Integrated Circuits: The Problem with Wires (and Some Solutions)

Paul A. Kohl
Todd Spencer
Tyler Osborn

School of Chemical and Biomolecular Engineering
Director, Interconnect Focus Center
Georgia Institute of Technology
Atlanta, GA 30332-0100
kohl@gatech.edu
404-894-2893
Strategy Based on Maintaining Historic Cost Improvements

- Logarithmic $ per Function
  - Selling Price
  - Mfg. Cost
  - Value

25-30% per Year Improvement

Past  Future
**Four Parameters (L, W, x_j, t_{ox})**

- **Source**
- **Gate**
- **Drain**

Diagram showing:
- **Metal**
- **SiO_2**
- **Silicon**
- **x_j**
- **L**
- **W**
- **t_{ox}**
## Scaling Parameters for CMOS

<table>
<thead>
<tr>
<th>Parameter</th>
<th>S Factor ( S \approx 1.15 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( W, L, T_{ox}, X_j )</td>
<td>( 1/S ) 87%</td>
</tr>
</tbody>
</table>

\( W, L, T_{ox}, X_j \) are scaling parameters for CMOS technology where \( S \) is a factor that scales with decreasing feature sizes.
## Scaling Parameters for CMOS

<table>
<thead>
<tr>
<th>Parameter</th>
<th>S Factor S~1.15</th>
</tr>
</thead>
<tbody>
<tr>
<td>* W, L, T(_{ox}), X(_j)*</td>
<td>* 1/S 87%</td>
</tr>
<tr>
<td>* Voltage (constant field)</td>
<td>* 1/S 87%</td>
</tr>
</tbody>
</table>
Scaling Parameters for CMOS

* **Parameter**
  * W, L, T_{ox}, X_{j}
  * Voltage (constant field)
  * Capacitance

* **S Factor S ~ 1.15**
  * 1/S 87%

* **1/S 87%**

\[
C_{ox} = \varepsilon \frac{W}{L} \frac{1}{T_{ox}} = \frac{1}{S} \frac{1}{S} = \frac{1}{S}
\]

* **1/S 87%**
Scaling Parameters for CMOS

* **Parameter**
  * W, L, T\(_{\text{ox}}\), X\(_{j}\)
  * Voltage (constant field)
  * Capacitance
  * Current

\[
I = \mu \frac{C_{\text{ox}} W V^2}{A L} \quad \text{and} \quad C_{\text{ox}} = \varepsilon \frac{W L}{T_{\text{ox}}} = \frac{1}{S} \frac{1}{S} \frac{1}{S} = 87%
\]

* **S Factor S~1.15**
  * 1/S 87%
  * 1/S 87%
  * 1/S 87%
  * 1/S 87%
Scaling Parameters for CMOS

* **Parameter**
  * W, L, T\textsubscript{ox}, X\textsubscript{j}
  * Voltage (constant field)
  * Capacitance
  * Current
  * Delay

* **S Factor S~1.15**
  * 1/S 87%

\[ C_{ox} = \varepsilon \frac{WL}{T_{ox}} = \frac{1}{S} \frac{1}{S} \]

\[ I = \mu \frac{C_{ox} W V^2}{A L} \quad \tau = C_{ox} \frac{V}{I} \]

* 1/S 87%
* 1/S 87%
* 1/S 87%
* 1/S 87%
## Scaling Parameters for CMOS

<table>
<thead>
<tr>
<th>Parameter</th>
<th>S Factor S~1.15</th>
</tr>
</thead>
<tbody>
<tr>
<td>W, L, T_{ox}, X_j</td>
<td>1/S 87%</td>
</tr>
<tr>
<td>Voltage (constant field)</td>
<td>1/S 87%</td>
</tr>
<tr>
<td>Capacitance</td>
<td>1/S 87%</td>
</tr>
<tr>
<td>Current</td>
<td>1/S 87%</td>
</tr>
<tr>
<td>Delay</td>
<td>1/S 87%</td>
</tr>
<tr>
<td>Power</td>
<td>1/S^2 76%</td>
</tr>
</tbody>
</table>

