CMPEN 411
VLSI Digital Circuits

Lecture 02: Design Metrics

Kyusun Choi

Overview of Last Lecture

- Digital integrated circuits experience exponential growth in complexity (Moore’s law) and performance.
- Design in the deep submicron (DSM) era creates new challenges:
  - Devices become somewhat different
  - Global clocking becomes more challenging
  - Interconnect effects play a more significant role
  - Power dissipation may be the limiting factor
- Our goal in this class will be to design digital integrated circuits in 0.5um CMOS technology and understand digital integrated circuits in advanced technologies, below 180 nanometer.
- Today we look at some basic design metrics.
Fundamental Design Metrics

- Functionality
- Cost
  - NRE (fixed) costs - design effort
  - RE (variable) costs - cost of material, parts, assembly, test
- Reliability, robustness
  - Noise margins
  - Noise immunity
- Performance/Power
  - Speed (delay)
  - Power consumption; energy
- Time-to-market
How a chip is manufactured?

1) Starting substrate - silicon wafer (purchased).

2) Wafer fabrication - fabricate IC’s on the wafer.

3) Wafer sort/test - test each IC, mark bad IC’s.

4) Packaging - assemble IC’s into packages.

5) Mark & class/final test - mark and final test packaged product.
Cost of Integrated Circuits

- **NRE (non-recurring engineering) costs**
  - Fixed cost to produce the design
    - design effort
    - design verification effort
    - mask generation
  - Influenced by the design complexity and designer productivity
  - More pronounced for small volume products

- **Recurring costs – proportional to product volume**
  - silicon processing, material
    - also proportional to chip area
  - assembly (packaging)
  - test

\[
\text{cost per IC} = \text{variable cost per IC} + \frac{\text{fixed cost}}{\text{volume}}
\]
NRE Cost is Increasing

"The club of people who can afford an extreme sub-micron ASIC or COTS design is getting pretty exclusive."

Ron Wilson, EE Times (May 2000)

70nm ASICs will have $4M NRE

www.InnovationRevolution.com
Silicon Wafer

From http://www.amd.com
Silicon Wafer

300mm wafer and Pentium 4 IC. Photos courtesy of Intel.
Recurring Costs

variable cost = \frac{\text{cost of die} + \text{cost of die test} + \text{cost of packaging}}{\text{final test yield}}

\text{cost of die} = \frac{\text{cost of wafer}}{\text{dies per wafer} \times \text{die yield}}

diameter dies per wafer = \frac{\pi \times (\text{wafer diameter}/2)^2}{\text{die area}} - \frac{\pi \times \text{wafer}}{\sqrt{2} \times \text{die area}}

die yield = (1 + \frac{\text{defects per unit area} \times \text{die area}}{\alpha})^{-\alpha}
Defect Density Trends

References [3] and [4] include data from multiple wafer fabrication sites. The best reported data from each year is used for this chart. Furthermore, for each data series later year data where trailing edge fabrication facilities have apparently entered the data set driving up defect densities has been filtered out. Leading edge fabs that drove the initial performance are believed to have moved on to smaller geometries by the time the trailing edge fabs have entered production at each linewidth.
Yield Example

- Example
  - wafer size of 12 inches, die size of 2.5 cm\(^2\), 1 defects/cm\(^2\), \(\alpha = 3\) (measure of manufacturing process complexity)
  - 252 dies/wafer (remember, wafers round & dies square)
  - die yield of 16%
  - \(252 \times 16\% = \text{only 40 dies/wafer die yield!}\)

- Die cost is strong function of die area
  - proportional to the third or fourth power of the die area
### Examples of Cost Metrics (1994)

