Power Fluctuation Minimization During Behavioral Synthesis using ILP-Based Datapath Scheduling

Saraju P. Mohanty, N. Ranganathan and Sunil K. Chappidi
Dept. of CSE, University of South Florida, Tampa, FL 33620
{smohanty,ranganat,chappidi}@csee.usf.edu

***Abstract*—**

We model the power fluctuation as cycle-to-cycle power gradient and minimize the mean of the power gradients using ILP. We propose scheduling schemes for three modes of datapath design : single supply voltage and single frequency (SVSF), multiple supply voltages and dynamic frequency clocking (MVDFC), and multiple supply voltages and multicycling (MVMC). Various experiments are conducted on selected high-level synthesis benchmarks. Experimental results in terms of several parameters, such as mean power gradient, mean cycle power, peak power, and power delay product, are presented.

I. Introduction

The power fluctuation is critical to reduce power supply noise and cross talk, and to increase battery life and reliability. Power fluctuation leads to larger $\frac{di}{dt}$ causing power supply noise, because of self inductance. Large $\frac{di}{dt}$ and $\frac{dv}{dt}$ due to high power fluctuation can introduce significant noise in signal lines due to mutual inductance and capacitance (cross-talk). More the power fluctuation lesser the electrochemical conversion, hence decrease in battery life. High current peaks in short time spans can cause high heat dissipation in a localised area of die which may lead to failures.

There are few research works minimizing peak power at behavioral level. In [4], the peak power reduction is achieved through simultaneous assignment and scheduling. In [11], ILP based scheduling and force directed scheduling have been proposed to minimize peak power. In [10], data monitor operations are used for simultaneous reduction of peak power and peak power differential. In [7], heuristic based scheme is proposed that minimizes peak power, peak power differential, average power, energy altogether.

II. Power Fluctuation Modeling

Let $x_1, x_2, ..., x_n$ be a set of n observations from a given distribution. The sample mean (an unbiased estimate of population mean μ) is $m = \frac{1}{n}\sum_{i=1}^{n} x_i$. The observation-to-observation gradient can be defined as, $|x_i - x_{i-1}|$, where $2 \leq i \leq n$. The mean gradient is given by $\frac{1}{n-1}\sum_{i=2}^{n}|x_i - x_{i-1}|$. This is the basis of our power fluctuation modeling. Let us assume that the datapath is represented in the form of a sequencing DFG. The following notations are used in the description : N : total number of control steps in the DFG, O : total number of operations in the DFG, c : a control step or a clock cycle in DFG $(1 \leq i \leq N)$, o_i : any operation i $(1 \leq i \leq O)$, P_c : the total power consumption in control step c, P_p : peak power consumption for the DFG $(= max\,(P_c)_{\forall c})$, P_a : mean power consumption of the DFG $\left(= \frac{1}{N}\sum_{c=1}^{N} P_c\right)$, PG_c : power gradient for cycle c (where, $c = 2 \rightarrow N$), PG_p : peak power gradient of the DFG, MPG : mean power gradient of the DFG over $c = 2 \rightarrow N$, $FU_{k,v}$: any functional unit of type k operating at voltage v, FU_i : any $FU_{k,v}$ needed by o_i for its execution, $FU_{i,c}$: any functional unit FU_i active in control step c, R_c : total number of functional units active in step c, $\alpha_{i,c}$: switching activity of resource $FU_{i,c}$, $V_{i,c}$: operating voltage of resource $FU_{i,c}$, $C_{i,c}$: load capacitance of resource $FU_{i,c}$, and f_c : frequency of control step c.