### Equations

- Capacitance:
  \[ C_{ox} = \varepsilon \frac{WL}{T_{ox}} = \frac{V}{S} \frac{1}{S} \]
- Current:
  \[ I = \mu \frac{C_{ox} W V^2}{A L} \]
- Delay:
  \[ \tau = C_{ox} \frac{V}{I} \]
- Power:
  \[ P = IV \]
Scaling Parameters for CMOS

* **Parameter**
  * W, L, T_{ox}, X_j
  * Voltage (constant field)
  * Capacitance
  * Current
  * Delay
  * Power
  * Energy

* **S Factor S~1.15**
  * 1/S 87%
  * 1/S 87%
  * 1/S 87%
  * 1/S 87%
  * 1/S \( \frac{1}{S^2} \) 76%
  * 1/S \( \frac{1}{S^3} \) 66%

\[
C = \varepsilon \frac{WL}{T_{ox}} = \frac{1}{S} \frac{1}{S} \frac{1}{S}
\]

\[
I = \mu \frac{C_{ox} W V^2}{A L} \frac{1}{S} \frac{1}{S}
\]

\[
\tau = C_{ox} \frac{V}{I}
\]

\[
P = IV
\]

\[
E = P \tau
\]
Scaling Parameters for CMOS

* **Parameter**
  * W, L, T_{ox}, X_j
  * Voltage (constant field)
  * Capacitance
  * Current
  * Delay
  * Power
  * Energy
  * Die Size

* **S Factor S ~ 1.15**
  * 1/S 87%
  * 1/S 87%
  * 1/S 87%
  * 1/S 87%
  * 1/S^2 76%
  * 1/S^3 66%
  * S_c 115%

\[
C_{ox} = \varepsilon \frac{W L}{T_{ox}} = \frac{1}{S} \frac{1}{S} \frac{1}{S}
\]

\[
I = \mu \frac{C_{ox} W V^2}{A \frac{L}{2}}
\]

\[
\tau = C_{ox} \frac{V}{I}
\]

\[
P = IV
\]

\[
E = P \tau
\]
Scaling Parameters for CMOS

* Parameter
  * W, L, T_{ox}, X_j
  * Voltage (constant field)
  * Capacitance
    \[ C = \epsilon \frac{WL}{T_{ox}} = \frac{1}{S} \frac{1}{S} \]
  * Current
    \[ I = \mu \frac{C_{ox} W V^2}{A L \frac{1}{2}} \]
  * Delay
    \[ \tau = C_{ox} \frac{V}{I} \]
  * Power
    \[ P = IV \]
  * Energy
    \[ E = P \tau \]
  * Die Size
  * Number per Die

* S Factor S~1.15
  * 1/S 87%
  * 1/S 87%
  * 1/S 87%
  * 1/S 87%
  * 1/S^2 76%
  * 1/S^3 66%
  * S_c 115%
  * S^2 S_c^2 175%
Interconnection Scaling
Scaling Parameters for Interconnections

<table>
<thead>
<tr>
<th>Parameter</th>
<th>S Factor</th>
<th>S~1.15</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wm, Ws, H, L, t</td>
<td>1/S</td>
<td>87%</td>
</tr>
<tr>
<td>Resistance/L = ( \rho/W_mH )</td>
<td>S^2</td>
<td>132%</td>
</tr>
<tr>
<td>Local Length, L</td>
<td>1/S</td>
<td>87%</td>
</tr>
<tr>
<td>Local RC Delay = RCLL</td>
<td>1</td>
<td>100%</td>
</tr>
<tr>
<td>Die Size</td>
<td>S_c</td>
<td>115%</td>
</tr>
<tr>
<td>Global RC Delay = RCS_c^2</td>
<td>S^2S_c^2</td>
<td>175%</td>
</tr>
</tbody>
</table>
Problems with Scaling

1. Global Wiring Crisis
The Problem with Wires

![Graph showing delay time vs. technology node (nm)].

