Topics

• Power Dissipation
• Technology Scaling
• Final Project
Where Does Power Go in CMOS?

- Dynamic Power Consumption
  Charging and Discharging Capacitors

- Short Circuit Currents
  Short Circuit Path between Supply Rails during Switching

- Leakage
  Leaking diodes and transistors
Power dissipation

- Static power dissipation
  - In theory, CMOS has no static power dissipation
  - There is a slight current (subthreshold leakage current and gate leakage current) on the order of .1-.5nA per device
  - At 5V supply voltage, .5-2.5 nW static power dissipation per device
  - Million gate chip will have .5-2.5 mW static power dissipation
Power dissipation

- Dynamic power dissipation
  - Proportional to load capacitance and frequency
  - Proportional to square of the supply voltage
    - Current trend is to reduce supply voltages to reduce power
    - Reduced supply voltage will increase delays however
  - Not dependent on device parameters
Power dissipation

- Dynamic power dissipation
  - Switching causes short bursts of current flow which will cause power dissipation

\[ V_{out}(t) = i_n(t) \]

\[ V_{out}(t) = i_p(t) \]
Power dissipation

• Dynamic power dissipation

\[ P = \frac{1}{T} \left[ \int_{0}^{T/2} i_n(t)V_{out} dt + \int_{T/2}^{T} i_p(t)(V_{DD} - V_{out}) \right] \]

\[ = \frac{1}{T} \left[ -C_L \int_{V_{DD}}^{0} V_{out} dV_{out} + C_L \int_{0}^{V_{DD}} (V_{DD} - V_{out}) dV_{out} \right] \]

\[ = \frac{C_L}{T} \left[ \frac{V_{DD}^2}{2} + \frac{V_{DD}^2}{2} \right] \]

\[ = C_L V_{DD}^2 f \]
Power dissipation

• Example

\[ V_{DD} = 5V \]
\[ f = 1GHz \]
\[ C_L = 1pF \]
\[ P = C_L V_{DD}^2 f \]
\[ = 1pf \cdot 5^2 \cdot 1GHz \]
\[ = 25mW \]
Modification for Circuits with Reduced Swing

\[ E_{01} \rightarrow = C_L \cdot V_{dd} \cdot (V_{dd} - V_t) \]

- Can exploit reduced swing to lower power (e.g., reduced bit-line swing in memory)
Power dissipation

• Short circuit current dissipation
  – Short circuit current occurs when both transistors are on temporarily
  – Proportional to the ratio of rise time to T
  – Since the rise time is usually much less than T, it can be usually ignored
Short Circuit Currents

\[ V_{in} \quad V_{dd} \quad C_L \quad V_{out} \]

---

© John A. Chandy
Dept. of Electrical and Computer Engineering
University of Connecticut
Minimizing Short-Circuit Power

- Keep the input and output rise/fall times the same (< 10% of Total Consumption) from [Veendrick84]
  \textit{(IEEE Journal of Solid-State Circuits, August 1984)}

- If $V_{dd} < V_{tn} + |V_{tp}|$ then short-circuit power can be \textit{eliminated!}
Leakage

Sub-threshold current one of most compelling issues in low-energy circuit design!
Reverse-Biased Diode Leakage

Reverse Leakage Current

\[ I_{DL} = J_S \times A \]

\[ J_S = 10-100 \text{ pA/\(\mu\text{m}^2\)} \text{ at 25 deg C for 0.25\(\mu\text{m}\) CMOS} \]

\[ J_S \text{ doubles for every 9 deg C!} \]
Subthreshold Leakage Component

- Leakage control is critical for low-voltage operation
Principles for Power Reduction

• **Prime choice: Reduce voltage!**
  - Recent years have seen an acceleration in supply voltage reduction
  - Design at very low voltages still open question (0.6 … 0.9 V by 2010!)

• Reduce switching activity

• Reduce physical capacitance
  - Device Sizing: for $F=20$
    - $f_{\text{opt}}(\text{energy})=3.53$, $f_{\text{opt}}(\text{performance})=4.47$
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

• Goals of scaling the dimensions by 30%:
  – Reduce gate delay by 30% (increase operating frequency by 43%)
  – Double transistor density
  – Reduce energy per transition by 30%

