Lec 1
Overview and Introduction
Outline

- **Course Contents**
  - Moore’s law
  - CMOS scaling theory
  - Impact of interconnects
  - The interconnect-driven design paradigm is needed

- **Related Articles**
  - *Chapter 5.4 of Circuits, Interconnects, and Packaging for VLSI*
Components in VLSI Circuits

- **Devices**
  - Transistors
  - Logic gates and cells
  - Functional blocks

- **Interconnects**
  - Local signals
  - Global signals
  - Clock signals
  - Power/Ground networks

![Diagram of VLSI Circuit Components](image-url)
Determining Factors of System Performance

- Clock Period $\geq t_{\text{logic}} + t_{\text{interconnect}} + t_{\text{skew}} + t_{\text{su}} + t_s$
  - $t_{\text{logic}} =$ gate delay of the combinational logic
  - $t_{\text{interconnect}} =$ interconnect delay of the combinational logic
  - $t_{\text{skew}} =$ clock skew
  - $t_{\text{su}} =$ set up time
  - $t_s =$ delay within storage elements
Importance of VLSI Interconnect Design

- **Technology Trends**
  - Deep sub-micron design: < 0.25 μm CMOS technology
  - Nanometer technology: < 0.1 μm.
  - High-speed: > 500 MHz

- **Impact on VLSI System Design**
  - Interconnect delay becomes the dominating factor in system performance
    - Local interconnects: Almost constant
    - Global interconnects: RC delay goes as 1/S or 1/S³
    - Consumes 50%~70% clock cycle
  - Distributed nature of interconnects becomes significant
    - Becomes distributed RCL circuits or lossy transmission lines
  - Parasitics give rise to a whole set of signal integrity issues

Design paradigm shifts from device-centric to interconnect-centric
Delay for Metal 1 and Global Wiring v.s. Feature Size

from the ITRS’2003
Moore’s Law: Driving Technology Advances

- **Moore’s Law** (Gordon Moore, Intel’s co-founder) (1975)
  - Number of transistors doubles every 2.3 years (acceleration over past years: 1.5 years)

- **D. House**: Computer performance doubles every 18 months (1975)

Source: [http://www.intel.com](http://www.intel.com)
# ITRS - International Technology Roadmap for Semiconductors (ITRS’10 Update)

<table>
<thead>
<tr>
<th>Technology (nm)</th>
<th>hp29</th>
<th>hp24</th>
<th>hp20</th>
<th>hp15.3</th>
<th>hp11.3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Year</td>
<td>2009</td>
<td>2011</td>
<td>2013</td>
<td>2016</td>
<td>2019</td>
</tr>
<tr>
<td>Transistors density (Mtransistors/cm²) (including 0n-chip SRAM)</td>
<td>851</td>
<td>1,701</td>
<td>3,403</td>
<td>6,806</td>
<td>13,612</td>
</tr>
<tr>
<td>On-Chip (GHz)</td>
<td>5.45</td>
<td>6.33</td>
<td>7.34</td>
<td>9.18</td>
<td>11.47</td>
</tr>
<tr>
<td>Area (mm²)</td>
<td>260</td>
<td>260</td>
<td>260</td>
<td>260</td>
<td>260</td>
</tr>
<tr>
<td>#metal levels</td>
<td>12</td>
<td>12</td>
<td>13</td>
<td>13</td>
<td>14</td>
</tr>
<tr>
<td>Allowable maximum power (High performance with heat sink) (W)</td>
<td>143</td>
<td>161</td>
<td>149</td>
<td>130</td>
<td>133</td>
</tr>
<tr>
<td>$V_{dd}$ (High Power/Low Power) (V)</td>
<td>1.0/0.95</td>
<td>0.93/0.85</td>
<td>0.87/0.80</td>
<td>0.78/0.75</td>
<td>0.71/0.65</td>
</tr>
<tr>
<td>Battery (low-cost/handheld) (W)</td>
<td>3.0</td>
<td>3.0</td>
<td>3.0</td>
<td>3.0</td>
<td>3.0</td>
</tr>
</tbody>
</table>

**Predictions** of the worldwide semiconductor/IC industry about its own future prospects.....
### Predictions of the worldwide semiconductor/IC industry about its own future prospects.....