- **Longest Interconnect Delay**
- **Typical Gate Delay**
The Problem with Wires

- Longest Interconnect Delay
- Typical Gate Delay

![Graph showing delay time vs. technology node](image)

![Diagram illustrating wire structure](image)

- Resistivity vs. wire width
- 22nm Node
- 32nm Node
- Grain Boundary Scattering
- Surface Scattering
- Cu (bulk)

Wire width [nm]

- 40
- 60
- 80
- 100
**Scaling Parameters for Interconnections**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>S Factor</th>
<th>S~1.15</th>
</tr>
</thead>
<tbody>
<tr>
<td>* Wm, Ws, H, L, t</td>
<td>* 1/S</td>
<td>87%</td>
</tr>
<tr>
<td>* Resistance/L &gt; ( \rho/W_mH )</td>
<td>* 5S²</td>
<td>x132%</td>
</tr>
<tr>
<td>* Local Length, L</td>
<td>* 1/S</td>
<td>87%</td>
</tr>
<tr>
<td>* Local RC Delay = RCLL</td>
<td>* 5</td>
<td>500%</td>
</tr>
<tr>
<td>* Die Size</td>
<td>* ( S_c )</td>
<td>115%</td>
</tr>
<tr>
<td>* Global RC Delay = RCS_c²</td>
<td>* ( S^2S_c^2 )</td>
<td>175%</td>
</tr>
</tbody>
</table>
Problems with Scaling

1. Global Wiring Crisis
2. Local Wiring Crisis
## ITRS projections and Industry Roadmap

<table>
<thead>
<tr>
<th>Year</th>
<th>2007 ITRS</th>
<th>2010</th>
<th>2013</th>
<th>2016</th>
<th>2019</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>45 nm</td>
<td>32 nm</td>
<td>22 nm</td>
<td>16 nm</td>
<td></td>
</tr>
<tr>
<td>Off-chip BW (GHz)</td>
<td>15.1</td>
<td>23.0</td>
<td>39.7</td>
<td>62.4</td>
<td></td>
</tr>
<tr>
<td>Package Pin Count</td>
<td>2783</td>
<td>3704</td>
<td>4930</td>
<td>6562</td>
<td></td>
</tr>
<tr>
<td>(Cost-Effective)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Package Pin Count</td>
<td>4851</td>
<td>5616</td>
<td>6501</td>
<td>7525</td>
<td></td>
</tr>
<tr>
<td>(High Performance)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Max Chip Power</td>
<td>198</td>
<td>198</td>
<td>198</td>
<td>198</td>
<td></td>
</tr>
</tbody>
</table>
• Exponential decay and change in magnitude and phase of signal

• Resistance (R) and Inductance (L) governed by metal properties

• Capacitance (C) and Shunt Conductance (G) governed by insulator properties

\[ E = E_0 e^{-\gamma z} \]

\[ \gamma = \alpha + j\beta = \sqrt{(R + j\omega L)(G + j\omega C)} \]

\[ \alpha_c = \frac{\sqrt{\omega \mu_0 \varepsilon_r}}{2 \sigma \mu_r} \eta_0 d \]

\[ \alpha_d = \frac{\omega \sqrt{\mu_r \varepsilon_r} \tan \delta}{2c} \]
Off-Chip Bandwidth

ITRS 2006

Off-Chip Freq. (GHz)

Technology (nm)

Off-Chip Bandwidth

Attenuation [dB]

9" FR4, via stub

26" FR4, via stub

9" FR4

26" FR4
Problems with Scaling

1. Global Wiring Crisis
2. Local Wiring Crisis
3. Off-chip Wiring Crisis
The Problem with Wires

- Interconnects: >50% dynamic pwr

N. Magen et al. *SLIP* 2004

- Global Signal: 34%
- Global Clock: 19%
- Local Signals: 27%
- Local Clock: 20%

Total power (Dynamic) 130 nm μP

N. Magen et al. *SLIP* 2004
Problems with Scaling

1. Global Wiring Crisis
2. Local Wiring Crisis
3. Off-chip Wiring Crisis
4. Power limitations: $V_{dd}$ and $t_{ox}$ are not scaling
   »  Power is now *everything*
   »  Heat: Steady-state operation no longer possible
Shorter Wires and Cooler Transistors
On-Chip Ultra Low-\(k\)
Cross-section Interconnect Structure

