Switching Activity Driven Gate Sizing and Vth Assignment for Low Power Design

Yu-Hui Huang
Po-Yuan Chen
TingTing Hwang

National Tsing Hua University
Taiwan
Outline

• Introduction and Motivation
• Related Work
• Algorithm
• Experimental Result
• Conclusion
Introduction

• Power = $\alpha \cdot$ Active Power + (1 - $\alpha$) $\cdot$ Idle Power

• Active Power
  • Dynamic power
    • Gate sizing
  • Leakage power
    • Vth Re-assignment

• Idle Power
  • Leakage power

• Minimize total power
Motivation

• To enhance the performance of a circuit, we can **size-up gates** or **replace the Vth of gates** from high to low.
  - *Size-up*:
    - Increase dynamic power and small leakage power
  - *Replace* the Vth of cells from high to low:
    - Increase leakage power

• Which one is better?
  - Depends on the **switching activity** of a gate.
Motivation

<table>
<thead>
<tr>
<th></th>
<th>Inverter A</th>
<th>Inverter B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vth</td>
<td>_</td>
<td>lower</td>
</tr>
<tr>
<td>Size</td>
<td>larger</td>
<td>_</td>
</tr>
</tbody>
</table>

- Inverter A and B have same delay and output loading

- Comparison function

\[
\text{dyn}(A) - \frac{\text{dyn}(B)}{\text{lea}(B) - \text{lea}(A)}
\]
Motivation

- **Switching Activity**
  - Gate Sizing
  - Vth re-assignment

<table>
<thead>
<tr>
<th>Switching activity (α)</th>
<th>Ratio (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>TOP</td>
</tr>
<tr>
<td>%&lt;(\alpha) %&lt;(2%)</td>
<td></td>
</tr>
<tr>
<td>%&lt;(\alpha) %&lt;(2%)</td>
<td></td>
</tr>
</tbody>
</table>
Outline

• Introduction and Motivation
• Related Work
• Algorithm
• Experimental Results
• Conclusions
Related Work

• Previous work focused on minimizing power on **non-critical path**.

• **We can** minimize **power both on critical path and non-critical path.**
  - **On critical path:**
    We can **re-assign Vth to high and up-size gates** which has small switching activity.
  - **On non-critical path:**
    Slack can be used to **down-size gate or assign Vth to high**.
Outline

- Introduction and Motivation
- Related Work
- Algorithm
- Experimental Result
- Conclusion
Algorithm – Design Flow

Library of cells with low Vth

Library of cells with high Vth

Circuit & Timing Constraint

Step1. Circuit Synthesis

Step2. Swap cells from low Vth to high Vth

Step3. Gate sizing and Vth re-assignment

Output Circuit
Algorithm – Step 3

Step 3.1 For Critical Path

Timing constraint satisfied?

No

Size up or Vth re-assignment tech. to solve the timing violation problem

Yes

Minimize the power on non-critical paths by slack

Step 3.2 For Non Critical Path
Algorithm for Critical Path (Step 3.1)

Timing constraint satisfied?

Yes

No

Construct **path balanced graph**

Find minimal cost **separator sets** of the **path balanced graph**

Replace the nodes in the **separator sets**

**Step 3.2**
Step 3.1: Constructing Path Balanced Graph

- **Path-Balanced Graph**

\[
ds(e) = \text{slack(head_node(e))} - \text{slack(tail_node(e))}
\]
Step 3.1: Computing Cost

- Set **cost** of each node

\[
\text{cost}(g) = \gamma \cdot \text{penalty}(g) + \\
\delta \cdot \text{delay_reDUCTION}(g)
\]

\[
\text{penalty}(g) = \alpha \cdot \text{p_penalty}(g) + \\
\beta \cdot \text{a_penalty}(g)
\]

\[
\text{p_penalty}(g) = \text{per} \cdot ( \\
\sum_{j \in \text{fanin}(g)} \text{E}(j) \cdot C_{\text{inc}}(g) \cdot V^2 \\
+ \text{leak}_{\text{inc}}(g)) \\
+ (1 - \text{per}) \cdot \text{leak}_{\text{inc}}(g)
\]

**E(j)** is the transition density of node j
Step 3.1: Finding Separator Set

- Find separator set of minimal cost in the graph

Delay improvement $\min\{0.7, 0.5\} = 0.5$

$(x, y, z)$ means (slack, delay-reduction, cost)
Step 3.1: Finding Separator Set

- Find separator set of minimal cost in the graph

\[
\text{Delay improvement } \min \{0.5, 0.25\} = 0.25
\]
Algorithm for Non Critical Path

Step 3.1 For Critical Path
- Construct path balanced graph
- Compute minimal cost separator sets of the path balanced graph
- Size Gate or re-assign Vth to solve the timing violation problem
- Replace the nodes in the separator sets
- Timing constraint satisfied?
  - Yes
  - No

Step 3.2 For Non Critical Path
- Minimize the power on non-critical paths by slack
Algorithm for Non Critical Path

Compute available slack

Compute delay penalty caused by down-sizing or re-assigning Vth to high

Compute power saving caused by down-sizing or re-assigning Vth to high

Down-sizing or re-assigning high Vth

Avail_slack (A) = Minimal delay reduction of \( \{P_1, P_2\} = \min\{0.5, 0.4\} = 0.4 \)

