Digital Integrated Circuits – EECS 312
http://robertdick.org/eecs312/

Teacher: Robert Dick
Office: 2417-E EECS
Email: dickrp@umich.edu
Phone: 734–763–3329
Cellphone: 847–530–1824

GSI: Shengshou Lu
Office: 2725 BBB
Email: luss@umich.edu

Lab 3

- Input inverters.
- Implications of sizing on energy consumption.

Derive and explain.

Inverter chain delay optimization

**Given**
- Size (width) of first inverter in chain,
- Driven load,
- Transistors are minimal length, and
- \( W_d/W_n = 2 \) approximately balances \( t_{\text{pHL}} \) and \( t_{\text{pLH}} \).

**Find**
- Optimal number of inverters in chain and
- Optimal size (width) of each inverter to minimize chain delay.

Review

- Design a non-trivial logic gate.
- What happens to inverter delay as the driving MOSFET widths are increased?
- What happens to inverter delay as the driven MOSFET widths are increased?
- What impact does non-instantaneous rise/fall time have on the propagation delay for the subsequent logic stage?

Dependence of delay on width (R)

- Fix \( R_L C_L \) and vary \( W \).
- Eventually, self-loading dominates.

Intuition

- Given two inverters (first fixed) and a large load (\( C_L \)), how should the second be sized to minimize delay?
- \( C_{G2} = C_{G1} \) (minimal)?
- \( C_{G2} > C_L \)?
- \( C_{G2} = C_L \)?
- Some other setting?
- Why?
Derivation I

Let \( W = W_n = W_p/2 \)
\( R = R_p = R_n \)
\( T_{pHL} = T_{pLH} = 0.69RC_L \)
\( C_i = 3/W_{init}C_{unit} \)
\( t_p = 0.69(R(C_i + C_L)) \)

Derivation II

Consider the impact of scaling factor \( S \).
\[
t_p = 0.69 \left( \frac{R}{S}SC_{int} \left( 1 + \frac{C_L}{SC_{int}} \right) \right)
\]
\[
t_p = 0.69RC_{int} \left( 1 + \frac{C_L}{SC_{int}} \right)
\]
\[
t_p = t_{p0} \left( 1 + \frac{C_L}{SC_{int}} \right)
\]
\( t_{p0} \): Intrinsic delay.
- Scaling doesn’t impact intrinsic delay.
- Scaling does impact total delay.
- \( t_p \to t_{p0} \) as \( S \to \infty \).
- Diminishing returns with increasing \( S \).

Consider chain of inverters I

\[
t_{p,chain} = t_{p1} + t_{p2} + \cdots + t_{pn} \]
\[
t_{pi} \approx t_{p0} \left( 1 + \frac{C_{g,i+1}}{\gamma C_{g,i}} \right)
\]
\[
t_{p,chain} = \sum_{i=1}^{N} t_{pi}
\]
\[
t_{p,chain} = t_{p0} \sum_{i=1}^{N} \left( 1 + \frac{C_{g,i+1}}{\gamma C_{g,i}} \right)
\]
Given that
\( C_{g,N+1} = C_L \)

Consider chain of inverters II

and
\[
\gamma = \frac{C_{int}}{C_g} \approx 1 \text{ (technology-dependent constant)}.
\]

Sketch of derivation

- For each \( i \), find \( \sigma \) \( t_{p,chain} \) \( / \sigma C_{g,i} \).
- Solve for \( \sigma t_{p,chain} / \sigma C_{g,i} = 0 \), \( \forall i = 1 \cdots N \).
- Result is \( \frac{C_{g,i+1}}{C_{g,i}} = \frac{C_{g,i}}{C_{g,i-1}} \).
- Each stage size geometric mean of previous and next:
  \( C_{g,i} = \sqrt{C_{g,i-1}C_{g,i+1}} \).
- Constant factor relates sizing of all adjacent gate pairs.
- Each stage has same delay.

