ESE370: Circuit-Level Modeling, Design, and Optimization for Digital Systems

Lec 17: October 16, 2015
Energy Optimization
Ratioed Logic
Previously

- Three components of power
  \[ P_{tot} = P_{static} + P_{dyn} + P_{sc} \]

- Restoration and Noise Margins

- CMOS Gates
  - Drive rail-to-rail
  - Only one transistor turned on in steady state
    - Only subthreshold current in steady state
Today

- Energy Tradeoffs
  - Reduce dynamic power

- Ratioed Logic
  - Break all the rules… (lose our nice properties)
    - Not rail-to-rail signals, steady-state-current…
  - Correctness
  - Performance
  - Power
  - Implications
Dynamic Power
Total Power

- \( P_{\text{tot}} = P_{\text{static}} + P_{\text{sc}} + P_{\text{dyn}} \)

- \( P_{\text{dyn}} + P_{\text{sc}} = a\left(\frac{1}{2}C_{\text{load}} + C_{\text{sc}}\right)V^2f \)

- \( P_{\text{tot}} \approx a\left(\frac{1}{2}C_{\text{load}} + C_{\text{sc}}\right)V^2f + VI_s(W/L)e^{-V_t/(nkT/q)} \)

\[
I_{DS} = I_s'\left(\frac{W}{L}\right)e^{\left(\frac{V_{GS} - V_T}{nkT/q}\right)}\left(1 - e^{-\left(\frac{V_{DS}}{kT/q}\right)}\right)\left(1 + \lambda V_{DS}\right)
\]
Switching Power

- \( P_{\text{dyn}} = \left( \frac{\text{#trans/clock}}{2} \right)^{1/2} CV^2 f \)
- Let \( a = \) activity factor
  \[ a = \text{average } \frac{\text{#tran/clock}}{} \]

- \( P_{\text{dyn}} = a^{1/2} CV^2 f \)
- \( P_{\text{sc}} = aC_{\text{sc}} V^2 f \)
Activity Factor

- Let $a = \text{activity factor}$
  - $a = \text{average } \#\text{tran}_{0\rightarrow1}/\text{clock}$

\[
a = p(out_i = 0)p(out_{i+1} = 1) = \frac{N_0}{2^N} \cdot \frac{N_1}{2^N} = \frac{N_0(2^N - N_0)}{2^{2N}}
\]
Reduce Dynamic Power?

- $P_{\text{dyn}} = a^{1/2}CV^2 f$

- How do we reduce dynamic power?
Reduce Activity Factor

Tree

\[ a = p(out_i = 0)p(out_{i+1} = 1) \]

\[ a = \frac{N_0}{2^N} \frac{N_1}{2^N} = \frac{N_0 (2^N - N_0)}{2^{2N}} \]

Chain
Reduce Activity Factor

Tree

Chain

\[ a = p(out_i = 0)p(out_{i+1} = 1) \]

\[ a = \frac{N_0}{2^N} \frac{N_1}{2^N} = \frac{N_0(2^N - N_0)}{2^{2N}} \]
Reduce Activity Factor

\[ a = p(out_i = 0) p(out_{i+1} = 1) \]

\[ a = \frac{N_0}{2^N} \frac{N_1}{2^N} = \frac{N_0(2^N - N_0)}{2^{2N}} \]
Slow Down

- $P_{\text{dyn}} = a^{1/2}CV^2f$

- What happens to power contributions as we reduce clock frequency?
Slow Down

- \( P_{\text{dyn}} = \alpha \frac{1}{2} CV^2 f \)

- What happens to power contributions as we reduce clock frequency?
  - Example: CMOS circuit consumes equal dynamic and leakage power, \( P \). No short circuit power. The energy consumed in \( T \) seconds is \( 2PT \).