(0.6/0.5/0) P1

(0.6/0.4/0) P2

A

B

C

D

E

F

G

H

Avail_slack (E) = 0

(x,y,z) means (slack ,delay-reduction ,cost)
## Step 3.2: Computing Penalty

<table>
<thead>
<tr>
<th>Compute available slack</th>
<th>Compute delay penalty caused by down-sizing or re-assigning Vth to high</th>
<th>Compute power saving caused by down-sizing or re-assigning Vth to high</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>delay-penalty($g$) = $\text{Delay(new}_g) - \text{Delay}(g)$</td>
<td>p_saving($g$) = p_penalty($g$) − p_penalty(new.g)</td>
</tr>
</tbody>
</table>

**down-sizing or re-assigning Vth to high**
Step 3.2: Replacing Cell

- Compute available slack

- Compute delay penalty caused by down-sizing or re-assigning Vth to high

- Compute power saving caused by down-sizing or re-assigning Vth to high

- down-sizing or re-assigning Vth to high

- If only one delay penalty of two options is less than available slack?
  - choose the available one

- If delay penalties of both options are less than available slack?
  - depends on the larger power saving
Outline

• Introduction and Motivation
• Related Work
• Algorithm
• Experimental Result
• Conclusion
Experimental Result

- **Benchmarks**

<table>
<thead>
<tr>
<th>Cir.</th>
<th>Cell Count</th>
<th>Characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>TOP</td>
<td>463</td>
<td>An Alarm Clock</td>
</tr>
<tr>
<td>MAC</td>
<td>2425</td>
<td>Multiplier and Accumulator</td>
</tr>
<tr>
<td>AVG</td>
<td>6361</td>
<td>Average Number Calculator</td>
</tr>
<tr>
<td>GCC</td>
<td>8204</td>
<td>Gravity Center Calculator</td>
</tr>
<tr>
<td>RSA</td>
<td>14815</td>
<td>Asymmetric Crypto-Processor</td>
</tr>
<tr>
<td>AES</td>
<td>16824</td>
<td>Advanced Encryption Core</td>
</tr>
</tbody>
</table>
Experimental Result

• TOOLS

  • DesignCompiler
  • TSMC 0.13um library
  • PrimeTime
  • PrimePower
Experimental Result

- **Power saving**

<table>
<thead>
<tr>
<th>Circuit</th>
<th>$\alpha = 100%$</th>
<th>$\alpha = 50%$</th>
<th>$\alpha = 10%$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$P$ (mW)</td>
<td>$Red$</td>
<td>$P$ (mW)</td>
</tr>
<tr>
<td>$\alpha_1$</td>
<td>0.363</td>
<td>11.95%</td>
<td>0.179</td>
</tr>
<tr>
<td>$\alpha_2$</td>
<td>0.790</td>
<td>18.56%</td>
<td>0.397</td>
</tr>
<tr>
<td>$\alpha_3$</td>
<td>1.65</td>
<td>5.75%</td>
<td>0.835</td>
</tr>
<tr>
<td>$\alpha_4$</td>
<td>0.753</td>
<td>6.48%</td>
<td>0.412</td>
</tr>
<tr>
<td>$\alpha_5$</td>
<td>2.12</td>
<td>39.20%</td>
<td>1.08</td>
</tr>
<tr>
<td>$\alpha_6$</td>
<td>13.4</td>
<td>15.60%</td>
<td>6.70</td>
</tr>
<tr>
<td>$\alpha_7$</td>
<td>16.26%</td>
<td>18.53%</td>
<td></td>
</tr>
</tbody>
</table>
## Experimental Result

- **Time Penalty**

  \( \alpha \) is the fraction of active time

<table>
<thead>
<tr>
<th>Circuit</th>
<th>Original T</th>
<th>( \alpha = 100% )</th>
<th>( \alpha = 50% )</th>
<th>( \alpha = 10% )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>T'</td>
<td>T penalty</td>
<td>T'</td>
</tr>
<tr>
<td>ꔿ ꔿ ꔿ</td>
<td>1.43</td>
<td>1.37</td>
<td>-4.2%</td>
<td>1.39</td>
</tr>
<tr>
<td>ꔿ ꔿ ꔿ</td>
<td>3.30</td>
<td>3.33</td>
<td>0.8%</td>
<td>3.33</td>
</tr>
<tr>
<td>ꔿ ꔿ ꔿ</td>
<td>23.78</td>
<td>23.13</td>
<td>-2.7%</td>
<td>23.46</td>
</tr>
<tr>
<td>ꔿ ꔿ ꔿ</td>
<td>26.30</td>
<td>26.65</td>
<td>1.3%</td>
<td>26.73</td>
</tr>
<tr>
<td>ꔿ ꔿ ꔿ</td>
<td>10.00</td>
<td>10.08</td>
<td>0.8%</td>
<td>10.03</td>
</tr>
<tr>
<td>ꔿ ꔿ ꔿ</td>
<td>2.29</td>
<td>2.21</td>
<td>-3.5%</td>
<td>2.27</td>
</tr>
</tbody>
</table>
Outline

• Introduction and Motivation
• Related Work
• Algorithm
• Experimental Result
• Conclusion
Conclusion

• **Switching activity** of a gate plays an important role in making decision to choose gate sizing or Vth assignment.

• Under the **timing constraint**, our circuit have 16% and 18% improvement as compared to the original circuits where the fraction of active time are 100% and 50%, respectively.
Thank you