<table>
<thead>
<tr>
<th>Chip</th>
<th>Metal layers</th>
<th>Line width</th>
<th>Wafer cost</th>
<th>Defects/cm²</th>
<th>Area (mm²)</th>
<th>Dies/wafer</th>
<th>Yield</th>
<th>Die cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>386DX</td>
<td>2</td>
<td>0.90</td>
<td>$900</td>
<td>1.0</td>
<td>43</td>
<td>360</td>
<td>71%</td>
<td>$4</td>
</tr>
<tr>
<td>486DX2</td>
<td>3</td>
<td>0.80</td>
<td>$1200</td>
<td>1.0</td>
<td>81</td>
<td>181</td>
<td>54%</td>
<td>$12</td>
</tr>
<tr>
<td>PowerPC 601</td>
<td>4</td>
<td>0.80</td>
<td>$1700</td>
<td>1.3</td>
<td>121</td>
<td>115</td>
<td>28%</td>
<td>$53</td>
</tr>
<tr>
<td>HP PA 7100</td>
<td>3</td>
<td>0.80</td>
<td>$1300</td>
<td>1.0</td>
<td>196</td>
<td>66</td>
<td>27%</td>
<td>$73</td>
</tr>
<tr>
<td>DEC Alpha</td>
<td>3</td>
<td>0.70</td>
<td>$1500</td>
<td>1.2</td>
<td>234</td>
<td>53</td>
<td>19%</td>
<td>$149</td>
</tr>
<tr>
<td>Super SPARC</td>
<td>3</td>
<td>0.70</td>
<td>$1700</td>
<td>1.6</td>
<td>256</td>
<td>48</td>
<td>13%</td>
<td>$272</td>
</tr>
<tr>
<td>Pentium</td>
<td>3</td>
<td>0.80</td>
<td>$1500</td>
<td>1.5</td>
<td>296</td>
<td>40</td>
<td>9%</td>
<td>$417</td>
</tr>
</tbody>
</table>
## Examples of Wafer Cost

<table>
<thead>
<tr>
<th>Wafer size</th>
<th>Line-width (µm)</th>
<th></th>
<th>Mask layers</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>18</td>
<td>20</td>
<td>22</td>
<td>24</td>
<td>26</td>
</tr>
<tr>
<td>200mm</td>
<td>0.25</td>
<td></td>
<td>$890</td>
<td>$980</td>
<td>$1,070</td>
<td>$1,155</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>0.18</td>
<td></td>
<td>-</td>
<td>$1,320</td>
<td>$1,440</td>
<td>$1,565</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>0.13</td>
<td></td>
<td>-</td>
<td>-</td>
<td>$1,815</td>
<td>$1,970</td>
<td>$2,130</td>
</tr>
<tr>
<td>300mm</td>
<td>0.13</td>
<td></td>
<td>-</td>
<td>-</td>
<td>$2,500</td>
<td>$2,690</td>
<td>$2,890</td>
</tr>
<tr>
<td></td>
<td>0.09</td>
<td></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>$2,860</td>
<td>$3,065</td>
</tr>
</tbody>
</table>

