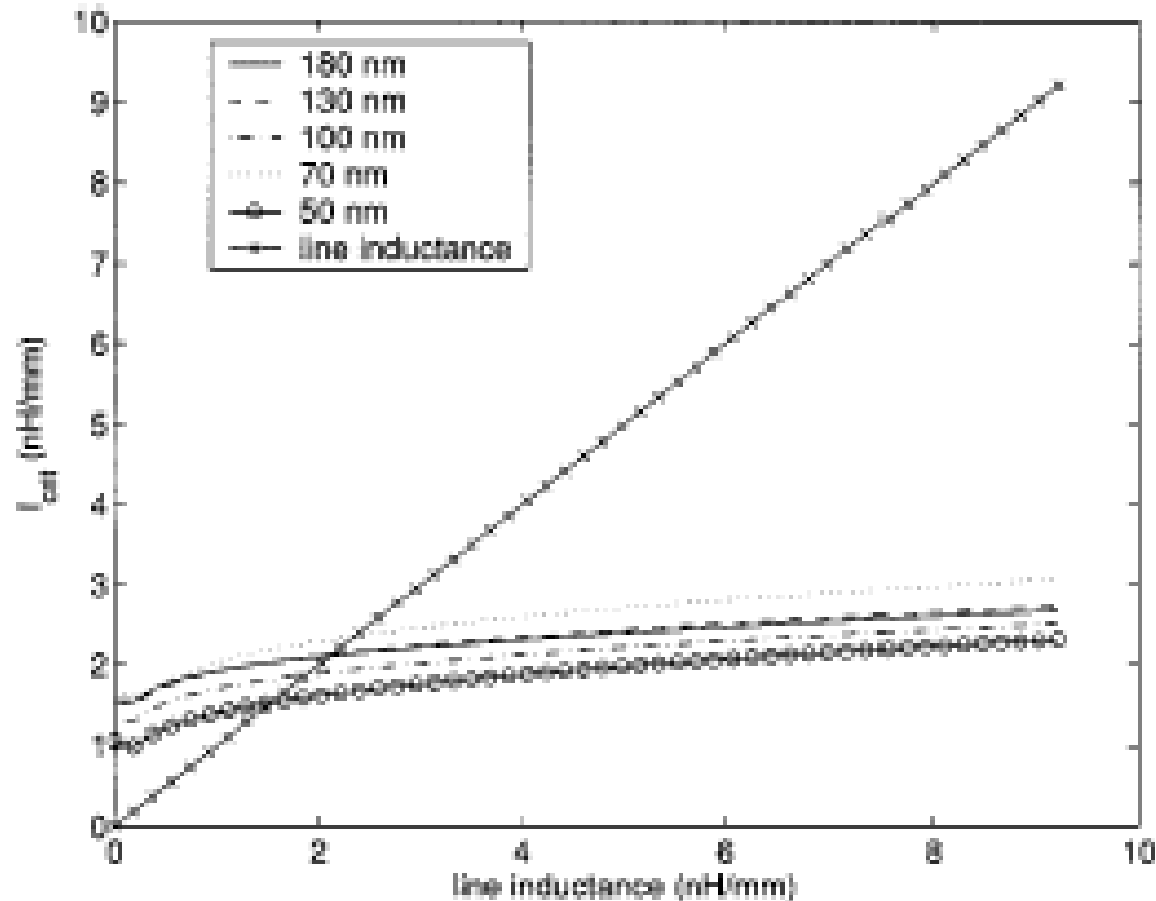
Minimum Width Global Wires



- As technology scales, $I < I_{crit}$ over larger range of line inductance
- Hence, impact of inductance **reduces** with scaling

Impact of scaling: Case 2

Unscaled Global Wires



- As technology scales, $I > I_{crit}$ over larger range of line inductance
- Hence, impact of inductance **increases** with scaling