<table>
<thead>
<tr>
<th>Technology (nm)</th>
<th>hp29</th>
<th>hp24</th>
<th>hp20</th>
<th>hp15.3</th>
<th>hp11.3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Year</td>
<td>2009</td>
<td>2011</td>
<td>2013</td>
<td>2016</td>
<td>2019</td>
</tr>
<tr>
<td>Transistors density (Mtransistors/cm²)(including On-chip SRAM)</td>
<td>851</td>
<td>1,701</td>
<td>3,403</td>
<td>6,806</td>
<td>13,612</td>
</tr>
<tr>
<td>On-Chip (GHz)</td>
<td>5.45</td>
<td>6.33</td>
<td>7.34</td>
<td>9.18</td>
<td>11.47</td>
</tr>
<tr>
<td>Area (mm²)</td>
<td>260</td>
<td>260</td>
<td>260</td>
<td>260</td>
<td>260</td>
</tr>
<tr>
<td>#metal levels</td>
<td>12</td>
<td>12</td>
<td>13</td>
<td>13</td>
<td>14</td>
</tr>
<tr>
<td>Allowable maximum power (High performance with heat sink)(W)</td>
<td>143</td>
<td>161</td>
<td>149</td>
<td>130</td>
<td>133</td>
</tr>
<tr>
<td>V&lt;sub&gt;dd&lt;/sub&gt; (High Power/Low Power) (V)</td>
<td>1.0/0.95</td>
<td>0.93/0.85</td>
<td>0.87/0.80</td>
<td>0.78/0.75</td>
<td>0.71/0.65</td>
</tr>
<tr>
<td>Battery (low-cost/handheld)(W)</td>
<td>3.0</td>
<td>3.0</td>
<td>3.0</td>
<td>3.0</td>
<td>3.0</td>
</tr>
</tbody>
</table>
## Predictions of the worldwide semiconductor/IC industry about its own future prospects.....

### ITRS- International Technology Roadmap for Semiconductors (ITRS’10 Update)

<table>
<thead>
<tr>
<th>Technology (nm)</th>
<th>hp29</th>
<th>hp24</th>
<th>hp20</th>
<th>hp15.3</th>
<th>hp11.3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Year</td>
<td>2009</td>
<td>2011</td>
<td>2013</td>
<td>2016</td>
<td>2019</td>
</tr>
<tr>
<td>Transistors density (Mtransistors/cm²)(including 0n-chip SRAM)</td>
<td>851</td>
<td>1,701</td>
<td>3,403</td>
<td>6,806</td>
<td>13,612</td>
</tr>
<tr>
<td>On-Chip (GHz)</td>
<td>5.45</td>
<td>6.33</td>
<td>7.34</td>
<td>9.18</td>
<td>11.47</td>
</tr>
<tr>
<td>Area (mm²)</td>
<td>260</td>
<td>260</td>
<td>260</td>
<td>260</td>
<td>260</td>
</tr>
<tr>
<td>#metal levels</td>
<td>12</td>
<td>12</td>
<td>13</td>
<td>13</td>
<td>14</td>
</tr>
<tr>
<td>Allowable maximum power (High performance with heat sink)(W)</td>
<td>143</td>
<td>161</td>
<td>149</td>
<td>130</td>
<td>133</td>
</tr>
<tr>
<td>V&lt;sub&gt;dd&lt;/sub&gt; (High Power/Low Power) (V)</td>
<td>1.0/0.95</td>
<td>0.93/0.85</td>
<td>0.87/0.80</td>
<td>0.78/0.75</td>
<td>0.71/0.65</td>
</tr>
<tr>
<td>Battery (low-cost/handheld)(W)</td>
<td>3.0</td>
<td>3.0</td>
<td>3.0</td>
<td>3.0</td>
<td>3.0</td>
</tr>
</tbody>
</table>
### ITRS-International Technology Roadmap for Semiconductors (ITRS’10 Update)