The power consumption for any control step c is given by Eqn. 1. The level converters are considered as resources operating in a cycle c, if the current resource is driven by a resource operating at lower voltage.

$$P_c = \sum_{i=1}^{R_c} \alpha_{i,c} C_{i,c} V_{i,c}^2 f_c \tag{1}$$

The power gradient PG_c for any step is defined as the absolute difference of power consumption from its previous step.

$$PG_c = |P_c - P_{c-1}| \quad (\forall c = 2, 3, ...N) \quad = \left|\sum_{i=1}^{R_c} \alpha_{i,c} C_{i,c} V_{i,c}^2 f_c - \sum_{i=1}^{R_{c-1}} \alpha_{i,c-1} C_{i,c-1} V_{i,c-1}^2 f_{c-1}\right| \tag{2}$$

The peak of the power gradients is denoted as (PG_p) :

$$\begin{aligned} PG_p &= max\left(|P_c - P_{c-1}|\right)_{\forall c=2,3,....N} \\ &= max\Big(\Big|\sum_{i=1}^{R_c} \alpha_{i,c} C_{i,c} V_{i,c}^2 f_c \\ &- \sum_{i=1}^{R_{c-1}} \alpha_{i,c-1} C_{i,c-1} V_{i,c-1}^2 f_{c-1}\Big|\Big)_{\forall c} \end{aligned} \tag{3}$$

The mean power gradient MPG is calculated as,

$$\begin{aligned} MPG &= \frac{1}{N-1}\sum_{c=2}^{N} PG_c = \frac{1}{N-1}\sum_{c=2}^{N}|P_c - P_{c-1}| \\ &= \frac{1}{N-1}\sum_{c=2}^{N}\Big(\Big|\sum_{i=1}^{R_c} \alpha_{i,c} C_{i,c} V_{i,c}^2 f_c \\ &- \sum_{i=1}^{R_{c-1}} \alpha_{i,c-1} C_{i,c-1} V_{i,c-1}^2 f_{c-1}\Big|\Big) \end{aligned} \tag{4}$$

Using the dynamic energy model proposed in [1] we express the effective switching capacitance as, $\alpha_i C_i = C_{swi}(\alpha_i{}^1, \alpha_i{}^2)$. The α_i and C_i are the parameters corresponding to the functional unit FU_i. The C_{swi} is a measure of the effective switching capacitance of FU_i, which is a function of the average switching activities on the first and second input operands. Using the above, we rewrite Eqn. 4 and get the

following objective function.

$$\begin{aligned} MPG &= \frac{1}{N-1}\sum_{c=2}^{N}\Big(\Big|\sum_{i=1}^{R_c} C_{swi,c}V_{i,c}^2 f_c \\ &- \sum_{i=1}^{R_{c-1}} C_{swi,c-1}V_{i,c-1}^2 f_{c-1}\Big|\Big) \end{aligned} \quad (5)$$

III. Scheduling Algorithm

Due to lack of space, we provide ILP formulations for MVDFC mode only; similar formulations are done for SVSF and MVMC schemes. Different decision variables are used for other two modes of formulations. We use the following notations for ILP formulation : $M_{k,v}$ = maximum number of functional units $FU_{k,v}$, S_i = ASAP time stamp for the operation o_i, E_i = ALAP time stamp for the operation o_i, $P(C_{swi}, v, f)$ = power consumption of functional unit FU_i at voltage v and frequency f used by o_i for its execution, and $x_{i,c,v,f}$ = decision variable which takes the value of 1 if operation o_i is scheduled in control step c using the functional unit $F_{k,v}$ and c has frequency f_c.

(a) *Objective Function* : The objective is to minimize the mean power gradient MPG described in Eqn. 5 of the whole DFG over all control steps.

$$\text{Minimize :} \quad MPG = \frac{1}{N-1}\sum_{c=2}^{N}|P_c - P_{c-1}| \quad (6)$$

This problem has a non-linearity in it because of the absolute function. This can be converted to an equivalent problem using the transformation suggested in [5], [9].

$$\begin{aligned} &\text{Minimize :} \quad \frac{1}{N-1}\sum_{c=2}^{N}(P_c + P_{c-1}) \\ &\text{Subject to :} \quad \text{Power gradient constraints} \end{aligned} \quad (7)$$