- What suggest about \( V_{\text{th}} \)?
Slow Down

- $P_{\text{dyn}} = \alpha^{1/2} CV^2 f$

- What happens to power contributions as we reduce clock frequency?
  - Example: CMOS circuit consumes equal dynamic and leakage power, $P$. No short circuit power. The energy consumed in $T$ seconds is $2PT$.
  - New power: $(P/2 + P)2T = 3PT$
    - Increased power!
Slow Down

- \( P_{\text{dyn}} = a^{1/2}CV^2 f \)

- What happens to power contributions as we reduce clock frequency?
  - Example: CMOS circuit consumes equal dynamic and leakage power, \( P \). No short circuit power. The energy consumed in \( T \) seconds is \( 2PT \).
  - New power: \( (P/2 + P)2T = 3PT \)
    - Increased power!

- What does this suggest about \( V_{\text{th}} \)?
Reduce $V_{dd}$

- What happens as reduce $V$?
  - Energy?
    - Static
    - Switching
  - Delay?
Reminder:

- $V_{dd}=1V$, $V_{thn}=|V_{thp}|=300mV$

<table>
<thead>
<tr>
<th>$V_{in}$</th>
<th>$I_{static}$</th>
<th>$I_{dynamic}$</th>
<th>$I_{sc}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0V</td>
<td>180pA</td>
<td>126uA</td>
<td></td>
</tr>
<tr>
<td>140mV</td>
<td>6nA</td>
<td>100uA</td>
<td></td>
</tr>
<tr>
<td>400mV</td>
<td></td>
<td>36uA</td>
<td>18uA</td>
</tr>
<tr>
<td>500mV</td>
<td></td>
<td>36uA</td>
<td></td>
</tr>
<tr>
<td>600mV</td>
<td></td>
<td>36uA</td>
<td>18uA</td>
</tr>
<tr>
<td>860mV</td>
<td>6nA</td>
<td>100uA</td>
<td></td>
</tr>
<tr>
<td>1V</td>
<td>180pA</td>
<td>126uA</td>
<td></td>
</tr>
</tbody>
</table>
Reduce $V_{dd}$:

- $V_{dd} = 520\text{mV}$, $V_{thn} = |V_{thp}| = 300\text{mV}$

<table>
<thead>
<tr>
<th>$V_{in}$</th>
<th>$I_{static}$</th>
<th>$I_{dynamic}$</th>
<th>$I_{sc}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0V</td>
<td>180pA</td>
<td>39.6uA</td>
<td></td>
</tr>
<tr>
<td>140mV</td>
<td>6nA</td>
<td>14.4uA</td>
<td></td>
</tr>
<tr>
<td>260mV</td>
<td>111nA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>360mV</td>
<td>6nA</td>
<td>10.8uA</td>
<td></td>
</tr>
<tr>
<td>500mV</td>
<td>180pA</td>
<td>36uA</td>
<td></td>
</tr>
</tbody>
</table>
Reduce V (no velocity saturation)

- $\tau_{gd} = \frac{Q}{I} = \frac{(CV)}{I}$
- $I_d = (\mu C_{OX}/2)(W/L)(V_{gs} - V_{TH})^2$
- $\tau_{gd}$ impact?

$\tau_{gd} \propto \frac{1}{V}$
Reduce V (no velocity saturation)

- \( \tau_{gd} = Q/I = (CV)/I \)
- \( I_d = (\mu C_{OX}/2)(W/L)(V_{gs} - V_{TH})^2 \)
- \( \tau_{gd} \) impact?

- \( \tau_{gd} \propto \frac{1}{V} \)

- Ignoring leakage:
  \[
  E \propto V^2
  \]
  \[
  E\tau^2 \approx \text{Const}
  \]
Reduce V (velocity saturation)

- $\tau_{gd} = \frac{Q}{I} = \frac{(CV)}{I}$
- $I_d = \left(\nu_{sat} C_{OX}\right) \left(W\right) \left(V_{gs} - V_{TH} - V_{DSAT}/2\right)$

![Graph showing the relationship between metric and V]

Penn ESE 370 Fall 2015 - Khanna
Reduce $V_{dd}$ (velocity saturation):