Interconnect (BEOL)

- Passivation
- Dielectric
- Etch Stop Layer
- Dielectric Capping Layer
- Copper Conductor with Barrier/Nucleation Layer
- Pre-Metal Dielectric
- Tungsten Contact Plug

Metal 1

Global

Intermediate
Sacrificial Materials

Thermogravimetric Analysis

% Weight

Temp (C)

0 50 100 150 200 250 300 350 400 450 500

0 10 20 30 40 50 60 70 80 90 100

R

O

R

O

R

O

R

O

R

O
Air-Channels in $\text{SiO}_2$
**Effective Dielectric Constant ($k_{\text{eff}}$)**

- **Standing alone Cu lines**
  - $k_{\text{eff}} = 1.81$

- **After RIE of SiO$_2$**
  - $k_{\text{eff}} = 1.23$

- **Homogeneous SiO$_2$**
  - $k_{\text{eff}} = 4.14$
  - Width of Cu : 200 nm
  - Aspect ratio : 1.8:1 (H:W)
  - Half pitch : 200 nm
  - Extended height : 80nm(top)/ 100nm(bottom)
  - Inter-layer dielectric : PECVD SiO$_2$

- **Air-gaps**
  - $k_{\text{eff}} = 2.42$

- **Extended air-gaps**
  - $k_{\text{eff}} = 2.17$
Off-Chip Ultra Low-$k$

Todd Spencer
### ITRS projections and Industry Roadmap

<table>
<thead>
<tr>
<th></th>
<th>2007 ITRS</th>
<th>2010</th>
<th>2013</th>
<th>2016</th>
<th>2019</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>45 nm</td>
<td>32 nm</td>
<td>22 nm</td>
<td>16 nm</td>
</tr>
<tr>
<td>Off-chip BW (GHz)</td>
<td></td>
<td>15.1</td>
<td>23.0</td>
<td>39.7</td>
<td>62.4</td>
</tr>
<tr>
<td>Package Pin Count</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Cost-Effective)</td>
<td></td>
<td>2783</td>
<td>3704</td>
<td>4930</td>
<td>6562</td>
</tr>
<tr>
<td>Package Pin Count</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(High Performance)</td>
<td></td>
<td>4851</td>
<td>5616</td>
<td>6501</td>
<td>7525</td>
</tr>
<tr>
<td>Max Chip Power</td>
<td></td>
<td>198</td>
<td>198</td>
<td>198</td>
<td>198</td>
</tr>
</tbody>
</table>
Off-Chip Bandwidth

ITRS 2006

Off-Chip Freq. (GHz) vs. Technology (nm)

Attenuation [dB]

9" FR4, via stub
26" FR4, via stub
Motivation for Air Insulation on FR4

* Primary loss due to dielectric in conventional fiberglass-epoxy substrates above 10 GHz

* Conductor loss increases by \( \text{freq}^{1/2} \)

* Dielectric loss scales linearly by \( \text{freq} \)

* Air isolation has lowest \( \varepsilon_r \) and \( \tan \delta \)

* Signal propagation velocity also increases, lowering latency

\[
\frac{E}{E_0} = e^{-\alpha \cdot \text{length}}
\]

\[
\alpha_{\text{conductor}} = \sqrt{\frac{\omega \mu_0 \varepsilon_r}{2 \sigma \mu_r}} \frac{\eta_0 d}{\eta_0 d}
\]

\[
\alpha_{\text{dielectric}} = \frac{\omega \sqrt{\mu_r \varepsilon_r} \tan \delta}{2c}
\]
Acid Catalyzed Decomposition

- Acid generation via PAG, aid decomposition onset temperature \(~100\,^\circ\text{C}\).
- Photo-acid & Thermal acid Generation.