Sizing for optimal inverter chain delay

- Optimal stage-wise sizing factor: \( \frac{C_{g,i}}{\sqrt{C_{g,i-1}/\gamma}} \).
- Minimum path delay: \( t_{p,chain} = Nt_{p0} \left( 1 + \frac{C_L}{\sqrt{C_{g,i}/\gamma}} \right) \)
Example of inverter sizing

Given
- $C_L = 16C_1$
- $N = 4$

Per-stage scaling factor: $\sqrt[4]{16C_1/C_1} = 2$

Optimizing $N_I$

Let

$$\Phi = \frac{C_L}{C_g1}$$

$$t_{p,\text{chain}} = Nt_0\left(1 + \frac{\sqrt[4]{\Phi}}{\gamma}\right)$$

$$t_{p,\text{chain}} \frac{d}{dN} = \gamma + \frac{1}{\gamma} - \frac{\sqrt[4]{\Phi} \ln(\Phi)}{N}$$

Optimizing $N_{II}$

Set this to zero.

Let $\phi = \sqrt[4]{\Phi}$

$$0 = \gamma + \phi \ln(\phi^N) - \frac{\phi}{N} \ln(\phi^N)$$

$$0 = \frac{\gamma}{\phi} + 1 - \frac{\ln(\phi^N)}{N}$$

$$0 = \frac{\gamma}{\phi} + 1 - \frac{N\ln(\phi)}{N}$$

$$0 = \frac{\gamma}{\phi} + 1 - \ln(\phi)$$

$$\ln(\phi) = \frac{\gamma}{\phi} + 1$$

Optimal stage sizing factor

- Optimal tapering factor for $\gamma = 0$: $e \approx 2.7$
- 3.6 for $\gamma = 1$

### Table

<table>
<thead>
<tr>
<th>$\Phi$</th>
<th>Unbuffered</th>
<th>$N = 2$</th>
<th>Optimal $N$</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>11</td>
<td>8.3</td>
<td>8.3</td>
</tr>
<tr>
<td>100</td>
<td>101</td>
<td>22</td>
<td>16.5</td>
</tr>
<tr>
<td>1,000</td>
<td>1,001</td>
<td>65</td>
<td>24.8</td>
</tr>
<tr>
<td>10,000</td>
<td>10,001</td>
<td>202</td>
<td>33.1</td>
</tr>
</tbody>
</table>
Buffering example

<table>
<thead>
<tr>
<th>N</th>
<th>f</th>
<th>Lp</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>64</td>
<td>65</td>
</tr>
<tr>
<td>2</td>
<td>8</td>
<td>18</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>15</td>
</tr>
<tr>
<td>4</td>
<td>2.8</td>
<td>15.3</td>
</tr>
</tbody>
</table>

### Upcoming topics
- Interconnect.
- Alternative logic design styles.

### Power consumption in synchronous CMOS

\[ P = P_{\text{SWITCH}} + P_{\text{SHORT}} + P_{\text{LEAK}} \]

\[ P_{\text{SWITCH}} = C \cdot V_{\text{DD}}^2 \cdot f \cdot A \]

\[ P_{\text{SHORT}} = \frac{b}{12} \left( V_{\text{DD}} - 2 \cdot V_T \right)^3 \cdot f \cdot A \cdot t \]

\[ P_{\text{LEAK}} = V_{\text{DD}} \cdot (I_{\text{SUB}} + I_{\text{GATE}} + I_{\text{JUNCTION}} + I_{\text{GIDL}}) \]

- C: total switched capacitance
- V_{\text{DD}}: high voltage
- f: switching frequency
- A: switching activity
- b: MOS transistor gain
- V_T: threshold voltage
- t: rise/fall time of inputs

\[ P_{\text{SHORT}} \text{ usually } \leq 10\% \text{ of } P_{\text{SWITCH}} \]

Smaller as \( V_{\text{DD}} \rightarrow V_T \)

\[ A < 0.5 \text{ for combinational nodes, 1 for clocked nodes.} \]

### Reasons for power consumption
- Dynamic
  - Charging and discharging RC loads.
  - \( E_{\text{dyn}} = C \cdot V_{\text{DD}}^2 \cdot f \).
  - \( P_{\text{dyn}} = C \cdot V_{\text{DD}}^2 \cdot f \).
  - \( \text{But } f \propto V_{\text{DD}} \).
  - So \( P_{\text{dyn}} = C \cdot V_{\text{DD}}^3 \).
- Static
  - Sub-threshold leakage.
  - Gate leakage.
- Short-circuit: Pull-up and pull-down networks briefly both on.