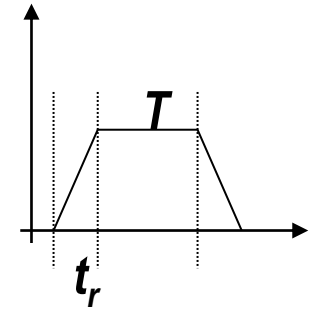
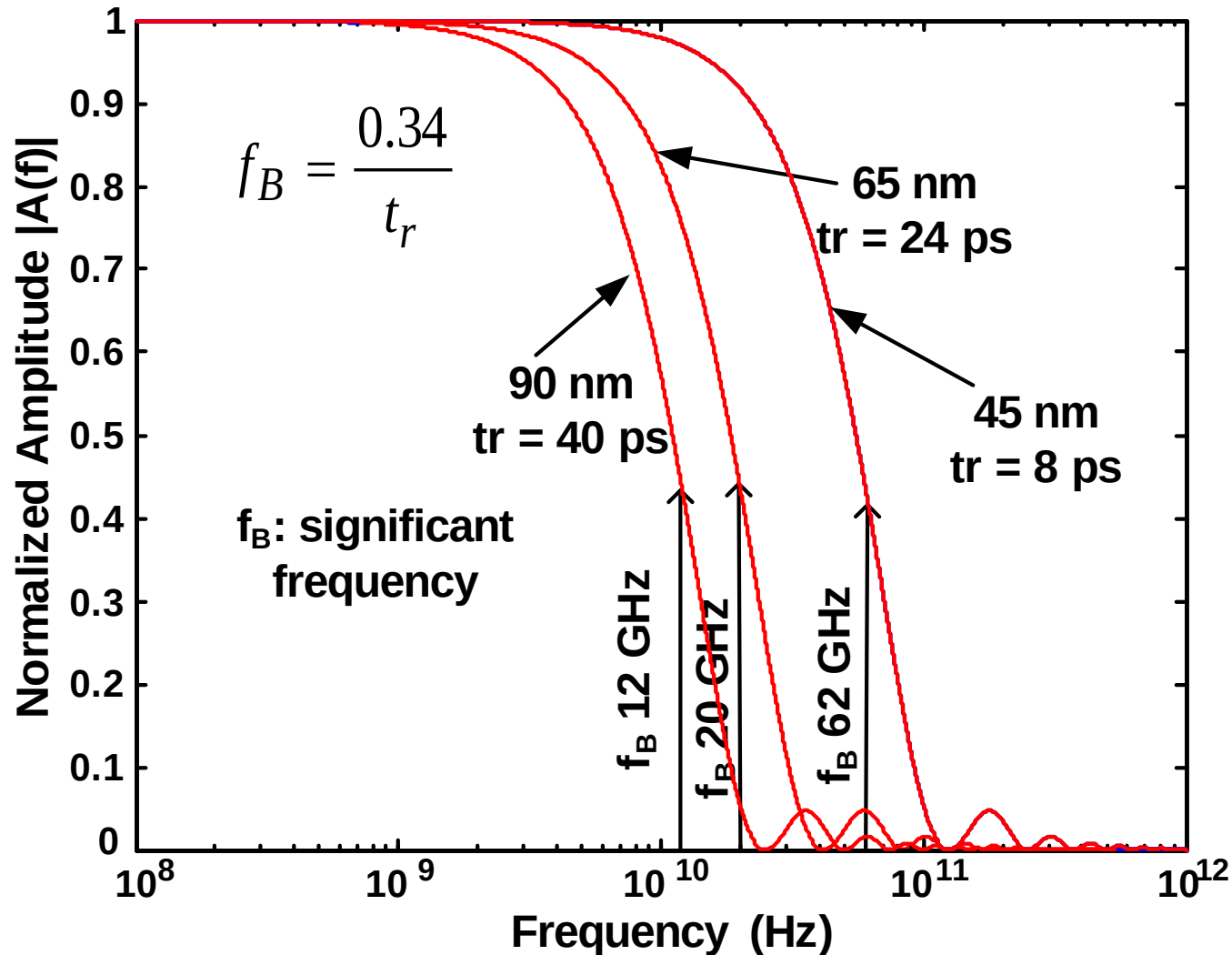
High-Frequency Issues

*K. Banerjee, S. Im and N. Srivastava, Proc. Advanced Metallization Conf., 2005.
S. Suaya, R. Escovar, S. Ortiz, K. Banerjee and N. Srivastava, Proc. Advanced
Metallization Conf., 2006.*

High-Frequency Effects

- Frequency of signals on interconnects is rising rapidly
- Frequency dependent impedance extraction is a major hurdle
 - Field solvers are unable to handle the complexity of VLSI interconnects
- Need to develop models for interconnect geometry dependent calculation of high frequency impedance

Signal Frequency on Interconnects



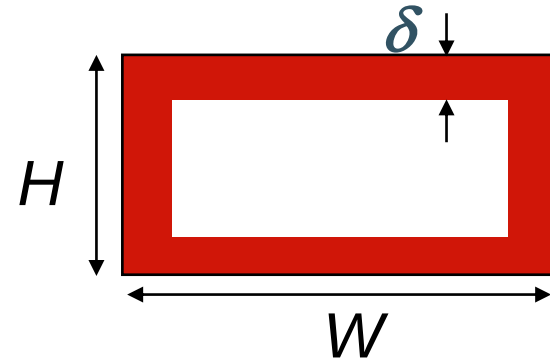
Digital signal composed of various frequency components.