Source: Icknowledge.com
**Examples of Cost Metrics  Intel Pentium4**

<table>
<thead>
<tr>
<th>Fab facility</th>
<th>200mm - 130nm - CMOS logic - 1 layer poly - 8 layer copper - FSG ILD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fab capacity (wafers/month)</td>
<td>30,000 United States</td>
</tr>
<tr>
<td>Fab utilization</td>
<td>90%</td>
</tr>
<tr>
<td><strong>Wafer costs</strong></td>
<td></td>
</tr>
<tr>
<td>Material</td>
<td>Labor</td>
</tr>
<tr>
<td>Unyielded wafer costs</td>
<td>$85.00</td>
</tr>
<tr>
<td>Wafer yield</td>
<td>97.7%</td>
</tr>
<tr>
<td>Yielded wafer costs</td>
<td>$87.00</td>
</tr>
<tr>
<td><strong>Die per wafer</strong></td>
<td></td>
</tr>
<tr>
<td>Area (mm²)</td>
<td>146.0</td>
</tr>
<tr>
<td><strong>Wafer sort costs</strong></td>
<td></td>
</tr>
<tr>
<td>Sort 1</td>
<td>Burn-in</td>
</tr>
<tr>
<td>$0.939</td>
<td>$0.000</td>
</tr>
<tr>
<td><strong>Package</strong></td>
<td></td>
</tr>
<tr>
<td>FCPGA - 35mm x 35mm x 478pins - Intel 130nm Pentium 4 with 512Kb cache</td>
<td></td>
</tr>
<tr>
<td>Package Cost</td>
<td>Yield</td>
</tr>
<tr>
<td>$6.797</td>
<td>95.3%</td>
</tr>
<tr>
<td><strong>Class test</strong></td>
<td></td>
</tr>
<tr>
<td>Class 1</td>
<td>Burn-in</td>
</tr>
<tr>
<td>$1.031</td>
<td>$0.080</td>
</tr>
</tbody>
</table>

**Final cost:** 25.57\$/chip  
**source:** www.icknowledge.com
Reliability
Noise in Digital Integrated Circuits

- **Noise** – unwanted variations of voltages and currents at the logic nodes

- From two wires placed side by side
  - capacitive coupling
    - voltage change on one wire can influence signal on the neighboring wire
    - cross talk
  - inductive coupling
    - current change on one wire can influence signal on the neighboring wire

- From noise on the power and ground supply rails
  - can influence signal levels in the gate
Example of Capacitive Coupling

- Signal wire glitches as large as 80% of the supply voltage will be common due to crosstalk between neighboring wires as feature sizes continue to scale.

**Crosstalk vs. Technology**

From Dunlop, Lucent, 2000
Static Gate Behavior

- Steady-state parameters of a gate – *static behavior* – tell how robust a circuit is with respect to both variations in the manufacturing process and to noise disturbances.

- Digital circuits perform operations on Boolean variables $x \in \{0,1\}$

- A logical variable is associated with a *nominal voltage level* for each logic state

  $1 \iff V_{\text{OH}}$ and $0 \iff V_{\text{OL}}$

- Difference between $V_{\text{OH}}$ and $V_{\text{OL}}$ is the logic or *signal swing* $V_{\text{sw}}$
DC Operation
Voltage Transfer Characteristics (VTC)

- Plot of output voltage as a function of the input voltage

\[ V_{OL} = f(V_{IH}) \]
\[ V_{OH} = f(V_{IL}) \]

Switching Threshold: \( V_M \)
Mapping Logic Levels to the Voltage Domain

- The regions of acceptable high and low voltages are delimited by $V_{IH}$ and $V_{IL}$ that represent the points on the VTC curve where the gain $= -1$.
Noise Margins

- For robust circuits, want the “0” and “1” intervals to be as large as possible

\[ \text{Noise Margin High (NM}_H = \text{V}_{\text{OH}} - \text{V}_{\text{IH}} \]

\[ \text{Noise Margin Low (NM}_L = \text{V}_{\text{IL}} - \text{V}_{\text{OL}} \]

- Large noise margins are desirable, but not sufficient …
The Regenerative Property

- A gate with regenerative property ensures that a disturbed signal converges back to a nominal voltage level.
Conditions for Regeneration

To be regenerative, the VTC must have a transient region with a gain greater than 1 (in absolute value) bordered by two valid zones where the gain is smaller than 1. Such a gate has two stable operating points.
Noise Immunity

- Noise margin expresses the ability of a circuit to overpower a noise source
  - noise sources: supply noise, cross talk, interference, offset

- Absolute noise margin values are deceptive
  - a floating node is more easily disturbed than a node driven by a low impedance (in terms of voltage)

- Noise immunity expresses the ability of the system to process and transmit information correctly in the presence of noise