Chemical Reaction:

\[
\text{PAG} \xrightarrow{\text{hv or } \Delta} \text{H}^+ \text{X}^- + \text{Other products}
\]
Fabrication of Positive-Tone Sacrificial Polymer

1. Spin-coat Unity
2. UV Expose through Mask (248 nm or 365 nm)
3. Bake develop exposed areas (110 °C)
4. Decompose unexposed areas (170 °C)
5. Overcoat
Air-Clad Transmission Lines on Organic Substrates
Air-Clad Transmission Lines on Organic Substrates

* Parallel plate and suspended ground microstrip lines
* Capacitance reduced by more than 30%
* Loss tangent reduced by more than 85%
* Reduces both conductor and dielectric loss contributions
Air Cavity Parallel Plate & Microstriplines

* Air cavity formed between signal and ground lines on FR4
* No polymer between signal and ground (all air)
* Improves on previously reported partial air-gap lines
* Air cavity reduces capacitance by up to 47% for ground line 3x signal line width
* Loss tangent reduced by up to 90%

<table>
<thead>
<tr>
<th>Ground Line Width (μm)</th>
<th>Capacitance Before Air Cavity (pF)</th>
<th>Capacitance with Air Cavity (pF)</th>
<th>Capacitance % reduction with air cavity</th>
</tr>
</thead>
<tbody>
<tr>
<td>650</td>
<td>16.42</td>
<td>9.02</td>
<td>46.6</td>
</tr>
<tr>
<td>650</td>
<td>15.94</td>
<td>10.16</td>
<td>38.0</td>
</tr>
<tr>
<td>650</td>
<td>15.29</td>
<td>8.37</td>
<td>46.6</td>
</tr>
<tr>
<td>220</td>
<td>12.56</td>
<td>10.27</td>
<td>21.7</td>
</tr>
<tr>
<td>220</td>
<td>12.83</td>
<td>10.45</td>
<td>21.5</td>
</tr>
<tr>
<td>220</td>
<td>12.85</td>
<td>10.12</td>
<td>25.6</td>
</tr>
</tbody>
</table>
* 1 dB/GHz channel loss allows high aggregate bandwidth at low power density
* 8 dB/GHz loss decreases bandwidth and increases power density
* Channel loss should be as small as possible to minimize energy per bit
* Channel loss can be minimized with air cavity strip

Figures courtesy Rizwan Bashirullah, U. of Florida
Air Cavity Lines for Off-chip Communication

- Research test vehicle designed at University of Florida
- Air cavity differential signal lines between for off-chip communication

- Test chips with 4x10Gb/s transmitters and cross-talk cancellation to test air cavity line performance
Future Multilayer Channel Buildup

* Low loss signal line build-up on multilayer boards uses existing infrastructure
* Channel cross-sectional area determines loss and channel density
  » Unshielded, nonplanar structures easy to build, higher loss
  » Shielded structures have lower loss, but larger cross-sectional area
  » Buildup process gives nonplanar lines
  » Inlay processes are more valuable, more difficult to make
Future Imprint Lithography: Complex geometries

- Coaxial geometry minimizes crosstalk noise and radiation losses
- Channel geometries with complex shapes can be built using imprint lithography
- Routing terminations and chip connections can be more easily defined using imprint
- Smooth, rounded transitions will minimize reflections and maximize power transmission
High Performance Chip-to-Substrate Connections

Tyler Osborn
**Introduction**

* Flip-Chip Interconnects are a critical interface between the IC and system
* Large numbers of connections are made over a small surface area
  » 1,000 – 10,000 I/O
  » Typical Die is ~ 225 mm²
* Solder melt-cast connections are the industry standard

<table>
<thead>
<tr>
<th>Year</th>
<th>2010</th>
<th>2015</th>
<th>2020</th>
</tr>
</thead>
<tbody>
<tr>
<td>Area Array Pitch (µm)</td>
<td>120</td>
<td>100</td>
<td>85</td>
</tr>
<tr>
<td>I/O per cm²</td>
<td>7000</td>
<td>1000</td>
<td>1400</td>
</tr>
<tr>
<td>Off Chip Freq. (GHz)</td>
<td>9.5</td>
<td>29.1</td>
<td>72.4</td>
</tr>
</tbody>
</table>
All-Copper Approach