Significant Frequency:

Frequency used for interconnect analysis

Frequency Dependence of R

- At high enough frequencies, R can increase due to **skin effect**
- Current flow constrained in a thin layer along the surface of the conductor: **R increases**
- The current falls off exponentially with depth into the interconnect
- **Skin depth**, δ is defined as the depth at which current falls off to a value of $1/e$ of its nominal value



$$\delta = \sqrt{\frac{\rho}{\pi f \mu}}$$

μ is the permeability of the surrounding dielectric

Resistance per unit length:

$$r(f) = \frac{\sqrt{\pi f \mu \rho}}{2(H+W)}, \quad \text{for } W, H \gg \delta$$

Skin Effect

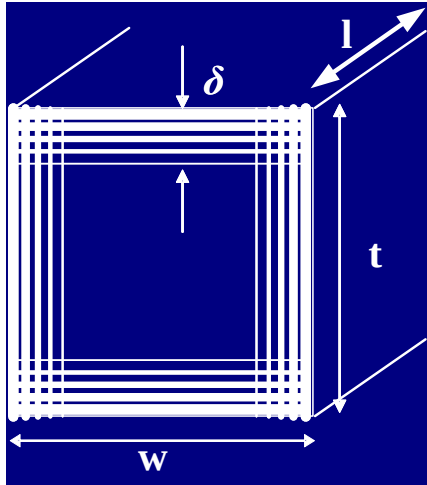
- Onset of skin effect:
at $f=f_s$, $\delta = 1/2 \times$
smallest dimension

$$f_s = \frac{4\rho}{\pi\mu(\min(W,H))^2}$$

- For Cu wires:

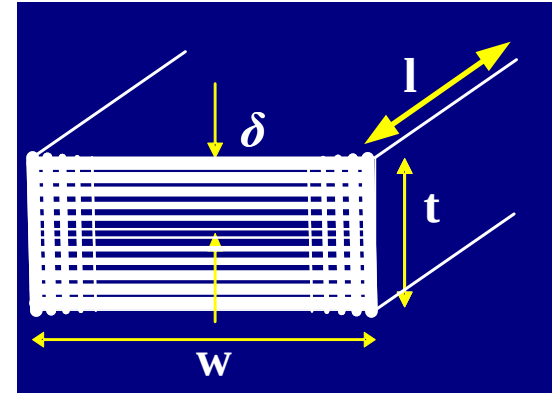
Freq	1 GHZ	5 GHZ	10 GHZ	15 GHZ	20 GHZ
Skin depth (μm)	2.08	0.93	0.66	0.53	0.46

Skin and Proximity Effects



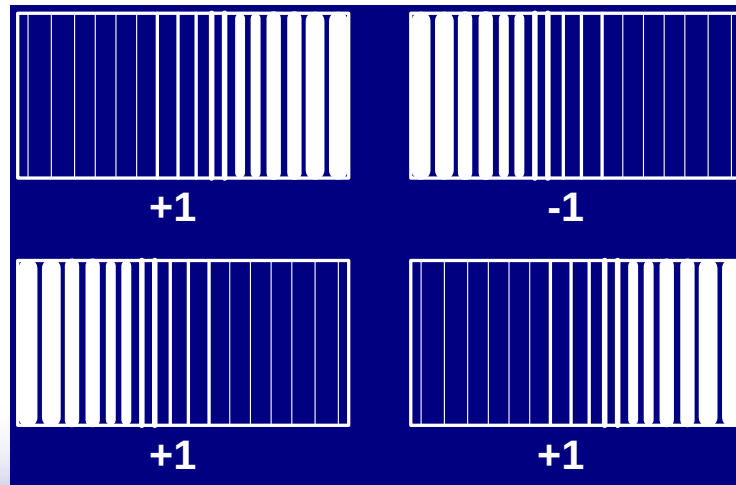
Isolated Wire $5\mu\text{m} \times 5\mu\text{m}$