The above problem in Eqn. 7 is simplified to :

$$\begin{aligned} &\text{Minimize :} \quad \frac{2}{N-1}\sum_{c=2}^{N-1} P_c + P_1 + P_N \\ &\text{Subject to :} \quad \text{Power gradient constraints} \end{aligned} \quad (8)$$

Using the decision variables, we have,

$$\begin{aligned} \text{Min: } &\Big(\frac{2}{N-1}\Big)\sum_{c=2}^{N-1}\sum_{i\in F_{k,v}}\sum_v\sum_f x_{i,c,v,f}P(C_{swi},v,f) \\ &+\sum_{i\in F_{k,v}}\sum_v\sum_f x_{i,1,v,f}P(C_{swi},v,f) \\ &+\sum_{i\in F_{k,v}}\sum_v\sum_f x_{i,N,v,f}P(C_{swi},v,f) \\ &\text{Subject to :Power gradient constraints} \end{aligned} \quad (9)$$

(b) *Uniqueness Constraints* : These constraints ensure that every operation o_i is scheduled to one unique control step within the mobility range (S_i, E_i) with a particular voltage and frequency. and are represented as, $\forall i,\ 1 \le i \le O$,

$$\sum_c\sum_v\sum_f x_{i,c,v,f} = 1 \quad (10)$$

(c) *Precedence Constraints* : These constraints guarantee that for an operation o_i, all its predecessors are scheduled in an earlier control step and its successors are scheduled in an later control step, and are modeled as, $\forall i, j, o_i \in Pred_{o_j}$,

$$\sum_v\sum_f\sum_{d=S_i}^{E_i} dx_{i,d,v,f} - \sum_v\sum_f\sum_{e=S_j}^{E_j} ex_{j,e,v,f} \le -1 \quad (11)$$

Input : DFG, Constraints, Voltage and Freq. Levels, Delays
Output : Scheduled DFG, f_{base}, N, cfi_c, Power estimates
Step 1: Construct effective switching capacitance look-up table.
Step 2: Calculate the switching activities for each node.
Step 3: Find ASAP and ALAP schedule of the UDFG.
Step 4: Determine the mobility graphs for different schemes.
Step 5: Calculate operating frequency of FUs using delays.
Step 6: Model the ILP formulations of DFG using AMPL [2].
Step 7: Solve the ILP formulations using LP-Solve.
Step 8: Obtain the scheduled DFG.
Step 9: Determine f_c, f_{base} and cfi_c for MVDFC scheme [7].
Step 10: Estimate the power and delay of the scheduled DFG.

Fig. 1. Scheduling for MPG minimization

(d) *Resource Constraints* : These constraints ensure that no control step contains more than $M_{k,v}$ operations of type k operating at voltage v, and are enforced as, $\forall c$ and $\forall v$,

$$\sum_{i\in F_{k,v}}\sum_f x_{i,c,v,f} \le M_{k,v} \quad (12)$$

(e) *Frequency Constraints* : This set ensures that if a functional unit is operating at higher voltage level then it can be scheduled in a lower frequency control step, whereas, a functional unit is operating at lower voltage level then it can not be scheduled in a higher frequency control step. We write these constraints as, $\forall i,\ 1 \le i \le O$, $\forall c,\ 1 \le c \le N$, if $f < v$, then $x_{i,c,v,f} = 0$.

(f) *Power Gradient Constraints* : To eliminate the non-linearity introduced due to the absolute function, we introduce these constraints $\forall c,\ 2 \le c \le N$,

$$\begin{aligned} &\sum_{i\in F_{k,v}}\sum_v\sum_f x_{i,c,v,f} * P(C_{swi},v,f) \\ &-\sum_{i\in F_{k,v}}\sum_v\sum_f x_{i,c-1,v,f} * P(C_{swi},v,f) \le PG_p \end{aligned} \quad (13)$$

The PG_p is peak power gradient constraint added to the objective function and minimized alongwith it.