* Replace solder with copper:
  » Electrical improvement
  » Thermo-Mechanical
* Copper-to-copper bonding:
  » Low Cost
  » Low Temperature
  » Mechanical Compliance
* Current SnPb and SnAgCu:
  » Undesirable properties:
    • Limited (H:W) aspect ratio
    • Brittle intermetallics
    • High capacitance (underfill)
    • Low electromigration resistance
  » Desirable properties:
    • Low processing temperature
    • Tolerance- x-y misalignment
    • Tolerance to height variation

Addition of underfill to Cu pillar system is possible for reliability
All Copper Interconnects

1. Flip Chip Alignment  
2. Electroless Plating  
3. Low T Annealing  
4. Shear Testing

Patent Pending Process:
Integrated Circuit Interconnection Devices and Methods
Ate He, Tyler Osborn, Paul Kohl. September 2006
US Patent Application # 20080073795
* No anneal: Unbonded interface between electrolessly plated pillars prior to annealing processes
Bonding Mechanism

» Additional annealing at 180°C for 30 minutes, interface is eliminated
Will Other Metals Bond?

- Electroless gold
- Annealed to form joint
- Gold-Gold seam closed
* The stress state within the copper pillar and at the chip & substrate surface is a function of:
  » Position
  » Shape
  » ‘Collar’ Material
GPD Model Results

- Pillar Height
- Pillar Diameter

- 8.8 fF
- 148 MPa
- 300 pH

- Low parasitic C region
- Low parasitic L region
- Mechanically compliant
**Single Pillar Model**

- A single pillar model - mimics previous GPD model
- Correctly matches maximum stresses for all GPD results
- **TARGET:** Lower stress from 148 MPa (adhesive failure) to below ‘solder & low-k’ values.
- **NEW GOAL:** Yield stress of solder is the only agreed upon value.
  - SnAgCu ~ 50 MPa
  - Eutectic SnPb ~ 25 MPa
Results – Single Polymer Collar

Effect of Collar Modulus on Max Stress

Overall Maximum Stress (MPa)

Elastic Modulus of the Polymer Collar (GPa)

- Copper Pillar
- Gold Pillar

Collar Modulus: 0.5 GPa, 8.0 GPa, 18.0 GPa
Comparing Polymer Collars

Aq. Avatrel

SU-8
Bi-Layer Collar Results

- SU-8 Top Collar with E=21 GPa Base
- SU-8 Top Collar with E=50 GPa Base

High E Material (~20 GPa)
High AR Definable Material

Isosurface: von Mises stress [MPa]
Max: 65.0
Min: 30.0

SU-8 Top Collar with E=50 GPa Base
Bi-Layer Collar Results

First layer modulus = 2.5 or 4 GPa

Copper Pillar with Bi-Layer Collar

Maximum Stress (MPa)

Elastic Modulus of the Bottom Collar Material (GPa)

- Aquatrel Top 2.5 GPa
- SU-8 Top 4.0 GPa
Changing Shape

- Cylindrical I/O has high stress point at leading and tailing edge of pillar.
- Distribute the stress at the chip interface over a greater area by using *Ellipse or Square I/O*
- **KEY:** Align the broad face directly in the direction of thermal expansion
**Ellipse Effect Results**

Benefit of Elliptical I/O Shape on Stress
Aquatrel Top (2.5 GPa) with 21 GPa Bottom

- **Objective**
- **Chip Surface Stress Max**
- **Structure Maximum Stress**
- **Yield of Eutectic SnPb**
- **Yield of SnAgCu**

Stress (MPa) vs. A/B Ratio of Ellipse graph.
High Frequency, Shielded I/O
Future Insulation, Cooling, and Conductors
3-D Electronics: Short Wires

Die #3
Die #2
Die #1

Electrical  Optical  Thermal

Substrate

Si Die

Kohl, Bakir, Bashirulla, GT and UFL
Comments

* Transistor scaling is slowing
* Need for Bandwidth (not clock speed) is increasing due to applications
* Critical needs:
  » Very low Permittivity and Loss materials
  » Ballistic transport in conductors (no scattering)
  » Advanced cooling (Fluids-to-the-chip)
  » Simple, high-yield processes