Administrative Stuff

- Office hours today: 4-5:30pm (511 Cory)
- Lab 4 this week – Mo lab to be made up next week
- Brian’s discussion session and OH this week:
  - Discussion: We 10-11am in 353 Cory
  - OH: 5-6pm in 353 Cory
- Homework #4 due 2/19
- Midterm 1 is next Thursday, Febr 26, 6:30-8pm 277 Cory
- Review session
  - Febr 25, 6-7:30pm TBD
  - Material up to Lecture 9 (Scaling) – Ch 1, Ch2 (partial), Ch3 (partial), Ch4
  - Open book, open notes
  - No labs next week
  - No new homework next week
  - Past midterms posted on the web

Impact of Technology Scaling

Last Lecture

- Last lecture
  - Buffer sizing
  - Power dissipation
- Today’s lecture
  - CMOS scaling
  - Introduction to wires
Goals of Technology Scaling

- Make things cheaper:
  - Want to sell more functions (transistors) per chip for the same money
  - Build same products cheaper, sell the same part for less money
  - Price of a transistor has to be reduced
- But also want to be faster, smaller, lower power

Technology Scaling

- Technology generation spans 2-3 years
- Benefits of scaling the dimensions by 30%:
  - Reduce gate delay by 30% (increase operating frequency by 43%)
  - Double transistor density
  - Reduce energy per transition by 65% (50% power savings @ 43% increase in frequency
- Die size used to increase by 14% per generation

Technology Generations

Technology Roadmap

International Technology Roadmap for Semiconductors 2002 data

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>DRAM ½ pitch [nm]</td>
<td>130</td>
<td>100</td>
<td>80</td>
<td>65</td>
<td>55</td>
<td>52</td>
<td>44</td>
</tr>
<tr>
<td>MPU transistor/chip</td>
<td>97M</td>
<td>153M</td>
<td>243M</td>
<td>366M</td>
<td>773M</td>
<td>1.55G</td>
<td>3.09G</td>
</tr>
<tr>
<td>Wiring levels</td>
<td>8</td>
<td>8</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>11</td>
<td>11</td>
</tr>
<tr>
<td>High-perf. phys. gate [nm]</td>
<td>65</td>
<td>45</td>
<td>32</td>
<td>25</td>
<td>18</td>
<td>11</td>
<td>9</td>
</tr>
<tr>
<td>High-perf. VDD [V]</td>
<td>1.2</td>
<td>1.0</td>
<td>0.9</td>
<td>0.7</td>
<td>0.5</td>
<td>0.4</td>
<td>0.4</td>
</tr>
<tr>
<td>Local clock [GHz]</td>
<td>1.7</td>
<td>3.1</td>
<td>5.2</td>
<td>6.7</td>
<td>11.5</td>
<td>19.3</td>
<td>28.8</td>
</tr>
<tr>
<td>High-perf. power (W)</td>
<td>130</td>
<td>150</td>
<td>170</td>
<td>190</td>
<td>218</td>
<td>251</td>
<td>288</td>
</tr>
<tr>
<td>Low-power phys. gate [nm]</td>
<td>90</td>
<td>65</td>
<td>45</td>
<td>32</td>
<td>22</td>
<td>18</td>
<td>11</td>
</tr>
<tr>
<td>Low-power VDD [V]</td>
<td>1.2</td>
<td>1.1</td>
<td>1.0</td>
<td>0.9</td>
<td>0.8</td>
<td>0.7</td>
<td>0.6</td>
</tr>
<tr>
<td>Low-power power (W)</td>
<td>2.4</td>
<td>2.8</td>
<td>3.2</td>
<td>3.3</td>
<td>7.0</td>
<td>7.0</td>
<td>3.0</td>
</tr>
</tbody>
</table>

ITRS Technology Roadmap Acceleration Continues

Technology Scaling (1)

Minimum Feature Size

Technology Scaling (2)

Number of components per chip

Technology Scaling (3)

Propagation Delay
Technology Scaling (4)

(a) Power dissipation vs. year.

(b) Power density vs. scaling factor.