- $V_{thn} = |V_{thp}| = 300\text{mV}$, $V_{in} = V_{dd}$, estimate $E\tau$

<table>
<thead>
<tr>
<th>$V_{dd}$</th>
<th>$I_{ds,vsat}$</th>
<th>$\tau/(\tau @ V_{dd}=1)$</th>
<th>$E_{\text{switch}}/(E_{\text{switch}} @ V_{dd}=1)$</th>
<th>$E\tau$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1V</td>
<td>126uA</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>700mV</td>
<td>72uA</td>
<td>1.225</td>
<td>0.49</td>
<td>.6</td>
</tr>
<tr>
<td>500mV</td>
<td>36uA</td>
<td>1.75</td>
<td>0.25</td>
<td>.437</td>
</tr>
<tr>
<td>350mV</td>
<td>9uA</td>
<td>4.9</td>
<td>0.12</td>
<td>.588</td>
</tr>
<tr>
<td>260mV</td>
<td>111nA</td>
<td>295</td>
<td>0.07</td>
<td>20.6</td>
</tr>
</tbody>
</table>
Reduce Short-Circuit Power?

- \( P_{sc} = aC_{sc}V^2f \)

\[
E = V_{dd} \times \left( I_{peak} \times t_{sc} \times \left( \frac{1}{2} \right) \right)
\]
Increase $V_{th}$?

- What is impact of increasing threshold on
  - Delay?
  - Leakage?
Increase $V_{th}$

- $\tau_{gd} = Q/I = (CV)/I$
- $I_{ds} = (\nu_{sat} C_{OX})(W)(V_{gs} - V_{TH} - V_{DSAT}/2)$
Increase $V_{th}$?

- What is impact of increasing threshold on
  - Delay?
  - Leakage?

- $V_{dd}=1\,V$, $V_{in}=V_{dd}$

<table>
<thead>
<tr>
<th>$V_{thn}=-V_{thp}$</th>
<th>$I_{ds,\text{vsat}}$</th>
<th>$\tau/(\tau@V_{th}=300,\text{mV})$</th>
<th>$I_{\text{static}}$</th>
<th>$I_{\text{static}}/(I_{\text{static}@V_{th}=300,\text{mV}})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>300mV</td>
<td>126uA</td>
<td>1</td>
<td>180pA</td>
<td>1</td>
</tr>
<tr>
<td>460mV</td>
<td>97uA</td>
<td>1.3</td>
<td>3pA</td>
<td>0.016</td>
</tr>
<tr>
<td>600mV</td>
<td>72uA</td>
<td>1.75</td>
<td>90fA</td>
<td>0.0005</td>
</tr>
</tbody>
</table>
Idea

- Tradeoff
  - Speed
  - Switching energy
  - Leakage energy

- Energy-Delay tradeoff: $E\tau^2$, $E\tau$
Ratioed Logic
Previously

- Restoration and Noise Margins
  - Allows for gate abstraction

- CMOS Gates
  - Drive outputs rail-to-rail
  - Only one transistor turned on in steady state
    - Only subthreshold leakage current in steady state
Today

- **Ratioed Gates**
  - Break all the rules… (nice properties)
    - No rail-to-rail outputs, steady-state-current is not subthreshold…
  - Logic correctness
  - Performance
  - Power
  - Implications
Idea

- Building both pull-up and pull-down can be expensive – many gates
- Seems wasteful to build logic function twice
  - Once in pullup, once in pulldown
  - Large gate capacitance
Idea

- Maybe only need to build one
- Build NFET pulldown
  - Exploit high N mobility
    - traditional
Ratioed Inverter

- Does this work?
  - What is $V_{out}$ for $V_{in} = \text{Gnd}$?
  - What is $V_{out}$ for $V_{in} = V_{dd}$?

\[ W_p = 1 \]
\[ W_N = 1 \]
Ratioed Inverter in 22nm

DC Transfer Curve for Ratioed Inverter

Vout

Vin

Wn=1
Ratioed Inverter

- How do we need to size N to make it work?

\[ W_P = 1 \]

\[ W_N = ? \]
Ratioed Inverter in 22nm

DC Transfer Curve for Ratioed Inverter

- \( W_n=1 \)
- \( W_n=2 \)
- \( W_n=4 \)
- \( W_n=8 \)

\[ \text{Vout} \]
\[ \text{Vin} \]
DC Transfer Function

![DC Transfer Curve for Ratioed Inverter](image.png)

- $W_n = 8$
Ratioed Inverter

How do we need to size P to make it work?