<table>
<thead>
<tr>
<th>Technology (nm)</th>
<th>hp29</th>
<th>hp24</th>
<th>hp20</th>
<th>hp15.3</th>
<th>hp11.3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Year</td>
<td>2009</td>
<td>2011</td>
<td>2013</td>
<td>2016</td>
<td>2019</td>
</tr>
<tr>
<td>Transistors density (Mtransistors/cm²) (including On-chip SRAM)</td>
<td>851</td>
<td>1,701</td>
<td>3,403</td>
<td>6,806</td>
<td>13,612</td>
</tr>
<tr>
<td>On-Chip (GHz)</td>
<td>5.45</td>
<td>6.33</td>
<td>7.34</td>
<td>9.18</td>
<td>11.47</td>
</tr>
<tr>
<td>Area (mm²)</td>
<td>260</td>
<td>260</td>
<td>260</td>
<td>260</td>
<td>260</td>
</tr>
<tr>
<td>#metal levels</td>
<td>12</td>
<td>12</td>
<td>13</td>
<td>13</td>
<td>14</td>
</tr>
<tr>
<td>Allowable maximum power (High performance with heat sink) (W)</td>
<td>143</td>
<td>161</td>
<td>149</td>
<td>130</td>
<td>133</td>
</tr>
<tr>
<td>V_{dd} (High Power/Low Power) (V)</td>
<td>1.0/0.95</td>
<td>0.93/0.85</td>
<td>0.87/0.80</td>
<td>0.78/0.75</td>
<td>0.71/0.65</td>
</tr>
<tr>
<td>Battery (low-cost/handheld) (W)</td>
<td>3.0</td>
<td>3.0</td>
<td>3.0</td>
<td>3.0</td>
<td>3.0</td>
</tr>
</tbody>
</table>

Predictions of the worldwide semiconductor/IC industry about its own future prospects.....
## ITRS- International Technology Roadmap for Semiconductors (ITRS’10 Update)

<table>
<thead>
<tr>
<th>Technology (nm)</th>
<th>hp29</th>
<th>hp24</th>
<th>hp20</th>
<th>hp15.3</th>
<th>hp11.3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Year</td>
<td>2009</td>
<td>2011</td>
<td>2013</td>
<td>2016</td>
<td>2019</td>
</tr>
<tr>
<td>Transistors density (Mtransistors/cm²)(including 0n-chip SRAM)</td>
<td>851</td>
<td>1,701</td>
<td>3,403</td>
<td>6,806</td>
<td>13,612</td>
</tr>
<tr>
<td>On-Chip (GHz)</td>
<td>5.45</td>
<td>6.33</td>
<td>7.34</td>
<td>9.18</td>
<td>11.47</td>
</tr>
<tr>
<td>Area (mm²)</td>
<td>260</td>
<td>260</td>
<td>260</td>
<td>260</td>
<td>260</td>
</tr>
<tr>
<td>#metal levels</td>
<td>12</td>
<td>12</td>
<td>13</td>
<td>13</td>
<td>14</td>
</tr>
<tr>
<td>Allowable maximum power (High performance with heat sink)(W)</td>
<td>143</td>
<td>161</td>
<td>149</td>
<td>130</td>
<td>133</td>
</tr>
<tr>
<td>$V_{dd}$ (High Power/Low Power) (V)</td>
<td>1.0/0.95</td>
<td>0.93/0.85</td>
<td>0.87/0.80</td>
<td>0.78/0.75</td>
<td>0.71/0.65</td>
</tr>
<tr>
<td>Battery (low-cost/handheld)(W)</td>
<td>3.0</td>
<td>3.0</td>
<td>3.0</td>
<td>3.0</td>
<td>3.0</td>
</tr>
</tbody>
</table>

Predictions of the worldwide semiconductor/IC industry about its own future prospects…..
Why?