• Die size used to increase by 14% per generation

• Technology generation spans 2-3 years
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, t&lt;sub&gt;ox&lt;/sub&gt;</td>
<td>1/S</td>
<td>1/S</td>
<td>1/S</td>
<td></td>
</tr>
<tr>
<td>V&lt;sub&gt;DD&lt;/sub&gt;, V&lt;sub&gt;T&lt;/sub&gt;</td>
<td>1/S</td>
<td>1/U</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>N&lt;sub&gt;SUB&lt;/sub&gt;</td>
<td>V/W&lt;sub&gt;dep&lt;/sub&gt;²</td>
<td>S</td>
<td>S²/U</td>
<td>S²</td>
</tr>
<tr>
<td>Area/Device</td>
<td>WL</td>
<td>1/S²</td>
<td>1/S²</td>
<td>1/S²</td>
</tr>
<tr>
<td>C&lt;sub&gt;ox&lt;/sub&gt;</td>
<td>1/t&lt;sub&gt;ox&lt;/sub&gt;</td>
<td>S</td>
<td>S</td>
<td>S</td>
</tr>
<tr>
<td>C&lt;sub&gt;L&lt;/sub&gt;</td>
<td>C&lt;sub&gt;ox&lt;/sub&gt;WL</td>
<td>1/S</td>
<td>1/S</td>
<td>1/S</td>
</tr>
<tr>
<td>k&lt;sub&gt;n&lt;/sub&gt;, k&lt;sub&gt;p&lt;/sub&gt;</td>
<td>C&lt;sub&gt;ox&lt;/sub&gt;W/L</td>
<td>S</td>
<td>S</td>
<td>S</td>
</tr>
<tr>
<td>I&lt;sub&gt;av&lt;/sub&gt;</td>
<td>k&lt;sub&gt;n,p&lt;/sub&gt;V²</td>
<td>1/S</td>
<td>S/U²</td>
<td>S</td>
</tr>
<tr>
<td>t&lt;sub&gt;p&lt;/sub&gt; (intrinsic)</td>
<td>C&lt;sub&gt;L&lt;/sub&gt;V/I&lt;sub&gt;av&lt;/sub&gt;</td>
<td>1/S</td>
<td>U/S²</td>
<td>1/S²</td>
</tr>
<tr>
<td>P&lt;sub&gt;av&lt;/sub&gt;</td>
<td>C&lt;sub&gt;L&lt;/sub&gt;V²/t&lt;sub&gt;p&lt;/sub&gt;</td>
<td>1/S²</td>
<td>S/U³</td>
<td>S</td>
</tr>
<tr>
<td>PDP</td>
<td>C&lt;sub&gt;L&lt;/sub&gt;V²</td>
<td>1/S³</td>
<td>1/SU²</td>
<td>1/S</td>
</tr>
</tbody>
</table>
## Transistor Scaling (velocity-saturated 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>$W$, $L$, $t_{ox}$</td>
<td>$1/S$</td>
<td>$1/S$</td>
<td>$1/S$</td>
<td>$1/S$</td>
</tr>
<tr>
<td>$V_{DD}$, $V_T$</td>
<td>$1/S$</td>
<td>$1/U$</td>
<td>$1$</td>
<td></td>
</tr>
<tr>
<td>$N_{SUB}$</td>
<td>$V/W_{depl}^2$</td>
<td>$S$</td>
<td>$S^2/U$</td>
<td>$S^2$</td>
</tr>
<tr>
<td>Area/Device</td>
<td>$WL$</td>
<td>$1/S^2$</td>
<td>$1/S^2$</td>
<td>$1/S^2$</td>
</tr>
<tr>
<td>$C_{ox}$</td>
<td>$1/t_{ox}$</td>
<td>$S$</td>
<td>$S$</td>
<td>$S$</td>
</tr>
<tr>
<td>$C_{gate}$</td>
<td>$C_{ox}WL$</td>
<td>$1/S$</td>
<td>$1/S$</td>
<td>$1/S$</td>
</tr>
<tr>
<td>$k_f$, $k_p$</td>
<td>$C_{ox}WL$</td>
<td>$S$</td>
<td>$S$</td>
<td>$S$</td>
</tr>
<tr>
<td>$V_{on}$</td>
<td>$V/L_{on}$</td>
<td>$1$</td>
<td>$1$</td>
<td>$1$</td>
</tr>
<tr>
<td>$I_{sat}$</td>
<td>$L_{sat}V$</td>
<td>$1/S^2$</td>
<td>$1/L^2$</td>
<td>$1$</td>
</tr>
<tr>
<td>$P$</td>
<td>$L_{sat}V$</td>
<td>$1/S^2$</td>
<td>$1/L^2$</td>
<td>$1$</td>
</tr>
</tbody>
</table>
Technology Generations

Table 2. Time overlap of semiconductor technology generations.

<table>
<thead>
<tr>
<th></th>
<th>95</th>
<th>96</th>
<th>97</th>
<th>98</th>
<th>99</th>
<th>00</th>
<th>01</th>
<th>02</th>
<th>03</th>
<th>04</th>
<th>05</th>
<th>06</th>
<th>07</th>
<th>08</th>
<th>09</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>350 nm</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>250 nm</td>
<td>-1</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>180 nm</td>
<td>-3</td>
<td>-2</td>
<td>-1</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>150 nm</td>
<td>-5</td>
<td>-4</td>
<td>-3</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>130 nm</td>
<td>-7</td>
<td>-6</td>
<td>-5</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100 nm</td>
<td>-9</td>
<td>-8</td>
<td>-7</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>-1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>70 nm</td>
<td>-11</td>
<td>-10</td>
<td>-9</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>-4</td>
<td>-3</td>
<td>-2</td>
<td>-1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50 nm</td>
<td>-13</td>
<td>-12</td>
<td>-11</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>-7</td>
<td>-6</td>
<td>-5</td>
<td>-4</td>
<td>-3</td>
<td>-2</td>
<td>-1</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Technology Generations
# Technology Evolution (2000 data)