Skin Effect in Isolated Conductors

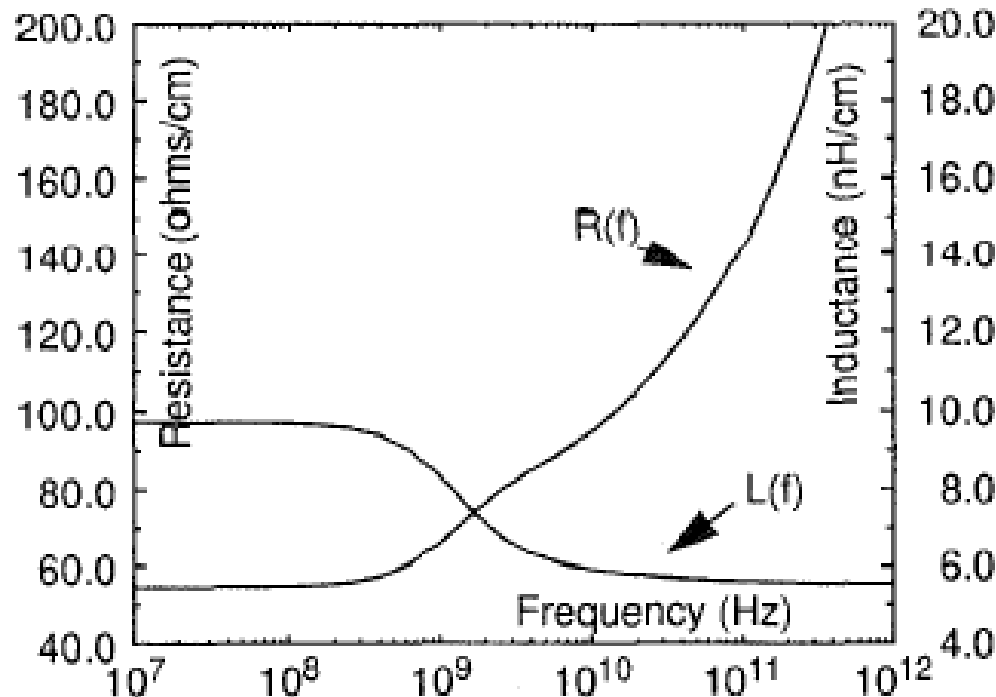


Isolated Wire $5\mu\text{m} \times 1\mu\text{m}$

Proximity Effect in Coupled Conductors



Frequency Dependence of R and L

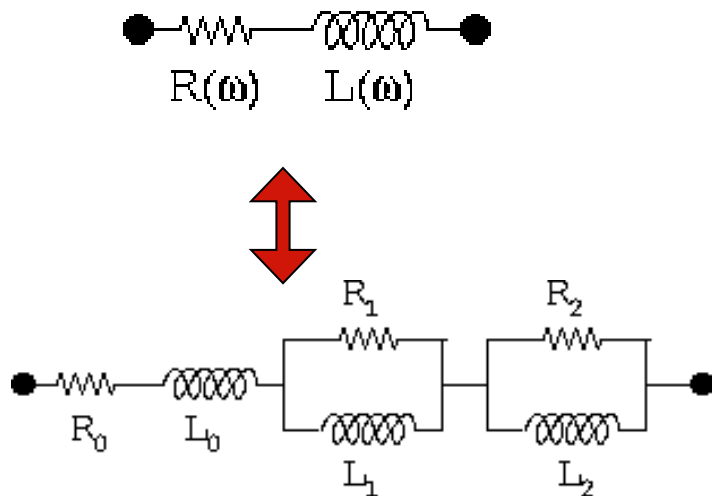


B. Krauter and S. Mehrotra, DAC, 1998.

- Resistance increases due to skin effect
- For lower frequencies, current return path is distributed between all ground lines---thus **loop inductance** is higher and return path resistance is smaller.....
- For higher frequencies, current returns through closest ground paths---thus loop area is smaller, hence loop inductance reduces and return path resistance increases