From Kuroda

Technology Scaling Models

- Full Scaling (Constant Electrical Field)
  - ideal model — dimensions and voltage scale together by the same factor $S$

- Fixed Voltage Scaling
  - most common model until recently — only dimensions scale, voltages remain constant

- General Scaling
  - most realistic for today's situation — voltages and dimensions scale with different factors

Scaling Relationships for Long Channel Devices

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Relation</th>
<th>Full Scaling</th>
<th>General Scaling</th>
<th>Fixed Voltage Scaling</th>
</tr>
</thead>
<tbody>
<tr>
<td>$W$, $L$, $g_m$</td>
<td>$L$</td>
<td>$L$</td>
<td>$L$</td>
<td>$L$</td>
</tr>
<tr>
<td>$V_{DD}$, $V_T$</td>
<td>$V$</td>
<td>$V$</td>
<td>$V$</td>
<td>$V$</td>
</tr>
<tr>
<td>$S_{CH}$</td>
<td>$S$</td>
<td>$S$</td>
<td>$S$</td>
<td>$S^2$</td>
</tr>
<tr>
<td>Area/Device</td>
<td>$A_L$</td>
<td>$A_L$</td>
<td>$A_L$</td>
<td>$A_L$</td>
</tr>
<tr>
<td>$C_{ox}$</td>
<td>$C_{ox}$</td>
<td>$C_{ox}$</td>
<td>$C_{ox}$</td>
<td>$C_{ox}$</td>
</tr>
<tr>
<td>$n_{sat}$</td>
<td>$n_{sat}$</td>
<td>$n_{sat}$</td>
<td>$n_{sat}$</td>
<td>$n_{sat}$</td>
</tr>
<tr>
<td>$l_{inf}$</td>
<td>$l_{inf}$</td>
<td>$l_{inf}$</td>
<td>$l_{inf}$</td>
<td>$l_{inf}$</td>
</tr>
<tr>
<td>$q$ (electronic)</td>
<td>$q/V_L$</td>
<td>$q/V_L$</td>
<td>$q/V_L$</td>
<td>$q/V_L$</td>
</tr>
<tr>
<td>$P_{on}$</td>
<td>$P_{on}$</td>
<td>$P_{on}$</td>
<td>$P_{on}$</td>
<td>$P_{on}$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Technology Scaling (Velocity-Saturated Devices)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parameter</td>
</tr>
<tr>
<td>-----------------</td>
</tr>
<tr>
<td>$W_L$, $g_m$</td>
</tr>
<tr>
<td>$V_{DD}$, $V_T$</td>
</tr>
<tr>
<td>$S_{CH}$</td>
</tr>
<tr>
<td>Area/Device</td>
</tr>
<tr>
<td>$C_{ox}$</td>
</tr>
<tr>
<td>$n_{sat}$</td>
</tr>
<tr>
<td>$l_{inf}$</td>
</tr>
<tr>
<td>$q$ (electronic)</td>
</tr>
<tr>
<td>$P_{on}$</td>
</tr>
</tbody>
</table>

Table 3.1: Scaling Relationships for Long Channel Devices
**μProcessor Scaling**

- 2X growth in 1.96 years!

**μProcessor Power**

- 5KW
- 1.5KW
- 500W

**μProcessor Performance**

- Performance 2X/16 months
- 1 TIP (terra instructions/s)
- 30 GHz clock

**2010 Outlook**

- Size
  - No of transistors: 2 Billion
  - Die: 40*40 mm
- Power
  - 10kW!!
  - Leakage: 1/3 of total Power
Some interesting questions

- What will cause this model to break?
- When will it break?
- Will the model gradually slow down?
  - Power and power density
  - Leakage
  - Process Variation

Wires

The Wire

Interconnect Impact on Chip
Wire Models

- All-inclusive model
- Capacitance-only

Impact of Interconnect Parasitics
- Interconnect parasitics
  - reduce reliability
  - affect performance and power consumption
- Classes of parasitics
  - Capacitive
  - Resistive
  - Inductive

Nature of Interconnect

- Local Interconnect
- Global Interconnect

INTERCONNECT

Capacitance
Capacitance of Wire Interconnect

Capacitance: The Parallel Plate Model

Permittivity

<table>
<thead>
<tr>
<th>Material</th>
<th>( \varepsilon_r )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Free space</td>
<td>1</td>
</tr>
<tr>
<td>Aerogels</td>
<td>(-1.5)</td>
</tr>
<tr>
<td>Polymides (organic)</td>
<td>3.4</td>
</tr>
<tr>
<td>Silicon dioxide</td>
<td>3.9</td>
</tr>
<tr>
<td>Glass-epoxy (PC board)</td>
<td>5</td>
</tr>
<tr>
<td>Silicon Nitride (Si₃N₄)</td>
<td>7.5</td>
</tr>
<tr>
<td>Alumina (package)</td>
<td>9.5</td>
</tr>
<tr>
<td>Silicon</td>
<td>11.7</td>
</tr>
</tbody>
</table>

Fringing Capacitance

\[
c_{\text{fring}} = \frac{2 \pi \varepsilon_0 \varepsilon_r L}{\ln\left(\frac{L}{w}\right)}
\]
Fringing versus Parallel Plate

(from [Bakoglu89])

Interwire Capacitance

Impact of Interwire Capacitance

Wiring Capacitances (0.25 µm CMOS)

(from [Bakoglu89])
Next Lecture

- Wires
  - Resistance, Inductance