The target architecture model assumed by the scheduling schemes is same as the one used in [3]. All functional units have one register each and one multiplexor. The register and the multiplexor operate at the same voltage level as that of the functional units. A controller decides which of the functional units are active in each control step and those that are not active are disabled using the multiplexors. For MVDFC scheme, the controller has a storage unit to store the parameters, cycle frequency index (cfi_c) obtained from the scheduling, which serves as clock dividing factor for the dynamic clocking unit. The cycle frequency f_c is generated dynamically and a functional unit operating at one of the supply voltages is activated.

The inputs to the algorithm are an unscheduled data flow graph, the resource constraints, the number of allowable voltage levels, the number of allowable frequencies, delays at different voltage levels. The delays of level converters is represented in the form of a matrix that shows the delay in converting one voltage level to another voltage level. The scheduling algorithm (Fig. 1) determines the proper time stamp for each operation, f_{base}, cfi_c and voltage level such that the function MPG is minimum.

TABLE I

POWER ESTIMATES FOR BENCHMARKS

	MPG Estimates (mW)					Peak Power (%)		Average Power (%)		PDP (%)	
	MPG_S	MPG_D	ΔMPG_D	MPG_M	ΔMPG_M	ΔP_{p_D}	ΔP_{p_M}	ΔP_{a_D}	ΔP_{a_M}	ΔPDP_D	ΔPDP_M
1	2	3	4	5	6	7	8	9	10	11	12
e	8.42	2.11	74.94	5.96	29.22	73.61	0	72.80	22.91	54.58	0
x	8.42	2.11	74.94	5.97	29.10	73.61	20.83	72.80	21.56	54.58	0
p	8.42	2.06	75.53	2.17	74.23	73.61	47.22	72.12	36.68	53.56	0
f	4.26	1.11	73.94	3.53	17.14	73.61	0	73.47	15.65	52.24	0
i	6.42	1.72	73.21	4.54	29.28	73.61	47.22	73.47	12.93	52.24	0
r	4.26	1.08	74.65	3.00	29.58	73.61	45.90	72.9	24.72	51.22	0
i	8.56	2.92	65.89	4.41	48.48	65.74	31.48	68.33	18.78	52.24	0
i	8.56	2.24	73.83	2.71	68.34	73.61	47.22	72.96	30.13	59.60	0
r	4.26	1.08	74.65	1.27	70.19	73.61	47.22	72.34	34.13	55.71	0
h	8.49	2.85	66.43	3.53	58.42	65.74	31.48	69.26	32.55	46.09	0
a	8.56	2.19	74.42	4.52	47.20	73.60	47.20	73.18	30.14	53.06	0
l	4.26	1.06	75.12	1.63	61.74	73.33	45.35	72.71	24.64	50.85	0
a	5.66	1.46	74.20	2.92	48.41	73.59	0	74.00	22.00	59.40	0
r	5.66	1.46	74.20	3.00	47.00	73.59	0	74.00	20.44	59.40	0
f	5.66	1.40	75.27	2.97	47.53	73.02	0	71.33	18.89	57.20	0
Average Results			**73.42**		**47.10**	**72.50**	**27.41**	**72.38**	**24.41**	**54.13**	**0**

IV. RESULTS AND CONCLUSIONS

The ILP based schedulers for all schemes are tested with five benchmark circuits [8]. The following notations are used to express results : S : subscript used for SVSF operation, D : subscript used for MVDFC operation, M : subscript used for MVMC operation, T : the critical path delay, $PDP = P_a * T$: the power delay product (in nJ), $\Delta P_{p_D} = \frac{(P_{p_S} - P_{p_D})}{P_{p_S}} * 100$: % peak power reduction, $\Delta P_{p_M} = \frac{(P_{p_S} - P_{p_M})}{P_{p_S}} * 100$: % peak power reduction, $\Delta PDP_D = \frac{(PDP_S - PDP_D)}{PDP_S} * 100$: % PDP reduction, and $\Delta PDP_M = \frac{(PDP_S - PDP_M)}{PDP_S} * 100$: % PDP reduction. We use the look-up table method for average switching capacitance calculation. The look-up table construction consists of two phases, such as input pattern generation and cell characterization. We generate the primary input signal of different correlations and perform the characterization of the physical implementations of the library modules available in [6]. Whenever necessary, we used interpolation method to find the average switching capacitance for any other values of $({\alpha_i}^1, {\alpha_i}^2)$ pairs that does not exist in the look-up table. The above generated signals are propagated through different operators in the DFG and the average switching activities are calculated.