Modeling Frequency Dependent Behavior

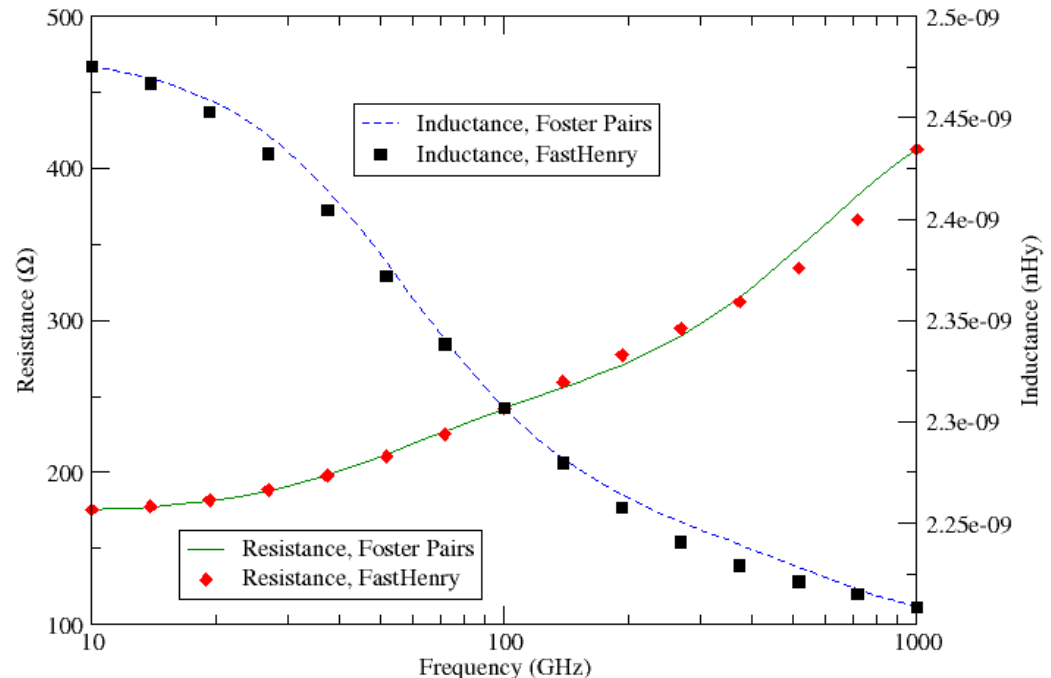
Replacing $R(\omega), L(\omega)$ with Foster Pairs



Note that at smaller frequencies L_1 ($Z=j\omega L$) will short R_1

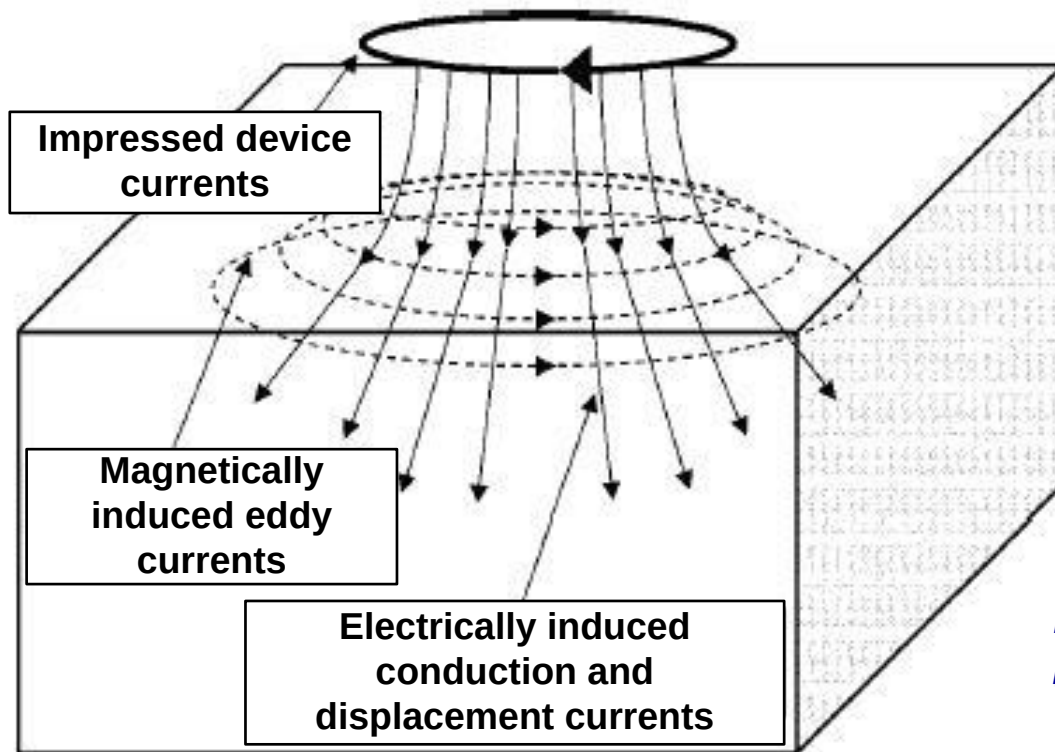
R_0 and $L_0+L_1+L_2$ are low-frequency values of R and L

At higher frequencies some current flows through R_1 and R_2 , hence R increases with frequency and effective inductance decreases



Suaya et al., Advanced Metallization Conf. 2006

Substrate Effects



Impacts interconnect impedance extraction

See, *N. Srivastava, R. Suaya and K. Banerjee, IEEE TCAD, Vol. 28, No. 7, pp. 1047-1060, July 2009.* ----Interconnect/inductor impedance extraction methodology in presence of multilayered conducting substrates – a **core technology in CALIBRE xL, xACT, tools by Mentor Graphics.**