Lecture 7: Transient Power Reduction

CSCE 6730
Advanced VLSI Systems

Instructor: Saraju P. Mohanty, Ph. D.

NOTE: The figures, text etc included in slides are borrowed from various books, websites, authors pages, and other sources for academic purpose only. The instructor does not claim any originality.
Outline of the Talk

• Different power parameters
• Cycle power profile function
• Heuristic to minimize CPF
• Experimental results
• Related research
• Conclusions

Different Power and Energy Parameters

- Peak power
- Cycle difference power
- Peak power differential
- Average Power
- Total Energy
Peak Power

The peak power is the maximum power consumption of the IC at any instance during its execution.

For a DFG, let $P_c$ denote power consumption in any control step $c$, then we define peak (cycle) power as:

$$P_{\text{peak}} = \text{maximum}(P_c), \text{ over all control steps}$$
Average Power and Total Energy

Average power (P) = Average of (cycle power consumption i.e. \( P_c \)) over all control steps

Total energy = Energy consumption for the DFG for all operations and control steps
Cycle Difference Power and Peak Power Differential

Let, $DP_c = \text{absolute } (P - P_c)$ denote the cycle difference power. This characterizes the power fluctuation for each cycle of DFG.

Peak power differential is defined as:

$$DP_{peak} = \text{maximum } (DP_c)$$
Transient Power?

Both the peak power and peak power differential drive the transient characteristic of a CMOS circuit.
We Aim At:

Simultaneous reduction of:

• Peak power
• Cycle difference power
• Peak power differential
• Average power
• Total energy
Our Approach

Define a new parameter (CPF) that captures all power parameters

Minimize the new parameter in using multiple supply voltage and dynamic frequency
Normalized Average Power ($P_{\text{norm}}$)

Normalized average power ($P_{\text{norm}}$)

= Average of cycle power consumption over all control steps / maximum power consumption in any control step

= $\text{Average} \left( P_c \right) / \text{maximum} ( P_c )$

= $P / P_{\text{peak}}$
Normalized Average Cycle Difference Power (DP_{norm})

Normalized average cycle difference power (DP_{norm})

= average cycle difference power over all control steps / maximum cycle difference power for any control step

= Average (DP_c) / Maximum (DP_c)

= DP / DP_{peak}
Normalized cycle power function (CPF)

Normalized cycle power profile function is defined as:

\[
CPF_{\text{norm}} = PF \times P_{\text{norm}} + (1-PF) \times DP_{\text{norm}}
\]

Where, \( PF \) = power profile factor used to make \( CPF_{\text{norm}} \)

either cycle power dominating (average and peak) or
difference power dominating (cycle difference and peak
differential)

\( P_{\text{norm}} \) = normalized average power

\( DP_{\text{norm}} \) = normalized average cycle difference power
Normalized CPF ............

Is a function of five different parameters:

• Average power power (P)
• Peak power (P\text{peak})
• Average cycle difference power (DP)
• Peak differential power (DP\text{peak})
• Power profile factor (PF)
Each Power is Determined by:

- $\alpha_{i,c} = \text{switching activity of resource } i \text{ active in control step } c$
- $C_{i,c} = \text{load capacitance of resource } i \text{ active in control step } c$
- $V_{i,c} = \text{operating voltage of resource } i \text{ active in control step } c$
- $f_c = \text{frequency of control step } c$
CPF Minimization

Minimization of the normalized cycle power profile function using multiple supply voltages and dynamic clocking frequency can minimize all the powers and energy parameters.
CPF-Scheduler

**Input:** Unscheduled data flow graph, resource constraint, number of allowable voltage levels, number of allowable frequencies, load capacitance of each resource, delay of each functional unit at different voltage levels, operating frequencies and voltages

**Output:** Scheduled data flow graph, base frequency, cycle frequency index, operating voltage for each operation
CPF-Scheduling Algorithm Flow

Step 1: Get the ASAP and ALAP schedule.

Step 2: Modify the ASAP and ALAP schedules using the number of resources without operating voltage constraint.

Step 3: Total No. of control steps = Maximum (ASAP steps, ALAP steps).