$W_N = 1$
Ratioed Inverter in 22nm

DC Transfer Curve for Ratioed Inverter

Vin

Vout

Wp=1
Wp=4
Wp=2
Wp=8
Wp=16
Wp=32
P vs. N

- **Conclude:** still prefer $N$ to $P$ for ratioed logic
Noise Margin Tradeoff

- What is impact of increasing noise margin?
  - On size
  - On input capacitance

![DC Transfer Curve for Ratioed Inverter](image)
Worst-Case Output Drive Strength?

- $R_{\text{drive}}$?

- $R_{\text{drive}}$ when input low?

- $R_{\text{drive}}$ when input high?

$W_p = 1$
Ratioed Inverter Sizing

- What causes knee in curve at high end?

![DC Transfer Curve for Ratioed Inverter](image-url)
Non-VSAT Equations

**Saturation**

\[ I_{DS} = \frac{\mu_n C_{OX}}{2} \left( \frac{W}{L} \right) \left[ (V_{GS} - V_T)^2 \right] \]

**Linear**

\[ I_{DS} = \mu_n C_{OX} \left( \frac{W}{L} \right) \left[ (V_{GS} - V_T)V_{DS} - \frac{V_{DS}^2}{2} \right] \]
Non-VSAT Equations

- P looks roughly like current source
- Must Match current with N
  - Determine $V_{ds}$
  - How does $W_n$ effect $V_{ds}$?

\[
I_{DS} = \frac{\mu_n C_{OX}}{2} \left( \frac{W}{L} \right) \left( V_{GS} - V_T \right)^2
\]

\[
I_{DS} = \mu_n C_{OX} \left( \frac{W}{L} \right) \left[ (V_{GS} - V_T)V_{DS} - \frac{V_{DS}^2}{2} \right]
\]
Size for $R_0/2$ drive?

- How do we size for $R_0/2$ drive?
Compare Static CMOS

For $R_{\text{drive}} = \frac{R_0}{2}$ inverter

- Total Transistor Width?
- Input capacitance load?
Power?

- $I_{\text{static}}$ ?

- Input low-Output high?
  - $I_{\text{leak}}$

- Input high-Output low?
  - $I_{\text{pmos\_on}}$
  - $V_{dd}/(R_0/2)$ -- for our sample case
Power

- $P_{tot} \approx a\left(\frac{1}{2}C_{load} + C_{sc}\right)V^2f$
  
  $\quad + p(V_{out}=\text{low})V^2/R_{pon}$

  $\quad + (1-p(V_{out}=\text{low}))Vt_s(W/L)e^{-Vt/(nkT/q)}$

$p(V_{out}=\text{low})$ – probability the output is low
How size for $R_0/2$ drive?
How size for $R_0/2$ drive?
How size for $R_0/2$ drive?

- How size K-input nor?
When better than CMOS nor-k?

- Better = smaller, lower input capacitance
Ideas

- There are other logic disciplines
- We have the tools to analyze
- Ratioed Logic
  - Tradeoff noise margin for
    - Reduced area? Capacitive load?
  - Dissipates static power in one mode
Admin

- **Project – Milestone due Wednesday**
  - Should have read it already
  - Build and simulate baseline design over weekend
    - Ch 11 in textbook
  - Start list of optimizations to try
    - If don’t have list talk come to office hours ready to talk about it
      (Khanna: Monday, Giesen: Tuesday)

- **Extra office hours next week on Thursday and Friday**
  - Giesen: Thursday 6-8pm
  - Khanna: Friday 4-6pm