- Why more transistors per IC?
  - Smaller transistors
  - Larger dice

- Why faster computers?
  - Smaller, faster transistors
  - Better micro-architecture (more IPC)
  - Fewer gate delays per cycle
Scaling

- The only constant in VLSI is constant change
- Feature size shrinks by 30% every 2-3 years
  - Transistors become cheaper
  - Transistors become faster
  - Wires do not improve (and may get worse)!!!
- Scale ratio $S$
  - Typically $S = 2^{1/2}$
    - The number of transistors per unit area is \textit{doubled} with each generation under constant field scaling.
    - The transistor performance is \textit{doubled} every two generations under constant field scaling.
  - Technology nodes
MOS Transistor Scaling

- **Constant-Field Scaling**: Attempts to preserve the magnitude of internal electric fields in the MOSFET.
  - All dimensions \((W, L, t_{g ox}, x_j)\) are reduced by \(S\).
  - Supply voltage \((V_{dd})\) and threshold voltage \((V_{T0})\) are reduced by \(S\).
  - Doping density \((N_A, N_B)\) is increased by \(S\).

- **Constant-Voltage Scaling**
  - All dimensions \((W, L, t_{g ox}, x_j)\) are reduced by \(S\).
  - Supply voltage \((V_{dd})\) and threshold voltage \((V_{T0})\) remain unchanged.
  - Doping density \((N_A, N_B)\) is increased by \(S^2\).
# MOS Transistors under Different Scaling Rules

<table>
<thead>
<tr>
<th>Quantity</th>
<th>Sensitivity</th>
<th>Constant Field</th>
<th>Constant Voltage</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Scaling Parameters</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Length</td>
<td>$L$</td>
<td>$1/S$</td>
<td>$1/S$</td>
</tr>
<tr>
<td>Width</td>
<td>$W$</td>
<td>$1/S$</td>
<td>$1/S$</td>
</tr>
<tr>
<td>Gate Oxide Thickness</td>
<td>$t_{g_{ox}}$</td>
<td>$1/S$</td>
<td>$1/S$</td>
</tr>
<tr>
<td>Supply Voltage</td>
<td>$v_{dd}$</td>
<td>$1/S$</td>
<td>$1$</td>
</tr>
<tr>
<td>Threshold Voltage</td>
<td>$V_{T0}$</td>
<td>$1/S$</td>
<td>$1$</td>
</tr>
<tr>
<td>Doping Density</td>
<td>$N_A, N_B$</td>
<td>$S$</td>
<td>$S^2$</td>
</tr>
</tbody>
</table>

## Device Characteristics

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Area (A)</td>
<td>$WL$</td>
<td>$1/S^2$</td>
<td>$1/S^2$</td>
</tr>
<tr>
<td>$\beta$</td>
<td>$W/Lt_{g_{ox}}$</td>
<td>$S$</td>
<td>$S$</td>
</tr>
<tr>
<td>D-S Current ($i_{DS}$)</td>
<td>$\beta(v_{dd} - v_{T})^2$</td>
<td>$1/S$</td>
<td>$S$</td>
</tr>
<tr>
<td>Gate Capacitance ($C_g$)</td>
<td>$WL/t_{g_{ox}}$</td>
<td>$1/S$</td>
<td>$1/S$</td>
</tr>
<tr>
<td>Transistor On-Resistance ($R_{tr}$)</td>
<td>$v_{dd}/i_{DS}$</td>
<td>$1$</td>
<td>$1/S$</td>
</tr>
<tr>
<td>Intrinsic Gate Delay ($\tau = C_g \Delta V/I_{av}$)</td>
<td>$R_{tr}C_g$</td>
<td>$1/S$</td>
<td>$1/S^2$</td>
</tr>
<tr>
<td>Clock Frequency</td>
<td>$f$</td>
<td>$f$</td>
<td>$F$</td>
</tr>
<tr>
<td>Dynamic Power Dissipation per Gate ($P$)</td>
<td>$(f/2)C_gV_{dd}^2$</td>
<td>$f/S^3$</td>
<td>$f/S$</td>
</tr>
<tr>
<td>Power Dissipation Density ($P/A$)</td>
<td>$P/A$</td>
<td>$f/S$</td>
<td>$fS$</td>
</tr>
</tbody>
</table>
Observations