Step 4: Find the vertices having zero and non-zero mobility.

Step 5: Use the CPF-Scheduler-Heuristic to assign time stamp, voltage level and cycle frequency such that $CPF_{norm}$ is minimum.

Step 6: Find cycle frequency index for each cycle.
CPF-Scheduler Heuristic

(01) initialize CurrentSchedule as ASAPSchedule;

(02) while (all mobile vertices are not time stamped) do

(03) for the CurrentSchedule

(04) if (vᵢ is a multiplication) then find the lowest available voltage for multipliers;

(05) if (vᵢ is add/sub) then find the highest available operating voltage for ALUs;

(06) find CurrentCPF<sub>norm</sub> for CurrentSchedule; Maximum = -∞;

(07) for each mobile vertex vᵢ

(08) c₁ = CurrentSchedule[vᵢ]; c₂ = ALAPSchedule[vᵢ];

(09) for c = c₁ to c₂ in steps of 1

(10) find a TempSchedule by adjusting CurrentSchedule in which vᵢ is scheduled in c;

(11) find next higher operating voltage for multiplication vertex (next lower for ALU operation) for the TempSchedule;

(12) find TempCPF<sub>norm</sub> for TempSchedule; DiffCPF = CurrentCPF<sub>norm</sub> - TempCPF<sub>norm</sub>

(13) if (DiffCPF > Maximum) then Maximum = DiffCPF; CurrentVertex = vᵢ;

CurrentCycle = c; CurrentVoltage = Operating voltage of vᵢ;

(14) adjust CurrentSchedule to accommodate vᵢ in c operating at voltage assigned above;
CPF-Scheduler Heuristic: Explanations

- The heuristic is used to find proper time stamp, operating voltage for mobile vertices such that the CPF\(_{\text{norm}}\) is minimum for whole DFG.
- Initially assumes the modified ASAP schedule (with relaxed voltage resource constrained) as the current schedule.
- The CurrentCPF\(_{\text{norm}}\) value for the current schedule is calculated.
- The heuristic finds CPF\(_{\text{norm}}\) values (TempCPF\(_{\text{norm}}\)) for each allowable control step of each mobile vertices and for each available operating voltages.
- The heuristic fixes the time step, operating voltage and hence cycle frequency for which CPF\(_{\text{norm}}\) is minimum.
Experimental Results : Resource Constraints Used

<table>
<thead>
<tr>
<th>Multipliers</th>
<th>ALUs</th>
<th>Serial</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.3V</td>
<td>5.0V</td>
<td>3.3V</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>
Notations Used to Describe the Results

- $\Delta P_p = \frac{(P_{pS} - P_{pD})}{P_{pS}} = \text{peak power reduction}$
- $\Delta DP = \frac{((P_{pS} - P_{mS}) - (P_{pD} - P_{pD}))}{(P_{pS} - P_{mS})} = \text{peak differential reduction}$
- $\Delta P = \frac{(P_S - P_D)}{P_S} = \text{average power reduction}$
- $\Delta E = \frac{(E_S - E_D)}{E_S} = \text{reduction in total energy}$

Where,

subscript S : single voltage and single freq operation
subscript D : multiple voltage and dynamic freq
Subscript m : minimum power power
% Reductions for Different Benchmarks