## International Technology Roadmap for Semiconductors

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Technology node [nm]</td>
<td>180</td>
<td>130</td>
<td>90</td>
<td>60</td>
<td>40</td>
<td>30</td>
<td></td>
</tr>
<tr>
<td>Supply [V]</td>
<td>1.5-1.8</td>
<td>1.5-1.8</td>
<td>1.2-1.5</td>
<td>0.9-1.2</td>
<td>0.6-0.9</td>
<td>0.5-0.6</td>
<td>0.3-0.6</td>
</tr>
<tr>
<td>Wiring levels</td>
<td>6-7</td>
<td>6-7</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>9-10</td>
<td>10</td>
</tr>
<tr>
<td>Max frequency [GHz],Local-Global</td>
<td>1.2</td>
<td>1.6-1.4</td>
<td>2.1-1.6</td>
<td>3.5-2</td>
<td>7.1-2.5</td>
<td>11-3</td>
<td>14.9-3.6</td>
</tr>
<tr>
<td>Max $\mu$P power [W]</td>
<td>90</td>
<td>106</td>
<td>130</td>
<td>160</td>
<td>171</td>
<td>177</td>
<td>186</td>
</tr>
<tr>
<td>Bat. power [W]</td>
<td>1.4</td>
<td>1.7</td>
<td>2.0</td>
<td>2.4</td>
<td>2.1</td>
<td>2.3</td>
<td>2.5</td>
</tr>
</tbody>
</table>

Technology Evolution (1999)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Channel length (μm)</td>
<td>0.4</td>
<td>0.3</td>
<td>0.25</td>
<td>0.18</td>
<td>0.13</td>
<td>0.1</td>
</tr>
<tr>
<td>Gate oxide (nm)</td>
<td>12</td>
<td>7</td>
<td>6</td>
<td>4.5</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>$V_{DD}$ (V)</td>
<td>3.3</td>
<td>2.2</td>
<td>2.2</td>
<td>1.5</td>
<td>1.5</td>
<td>1.5</td>
</tr>
<tr>
<td>$V_T$ (V)</td>
<td>0.7</td>
<td>0.7</td>
<td>0.7</td>
<td>0.6</td>
<td>0.6</td>
<td>0.6</td>
</tr>
<tr>
<td>NMOS $I_{Dsat}$ (mA/μm) (@ $V_{GS} = V_{DD}$)</td>
<td>0.35</td>
<td>0.27</td>
<td>0.31</td>
<td>0.21</td>
<td>0.29</td>
<td>0.33</td>
</tr>
<tr>
<td>PMOS $I_{Dsat}$ (mA/μm) (@ $V_{GS} = V_{DD}$)</td>
<td>0.16</td>
<td>0.11</td>
<td>0.14</td>
<td>0.09</td>
<td>0.13</td>
<td>0.16</td>
</tr>
</tbody>
</table>
Technology Scaling (1)

Minimum Feature Size
Technology Scaling (2)

Number of components per chip

- Processor Transistors
- Memory Transistors

Year:
- 1990
- 1991
- 1992
- 1993
- 1994
- 1995
- 1996
- 1997
- 1998
- 1999
- 2000

Technology:
- 0.5 um technology
- 0.35 um
- 0.25 um
- 0.18 um
Technology Scaling (3)

$\tau_p$ decreases by 13%/year
50% every 5 years!

Propagation Delay

gate delay (ns)
Technology Scaling (4)

(a) Power dissipation vs. year.

(b) Power density vs. scaling factor.

From Kuroda
µProcessor Scaling

2X Growth in 1.96 Years!

P. Gelsinger: µProcessors for the New Millenium, ISSCC 2001
µProcessor Power

P. Gelsinger: µProcessors for the New Millenium, ISSCC 2001
2010 Outlook

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

- Size
  - No of transistors: 2 Billion
  - Die: 40*40 mm

- Power
  - 10kW!!
  - Leakage: 1/3 active Power

P.Gelsinger: μProcessors for the New Millennium, ISSCC 2001
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
Final Project

- Teams of two
- Choose your own project
- If you want to fabricate chip, you are limited to a 1.5 mm square - roughly 5-10000 transistors
Final Project

- Spring 2003 - Search Engine Processor
Final Project

• Spring 2004 - Encoder/Decoder
Final Project

• Important Dates
  – Proposal due March 3rd
  – Architecture due March 17th
  – Logic Design due March 31st
  – Demonstrations April 26-28th
  – Final Project Report due April 29th
  – Presentation April 28th
Next Class

• Exam 1
  – Lectures 1-10
  – HW1-3
  – Chapters 1-6