- Gates get faster with scaling (good)
- Dynamic power goes down with scaling (good)
- Current density goes up with scaling (bad)
- Velocity saturation makes lateral scaling unsustainable
Interconnect Scaling

- **Ideal Scaling**
  - Cross sectional dimensions \((W_{int}, H_{int}, W_{sp}, t_{ox})\) are reduced by \(S\).
  - Global interconnect length is increased by \(D\).

- **Constant Thickness Scaling**
  - Cross sectional dimensions \((W_{int}, W_{sp}, t_{ox})\) are reduced by \(S\).
  - Global interconnect length is increased by \(D\).
  - Wire thickness \((H_{int})\) remains unchanged.
## Interconnect under Different Scaling Rules

<table>
<thead>
<tr>
<th>Quantity</th>
<th>Sensitivity</th>
<th>Ideal</th>
<th>Constant Thickness</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cross Sectional Dimensions</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Width</td>
<td>$W_{int}$</td>
<td>$1/S$</td>
<td>$1/S$</td>
</tr>
<tr>
<td>Thickness</td>
<td>$H_{int}$</td>
<td>$1/S$</td>
<td>$1$</td>
</tr>
<tr>
<td>Spacing</td>
<td>$W_{sp}$</td>
<td>$1/S$</td>
<td>$1/S$</td>
</tr>
<tr>
<td>Interlayer Oxide Thickness</td>
<td>$t_{ox}$</td>
<td>$1/S$</td>
<td>$1/S$</td>
</tr>
<tr>
<td><strong>Characteristics per Unit Length</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wire Resistance per Unit Length ($R_{int}$)</td>
<td>$1/W_{int}H_{int}$</td>
<td>$S^2$</td>
<td>$S$</td>
</tr>
<tr>
<td>Fringing Capacitance per Unit Length ($C_{Fint}$)</td>
<td>$H_{int} / W_{sp}$</td>
<td>$1$</td>
<td>$S$</td>
</tr>
<tr>
<td>Parallel Plate Capacitance per Unit Length ($C_{Pint}$)</td>
<td>$W_{int} / t_{ox}$</td>
<td>$1$</td>
<td>$1$</td>
</tr>
<tr>
<td>Total Wire Capacitance per Unit Length ($C_{int}$)</td>
<td>$C_{Fint} + C_{Pint}$</td>
<td>$1$</td>
<td>$1$</td>
</tr>
<tr>
<td>Unrepeated RC Delay per Unit Length ($\tau_{wu}$)</td>
<td>$R_{int}C_{int}$</td>
<td>$S^2$</td>
<td>$S \sim S^2$</td>
</tr>
<tr>
<td>Repeated Wire RC Delay per Unit Length ($\tau_{wr}$)</td>
<td>$(R_{tr}C_{g}R_{int}C_{int})^{1/2}$</td>
<td>$S^{1/2}$</td>
<td>$1 \sim S^{1/2}$</td>
</tr>
<tr>
<td>Cross Talk Noise</td>
<td>$H_{int} / W_{sp}$</td>
<td>$1$</td>
<td>$S$</td>
</tr>
</tbody>
</table>
### Interconnect Delay under Different Scaling Rules