<table>
<thead>
<tr>
<th></th>
<th>RCs</th>
<th>( \Delta P_p )</th>
<th>( \Delta DP )</th>
<th>( \Delta P )</th>
<th>( \Delta E )</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ARF</strong></td>
<td>1</td>
<td>63</td>
<td>68</td>
<td>71</td>
<td>47</td>
</tr>
<tr>
<td>(1)</td>
<td>3</td>
<td>70</td>
<td>72</td>
<td>69</td>
<td>47</td>
</tr>
<tr>
<td><strong>BPF</strong></td>
<td>1</td>
<td>73</td>
<td>79</td>
<td>66</td>
<td>46</td>
</tr>
<tr>
<td>(2)</td>
<td>3</td>
<td>73</td>
<td>87</td>
<td>71</td>
<td>46</td>
</tr>
<tr>
<td><strong>DCT</strong></td>
<td>1</td>
<td>63</td>
<td>68</td>
<td>50</td>
<td>41</td>
</tr>
<tr>
<td>(3)</td>
<td>3</td>
<td>61</td>
<td>72</td>
<td>67</td>
<td>41</td>
</tr>
<tr>
<td><strong>EWF</strong></td>
<td>1</td>
<td>73</td>
<td>79</td>
<td>41</td>
<td>44</td>
</tr>
<tr>
<td>(4)</td>
<td>3</td>
<td>69</td>
<td>72</td>
<td>55</td>
<td>44</td>
</tr>
<tr>
<td><strong>FIR</strong></td>
<td>1</td>
<td>70</td>
<td>75</td>
<td>58</td>
<td>46</td>
</tr>
<tr>
<td>(5)</td>
<td>3</td>
<td>77</td>
<td>84</td>
<td>54</td>
<td>46</td>
</tr>
<tr>
<td><strong>HAL</strong></td>
<td>1</td>
<td>73</td>
<td>94</td>
<td>73</td>
<td>51</td>
</tr>
<tr>
<td>(6)</td>
<td>3</td>
<td>76</td>
<td>97</td>
<td>70</td>
<td>51</td>
</tr>
</tbody>
</table>
Average Reductions for Benchmarks

- Peak Pow Red (%)
- Avg Pow Red (%)
- Time Penalty (unit)
- Energy Savings (%)
- Peak Pow Diff Red (%)
- Normalized CPF (unit)

Different Benchmark Circuits →

1 2 3 4 5 6
Power Profiles for Benchmarks
($\alpha = 0.5$, PF = 0.5, RC1)
Power Profiles ……
(α = 0.5, PF = 0.4, RC2)
Power Profiles .......
(\(\alpha = 0.3\), PF = 0.8, RC3)
Power Profiles ……
($\alpha = 0.4$, $PF = 0.2$, RC4)
Power Profiles .......
(α = 0.4, PF = 0.7, RC5)
CPF Vs PF plot
(\(\alpha = 0.5\), RC3 and RC4)
## Reductions Using Different Algorithms
(Only peak power reduction avg data given)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>ARF</td>
<td>68</td>
<td>50</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>BPF</td>
<td>71</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>DCT</td>
<td>64</td>
<td>50</td>
<td>71</td>
<td>28</td>
</tr>
<tr>
<td>EWF</td>
<td>72</td>
<td>0</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>FIR</td>
<td>71</td>
<td>63</td>
<td>45</td>
<td>23</td>
</tr>
<tr>
<td>HAL</td>
<td>73</td>
<td>28</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
Related Research …
(Peak power reduction at behavioral level)

• Martin & Knight [7], 1996 – simultaneous assignment and scheduling

• Raghunathan and et al. [13], 2001 – also address peak power differential

• Shiue [15], 2000 – ILP formulation to reduce peak power under latency constraints

• And many other works
Related Research ... : Martin and Knight [7]

• Peak power reduction is achieved through simultaneous assignment and scheduling

• Use minimization at one level of abstraction to achieve optimization at other level (specifically, simultaneous use of SPICE and behavioral synthesis tool)

• Genetic algorithm has been used for optimization

• Peak power reduction : 40-60%,

• Average power penalty : 0.3-2.7%
Related Research … : Raghunathan [13]

• Simultaneous minimization of peak power and peak power differential

• Use data-monitor operations

• Peak power reduction : 17-32%

• Peak power differential reduction : 25-58%

• Judicious use of transient power metric needed for minimization of area and performance overhead
Related Research ... : Shiue [15]

• ILP based scheduling and modified force-directed scheduling
• Peak power minimization under latency constraints
• Single supply voltage, multicycling and pipelining
• Peak power reduction : 0-75 %
Conclusions

• This work is a unified framework for simultaneous power and energy reduction

• The CPF parameter defined and used in this work facilitates such simultaneous reduction

• CPF-Scheduler algorithm developed that takes resources constraints, minimizes CPF

• The average time penalty is estimated to be 40%

• Future works needs to be done using better optimization technique