### Fixed voltage charging

\[ E_{\text{step}} = \int_{t=0}^{\infty} V_R(t) l_R(t) \, dt \]

\[ E_{\text{R}} = \int_{t=0}^{\infty} V_R(t) \frac{V_R(t)}{R} \, dt \]

\[ E_{\text{R}}^\text{step} = \int_{t=0}^{\infty} V_{\text{DD}} e^{-t/RC} \frac{V_{\text{DD}}}{R} e^{-t/RC} \, dt \]

\[ E_{\text{R}}^\text{step} = \frac{V_{\text{DD}}^2}{R} \int_{t=0}^{\infty} e^{-t/RC} \left( e^{-t/RC} \right)^\infty \]

\[ E_{\text{R}}^\text{step} = \frac{V_{\text{DD}}^2}{2} (0 - 1) \]

\[ E_{\text{R}}^\text{step} = \frac{V_{\text{DD}}^2}{2} \]
Fixed current charging I

\[ E_{R}^{\text{ramp}} = \int_{t=0}^{\infty} V_R(t) \, l_R(t) \, dt \]

Let \( T \) be the voltage ramp duration, \( l_R \) is fixed. \( V_R \) is fixed.

Fixed current charging II

\[ \Delta V = \frac{\Delta q}{C} \]
\[ V_{DD} = \frac{l_R \, T}{C} \]
\[ I_R = \frac{CV_{DD}}{T} \]
\[ V_R = Rl_R \]
\[ E_R^{\text{ramp}} = \int_{t=0}^{T} l_R(t) \, V_R(t) \, dt \]
\[ E_R^{\text{ramp}} = \int_{t=0}^{T} CV_{DD} \, RCV_{DD} \, dt \]
\[ E_R^{\text{ramp}} = \frac{R V_{DD}^2 C^2}{T^2} \int_{t=0}^{T} 1 \, dt \]

Fixed current charging III

\[ E_R^{\text{ramp}} = \frac{V_{DD}^2 C^2 R}{T^2} \]
\[ E_R^{\text{ramp}} = \frac{V_{DD}^2 C^2 R}{T} \]
\[ E_R^{\text{ramp}} = \frac{V_{DD}^2 C^2 2RC}{T} \]

Break-even point

\[ E_{\text{step}} = \frac{E_{R}^{\text{ramp}}}{2} \]
\[ V_{DD}^2 C = \frac{V_{DD}^2 C \, 2RC}{T} \]
\[ 1 = \frac{2RC}{T} \]
\[ T = 2RC \]

- Properly controlling \( V_R(t) \).
- Performance.
- In limit, permits reversible computation with low/no power consumption during charging and discharging.

Charging methods summary

- \( l(t) \) influences energy consumption for same change in \( V \).
- In theory, keeping voltage differences very small can permit extremely low-power operation.
- Leakage, current control, and preserving reversibility make this challenging.

Capacitive load modeling

- Simplified Model

Fanout

Interconnect

C_{M1} C_{M2}
Interconnect modeling

![Interconnect Circuit Diagram]

Homework

Permittivity ($k$)

<table>
<thead>
<tr>
<th>Material</th>
<th>$\epsilon$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vacuum</td>
<td>1</td>
</tr>
<tr>
<td>Aerogels</td>
<td>~1.5</td>
</tr>
<tr>
<td>Polyimides 3–4</td>
<td>3–4</td>
</tr>
<tr>
<td>SiO$_2$</td>
<td>3.9</td>
</tr>
<tr>
<td>Glass-epoxy</td>
<td>5</td>
</tr>
<tr>
<td>Si$_3$N$_4$</td>
<td>7.5</td>
</tr>
<tr>
<td>Alumina</td>
<td>9.5</td>
</tr>
<tr>
<td>Silicon</td>
<td>11.7</td>
</tr>
</tbody>
</table>

Fringing

\[
C_{\text{wire}} = C_{\text{pp}} + C_{\text{fringe}} = \frac{W\epsilon_{\text{ox}}}{t_{\text{ox}}} + \frac{2\pi\epsilon_{\text{ox}}}{\log\left(t_{\text{ox}}/H\right)}
\]

Trends in interconnect design

- More metal layers.
- Lower aspect ratios.
  - More coupling.
- Smaller transistors, but similar-length global interconnect.

Upcoming topics

- Alternative logic design styles.
- Latches and flip-flops.
- Memories.
Homework assignment

- 10 October: Homework 2.
- 22 October: Lab 3.