<table>
<thead>
<tr>
<th>Quantity</th>
<th>Sensitivity</th>
<th>Ideal</th>
<th>Constant Thickness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cross Sectional Dimensions</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Width</td>
<td>$W_{int}$</td>
<td>$1/S$</td>
<td></td>
</tr>
<tr>
<td>Thickness</td>
<td>$H_{int}$</td>
<td>$1/S$</td>
<td>$1$</td>
</tr>
<tr>
<td>Spacing</td>
<td>$W_{sp}$</td>
<td>$1/S$</td>
<td></td>
</tr>
<tr>
<td>Interlayer Oxide Thickness</td>
<td>$t_{ox}$</td>
<td>$1/S$</td>
<td></td>
</tr>
<tr>
<td>Local Interconnect Characteristics</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Length ($l_{loc}$)</td>
<td></td>
<td>$1/S$</td>
<td></td>
</tr>
<tr>
<td>Unrepeated Wire RC Delay</td>
<td>$l_{loc}^2\tau_{wu}$</td>
<td>$1$</td>
<td>$1/S\sim1$</td>
</tr>
<tr>
<td>Repeated Wire Delay</td>
<td>$l_{loc}\tau_{wr}$</td>
<td>$1/S^{1/2}$</td>
<td>$1/S\sim1/S^{1/2}$</td>
</tr>
<tr>
<td>Global Interconnect Characteristics</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Length ($l_{int}$)</td>
<td></td>
<td>$D$</td>
<td></td>
</tr>
<tr>
<td>Unrepeated Wire RC Delay</td>
<td>$l_{int}^2\tau_{wu}$</td>
<td>$D^2S^2$</td>
<td>$D^2S\sim D^2S^2$</td>
</tr>
<tr>
<td>Repeated Wire Delay</td>
<td>$l_{int}\tau_{wr}$</td>
<td>$DS^{1/2}$</td>
<td>$D\sim DS^{1/2}$</td>
</tr>
</tbody>
</table>
Observations

- Capacitance per micron is remaining constant
  - About 0.2 fF/µm
  - Roughly 1/10 of gate capacitance

- Local wires are getting faster
  - Not quite tracking transistor improvement
  - But not a major problem

- Global wires are getting slower
  - No longer possible to cross chip in one cycle
## Interconnect Parameters from ITRS’2008 Update

<table>
<thead>
<tr>
<th>Technology (nm)</th>
<th>32</th>
<th>27</th>
<th>22</th>
<th>18</th>
<th>15</th>
</tr>
</thead>
<tbody>
<tr>
<td>Year</td>
<td>2007</td>
<td>2009</td>
<td>2011</td>
<td>2013</td>
<td>2015</td>
</tr>
<tr>
<td>Res. $\rho$ ((\mu\Omega\cdot\text{cm}))</td>
<td>3.43</td>
<td>3.80</td>
<td>4.3</td>
<td>4.83</td>
<td>5.58</td>
</tr>
<tr>
<td>Dielectric constant</td>
<td>2.9-3.3</td>
<td>2.6-2.9</td>
<td>2.6-2.9</td>
<td>2.4-2.8</td>
<td>2.1-2.5</td>
</tr>
<tr>
<td>Intermediate wire pitch (nm)</td>
<td>136</td>
<td>104</td>
<td>80</td>
<td>64</td>
<td>50</td>
</tr>
<tr>
<td>1mm Cu wire RC delay (ps)</td>
<td>741</td>
<td>1320</td>
<td>2524</td>
<td>4044</td>
<td>6771</td>
</tr>
<tr>
<td>Intermediate wire aspect ratio (A/R)</td>
<td>1.8</td>
<td>1.8</td>
<td>1.8</td>
<td>1.9</td>
<td>1.9</td>
</tr>
<tr>
<td>Via aspect ratio (A/R)</td>
<td>1.6</td>
<td>1.6</td>
<td>1.6</td>
<td>1.7</td>
<td>1.7</td>
</tr>
<tr>
<td>$V_{dd}$ (High Power/Low Power) (V)</td>
<td>1.1/0.9</td>
<td>1.1/0.8</td>
<td>1.0/0.77</td>
<td>1.0/0.70</td>
<td>1.0/0.60</td>
</tr>
</tbody>
</table>
Scaling: Influence of Interconnect