The schedulers were tested using different sets of resource constraints : (R1) multipliers (2 at $2.4V$ and 1 at $3.3V$) and ALUs (1 at $2.4V$ and 1 at $3.3V$), (R2) multipliers (3 at $2.4V$) and ALUs (1 at $2.4V$ and 1 at $3.3V$), and (R3) multipliers (2 at $2.4V$) and ALUs (2 at $3.3V$). The number of allowable voltage levels being two ($2.4V, 3.3V$) and maximum number of allowable frequencies being three. The experimental results for various benchmark circuits are reported in Table I for all three schemes. In the table, the results are expressed in the order R1, R2, and R3 for each benchmark. The power estimations include the power consumption of the overheads. In case of MVDFC scheduler the frequencies found are $4.5MHz, 9MHz$ and $18MHz$. For MVMC and SVSF schedulers the operating frequency is $9MHz$.

This paper addressed power fluctuation reduction at behavioral level using low power datapath scheduling techniques. We used ILP based optimizations for the three modes of datapath operations. In dynamic frequency clocking scheme significant reduction could be achieved in mean power gradient, peak power and average power alongwith reductions in power delay product. The results clearly indicate that the dynamic frequency clocking is a better scheme than the multicycling approach for power minimization.

REFERENCES

[1] J. M. Chang and M. Pedram. Energy minimization using multiple supply voltages. *IEEE Trans. on VLSI Systems*, 5(4):436-443, Dec 1997.

[2] R. Fourer, D. M. Gay, and B. W. Kernighan. *AMPL : A Modeling Language for Mathematical Programming.* Brooks/Cole-Thomson Learning, CA, USA, 2003.

[3] M. Johnson and K. Roy. Datapath scheduling with multiple supply voltages and level converters. *ACM TODAES*, 2(3):227-248, July 1997.

[4] R. S. Martin and J. P. Knight. Optimizing power in asic behavioral synthesis. *IEEE Design & Test of Computers*, 13(2):58-70, Summer 1996.

[5] B. A. McCarl and T. H. Spreen. *Applied Mathematical Programming using Algebric Systems.* Online Book at : http://agecon.tamu.edu/faculty/mccarl/regbook.htm, 1997.

[6] S. P. Mohanty and N. Ranganathan. Energy efficient scheduling for datapath synthesis. In *Proc. of Intl. Conf. on VLSI Design*, pp. 446-451, 2003.

[7] S. P. Mohanty and N. Ranganathan. A framework for energy and transient power reduction during behavioral synthesis. In *Proc. of Intl. Conf. on VLSI Design*, pp. 539-545, 2003.

[8] S. P. Mohanty, N. Ranganathan, and S. K. Chappidi. Simultaneous Peak and Average Power Minimization During Datapath Scheduling for DSP Processors. In *Proc. of GLSVLSI*, pp. 215-220, 2003.

[9] M. J. Panik. *Linear Programming : Mathematics, Theory and Practice.* Kluwer Academic Publishers, 1996.

[10] V. Raghunathan, S. Ravi, A. Raghunathan, and G. Lakshminarayana. Transient power management through high level synthesis. In *Proc. of ICCAD*, pp. 545-552, 2001.

[11] W. T. Shiue. High level synthesis for peak power minimization using ILP. In *Proc. of IEEE Intl. Conf. on ASSAP*, pp. 103-112, 2000.