- For good noise immunity, the signal swing (i.e., the difference between $V_{\text{OH}}$ and $V_{\text{OL}}$) and the noise margin have to be large enough to overpower the impact of fixed sources of noise
Fan-In and Fan-Out

- **Fan-out** – number of load gates connected to the output of the driving gate
  - gates with large fan-out are slower

- **Fan-in** – the number of inputs to the gate
  - gates with large fan-in are bigger and slower
The Ideal Inverter

- The ideal gate should have
  - infinite gain in the transition region
  - a gate threshold located in the middle of the logic swing
  - high and low noise margins equal to half the swing

\[ g = -\infty \]

\[ R_i = \infty \]
\[ R_o = 0 \]
\[ \text{Fanout} = \infty \]
\[ \text{NM}_{H} = \text{NM}_{L} = \frac{VDD}{2} \]
Delay Definitions

Propagation delay?

signal slopes?

input waveform

output waveform
Modeling Propagation Delay

- Model circuit as first-order RC network

\[
v_{\text{out}}(t) = (1 - e^{-t/\tau})V
\]

where \( \tau = RC \)

Time to reach 50% point is
\[
t = \ln(2) \tau = 0.69 \tau
\]

Time to reach 90% point is
\[
t = \ln(9) \tau = 2.2 \tau
\]

- Matches the delay of an inverter gate
Power and Energy Dissipation

- Power consumption: how much energy is consumed per operation and how much heat the circuit dissipates
  - supply line sizing (determined by peak power)
    \[ P_{\text{peak}} = V_{\text{dd}} i_{\text{peak}} \]
  - battery lifetime (determined by average power dissipation)
    \[ p(t) = v(t)i(t) = V_{\text{dd}} i(t) \quad P_{\text{avg}} = \frac{1}{T} \int p(t) \, dt = \frac{V_{\text{dd}}}{T} \int i_{\text{dd}}(t) \, dt \]
  - packaging and cooling requirements

- Two important components: static and dynamic

\[
E \text{ (joules)} = C_L V_{\text{dd}}^2 P_{0 \rightarrow 1} + t_{sc} V_{\text{dd}} I_{\text{peak}} P_{0 \rightarrow 1} + V_{\text{dd}} I_{\text{leakage}}
\]

\[
P \text{ (watts)} = C_L V_{\text{dd}}^2 f_{0 \rightarrow 1} + t_{sc} V_{\text{dd}} I_{\text{peak}} f_{0 \rightarrow 1} + V_{\text{dd}} I_{\text{leakage}}
\]
Power and Energy Dissipation

- Propagation delay and the power consumption of a gate are related.

- Propagation delay is (mostly) determined by the speed at which a given amount of energy can be stored on the gate capacitors.
  - the faster the energy transfer (higher power dissipation) the faster the gate.

- For a given technology and gate topology, the product of the power consumption and the propagation delay is a constant.
  - Power-delay product (PDP) – energy consumed by the gate per switching event.

- An ideal gate is one that is fast and consumes little energy, so the ultimate quality metric is.
  - Energy-delay product (EDP) = power-delay $^2$.
Digital integrated circuits have come a long way and still have quite some potential left for the coming decades

Some interesting challenges ahead
- Getting a clear perspective on the challenges and potential solutions is the purpose of this course

Understanding the design metrics that govern digital design is crucial
- Cost, reliability, speed, power and energy dissipation
Next Lecture and Reminders

- Next lecture
  - MOS transistor
    - Reading assignment – Rabaey et al, 3.1-3.3.2
    - I will not be covering 3.2 in class (EE 310 material)
Design Abstraction Levels

SYSTEM

MODULE

GATE

CIRCUIT

DEVICE

\[ V_{in} \rightarrow V_{out} \]

\[ n^+ \rightarrow G \rightarrow n^+ \]
Device: The MOS Transistor

CROSS-SECTION of NMOS Transistor
Circuit: The CMOS Inverter

V_{DD} - V_{in} - V_{out} - C_L