- Scaling dimensions though resulted in many advantages it also resulted in significant *degradation in interconnect related circuit parameters*
  - Propagation delays (Timing) (Increasing)
    - wire $R$, $C$ and $L$ (Tpd $= F(R_{tr}, C_{tr}, R_{int}, C_{int})$)
  - Cross-talk effects (Signal Integrity)
    - (Aggravated)
      - False switching due to $C_c$, $L$
        ($\Delta V_{\text{victim}} = \alpha V_{\text{aggressor}}$)
  - Clock skew (Signal Integrity)
    - wire $R$, $C$ and $L$
  - Electromigration effect (Reliability) (Degraded)
    - wire $R$, $C$ (MTTF $= \alpha j^2 \exp(E_a/kT)$)
  - IR Voltage Drop (Reliability) (Significant)
    - wire $R$, $C$ ($\Delta V = IR$)
  - Power Consumption (Reliability) (Enhanced)
    - wire $C$

Which in fact are chip design issues.  K. Saraswat et al. TFUG 3/20/02
Impact of Interconnect on Propagation Delay

- Effects of scaling both geometry and materials:
  \((\text{Al/SiO}_2 \Rightarrow \text{Cu/low k})\)

- The use of Cu and low-k helps to reduce the RC delay, but the trend for local and global wires remain the same
Neighboring net (aggressor) switching results in **charge injection** into the victim net via the coupling capacitance. By consequence, voltage overshoot (or undershoot) occurs at both the input of the load and the output of the driver.
Crosstalk via Interconnect Coupling (Cont.)

- **Less interconnect spacing**
  - Capacitance increases
    - Plate capacitor formula: $C = \varepsilon * A / d$
  - Crosstalk
    - Voltage spikes couple to neighboring wires
  - Delay variation: Miller effect
Effects of Capacitance Crosstalk

- **Effects**
  - Noise
  - Timing uncertainty
  - Maximum delay
  - Slew rate degradation
Effects of Miller Effect

- Neighboring signals switch in opposite direction
  - Influences signal delay
  - Both capacitor terminals switch in opposite direction
  - Effective voltage doubled -> additional charge needed

- Related Article
  - Chapter 7.1 of Circuits, Interconnects, and Packaging for VLSI
Objective of clock distribution is to deliver clock at each node at the same time. (zero skew)

Topologies like H-tree and Grid used to deliver clock efficiently.

Clock distribution is a difficult problem.

Consumes significant resources
  - Almost 50% of total chip power.
  - Couples of metal layers used by clock and power.
  - A set of specialized design team.
Clock Distribution (Cont.)

- Mainly two topologies used to distribute clock
  - H – Tree
  - Clock Grid

- H – Tree requires significant amount of load balancing as the wire cap is comparable to the load cap.

- Grids don’t require significant amount of load balancing as wire cap is significantly higher than the load cap but they are power hungry.
Clock Drivers

- Clock distribution consumes significant amounts of power.
- Probability of switching = 1.
- Special care to be taken for the driver layout for reliability concerns.
Issues in Clock Distribution

- Symmetric trees and grids are only possible theoretically.
- Take into account reliability and productivity: electromigration, yield etc.
- Remember it’s skew that is more important to control and not delay.
- Process and power supply variation can alter the skew estimates significantly.
- How to handle multiple clocks domains?
Electromigration Effect

- Wire can tolerate only a certain amount of current density.
- Direct current for a long time causes ion movement breaking the wire over time.
- Contacts are more vulnerable to electromigration as the current tends to run through the perimeter.
- Perimeter of the contact, not the area important.
- Use of copper (heavier ions) have helped in tolerating electromigration.
Observed Electromigration Failure

A wire broken off due to electromigration

A contact (via) broken up due to electromigration

These figures are derived from *Digital integrated circuit – a design perspective*, J. Rabaey Prentice Hall
Self Heating

- Electromigration depends on the directionality of the current.
- Self-heating is just proportional to the amount of current the wire carries.
- Current flow causes the wire to get heated up and can result in providing enough energy to carriers to make them hot carriers.
- Self-heating effect can be reduced by sizing the wire (same as electromigration).
IR Voltage Drop and Ground Bounce

- IR Drop (ground bounce) increases clock skew
  - Hold time violations
- IR Drop (ground bounce) increases signal skew
  - Setup time violations

Impact timing and leads to failed silicon!!!
Power Distribution

- IR drop and electromigration are the biggest challenges for the power distribution.
- Capacitance is actually good for the power distribution.
- Many pins (at times more than 50% of total) are used due to significant amount of current.
- Finger-shaped network or grids can be used to reduce IR drop.
- About 10% of voltage is lost in IR drop.
Implications

- Communication is more expensive than computation
- Interconnect performance has become the bottleneck of the overall circuit/system performance
- Interconnect modeling and optimization has become very difficult due to the distributed nature of interconnects and their spatial and temporal interactions
- A new design paradigm/flow is needed to handle interconnect-dominated designs in the future
Design for Manufacturability: Process Variation

- Results from subwavelength lithography
- May create unexpected circuit behavior
- Requires design insensitive to process variation
Conventionally Simple VLSI Design Cycle

1. System Specification
2. Architectural Design
   - Behavioral VHDL, C
3. Functional Design
4. Logic Design
   - Structural VHDL
   - \( X = (AB*CD) + (A+D) \)
   - \( Y = (A(B+C) + AC+D) \)
5. Physical Design
6. Circuit Design
7. Fabrication
8. Packing & Testing

Figs. [©Sherwani]
Need a New Design Paradigm

Conventional Approach

Interconnection
Transistors/Cells

New Approach

Transistors/Cells
Interconnection

Interconnect-Driven Design
Appendix
Delay Modeling with Repeaters

- Without repeaters

\[ T_{50\%} = 0.4R_{\text{int}}C_{\text{int}} + 0.7(R_{\text{tr}}C_{\text{int}} + R_{\text{tr}}C_{L} + R_{\text{int}}C_{L}) \]

\[ \approx 0.7(R_{\text{tr}} + 0.4R_{\text{int}})C_{\text{int}} \quad \text{for } C_{L} \ll C_{\text{int}} \]

- With \( k \) uniform size repeaters

\[ T_{50\%} = k[0.7R_{\text{tr}}(C_{\text{int}}/k + C_{g}) + (0.4C_{\text{int}}/k + 0.7C_{g})R_{\text{int}}/k] \]

Setting \( dT/dk = 0 \),

\[ 0.4R_{\text{int}}C_{\text{int}}/k^2 = 0.7R_{\text{tr}}C_{g} \Rightarrow k = \sqrt{\frac{0.4R_{\text{int}}C_{\text{int}}}{0.7R_{\text{tr}}C_{g}}} \]
Appendix
Delay Modeling with Repeaters

- Therefore,
  \[ T_{50\%} = 0.7R_{tr}C_{int} + 1.1\sqrt{R_{tr}C_gR_{int}C_{int}} + 0.7R_{int}C_g \]

- When the W/L ratios of the transistors are increased by a factor \( h \)
  \[ T_{50\%} = k[0.7(R_{tr}/h)(C_{int}/k+hC_g)+(0.4C_{int}/k+0.7hC_g)R_{int}/k] \]

- By setting \( dT/dk = 0 \), and \( dT/dh = 0 \)
  \[ k = \sqrt{\frac{0.4R_{int}C_{int}}{0.7R_{tr}C_g}} \]
  \[ h = \sqrt{\frac{R_{tr}C_{int}}{R_{int}C_g}} \]

- Finally
  \[ T_{50\%} = 2.5\sqrt{R_{tr}C_gR_{int}C_{int}} \]