ILP Models for Energy and Transient Power Minimization During Behavioral Synthesis

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

*The reduction of peak power, peak power differential, average power and energy are equally important in the design of low-power battery driven portable applications. In this paper, we introduce a parameter called "cycle power function" (*CPF-DFC*) that captures the above power characteristics in the context of multiple supply voltage (MV) and dynamic frequency clocking (DFC) based designs. Further, we present ILP formulations for the minimization of* CPF-DFC *during datapath scheduling. We conducted experiments on selected high-level synthesis benchmarks for various resource constraints. Experimental results show that significant reduction in power, energy, and energy delay product, can be obtained using the proposed method.*

1 Introduction

Low power design and synthesis is driven by several factors such as battery life, increase in chip density, thermal considerations and environmental concerns, etc. In the work reported in [1], peak power reduction is achieved through simultaneous assignment and scheduling. ILP based models to minimize peak power and peak area have been proposed in [2] for latency constrained scheduling. The authors in [3] propose the use of data monitor operations for simultaneous peak power reduction and peak power differential. In [4], a heuristic based scheme is proposed that minimizes a parameter called "cycle power profile function" that captures the peak power, peak power differential, and average power. In [5], ILP-based datapath scheduling schemes are proposed to minimize both peak and average power. In [6], ILP formulations of the multiple voltage scheduling problem is given.

Most works discussed above address either average power, energy or peak power, but do not address all of the power parameters (average power, energy, peak power, and peak power differential) together. In this paper, we describe a framework for simultaneous minimization of energy, average power, peak power, and peak power differential using ILP-based minimization.

2 Cycle Power Function (*CPF-DFC*)

In this section, we define a parameter called *CPF-DFC*, which captures the peak power, the peak power differential and the average power of the datapath circuit. Further, we study its nonlinear behavior and modify it so as to use ILP for its minimization. The datapath is represented as a sequencing data flow graph (DFG). The following definitions and notations are needed.

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
o_i	: any operation, where $1 \leq i \leq O$,
P_c	: the power consumption of all functional units active in step c (cycle power consumption)
P_{peak}	: peak power consumption for DFG ($= max(P_c)_{\forall c}$)
P	: mean power consumption of the DFG
P_{norm}	: normalised mean power consumption of the DFG
DP_c	: cycle difference power
DP_{peak}	: peak differential power for DFG
DP	: mean of the cycle difference power for all c in DFG
DP_{norm}	: normalised mean of the cycle difference power
$FU_{k,v}$	: any functional unit of type k operating at voltage v
FU_i	: any functional unit $FU_{k,v}$ needed by o_i for its execution ($o_i \in FU_{k,v}$)
$FU_{i,c}$	: any functional unit FU_i active in control step c
R_c	: total number of functional units active in step c
α_i	: switching activity of resource FU_i
V_i	: operating voltage of resource FU_i
C_i	: load capacitance of resource FU_i
f_c	: frequency of control step c
${\alpha_i}^1, {\alpha_i}^2$	: the average switching activities on the first and second input operands of resource FU_i
$C_{sw\,i}$	: a measure of the effective switching capacitance of the functional unit FU_i

Using the dynamic energy model proposed in [7], the effective switching capacitance can be expressed as,

$$\alpha_i C_i = C_{sw\,i}({\alpha_i}^1, {\alpha_i}^2) \quad (1)$$

It should be noted that in the above switching model (Eqn. 1), the input pattern dependencies can be handled.

The power consumption for any control step c is given by Eqn. 2. This is the total power consumption of all the functional units active in control step c and also includes the power consumption of the level converters. If a current resource is driven by a resource operating at lower voltage, then level converters are needed as additional resources operating in the current cycle.

$$P_c = \sum_{i=1}^{R_c} \alpha_{i,c} C_{i,c} V_{i,c}^2 f_c = \sum_{i=1}^{R_c} C_{sw\,i,c} V_{i,c}^2 f_c \quad (2)$$

The peak power consumption of a DFG is the maximum power consumption over all control steps, as given below.

$$P_{peak} = max(P_c)_{\forall c} = max\left(\sum_{i=1}^{R_c} C_{swi,c} V_{i,c}^2 f_c\right)_{\forall c} \quad (3)$$

The mean cycle power (P) which is an unbiased estimate of the average power consumption of the DFG, is defined as,

$$P = \frac{1}{N}\sum_{c=1}^{N} P_c = \frac{1}{N}\sum_{c=1}^{N}\left(\sum_{i=1}^{R_c} C_{swi,c} V_{i,c}^2 f_c\right) \quad (4)$$

The normalised mean cycle power (P_{norm}) is calculated by dividing P by maximum cycle power (P_{peak}).

The cycle difference power (DP_c) for any control step can be defined as the absolute deviation of the cycle power from the mean cycle power consumption of the DFG. This is a measure of the cycle power fluctuation of the DFG.

$$DP_c = |P - P_c| \quad (5)$$

The peak differential power (DP_{peak}) that characterises the maximum power fluctuation of DFG over all control steps is,

$$DP_{peak} \quad = \quad max(|P - P_c|)_{\forall c} \quad (6)$$

The mean cycle difference power (DP) is calculated as sample mean of DP_c. This is a measure of the power spread or distribution of the cycle power over all control steps.

$$DP \quad = \quad \frac{1}{N}\sum_{c=1}^{N} DP_c = \frac{1}{N}\sum_{c=1}^{N} |P - P_c| \quad (7)$$

The cycle power function *CPF-DFC* is modeled as an equally weighted sum of the normalized mean cycle power and the normalized mean cycle difference power.

$$CPF\text{-}DFC(P_{norm}, DP_{norm}) = P_{norm} + DP_{norm} \quad (8)$$

The *CPF-DFC* has a value in the range [0,2]. In terms of peak cycle power (P_{peak}) and peak cycle difference power (DP_{peak}), it can be expressed as :

$$CPF\text{-}DFC \quad = \quad \frac{\frac{1}{N}\sum_{c=1}^{N} P_c}{P_{peak}} + \frac{\frac{1}{N}\sum_{c=1}^{N} |P - P_c|}{DP_{peak}} \quad (9)$$

The above function can serve as the objective function for low power datapath scheduling. *The minimization of this objective function using multiple supply voltages, dynamic frequency clocking can reduce power and energy.* From Eqns. 9 we observe that *CPF-DFC* is a *non-linear* function of four parameters, such as, P, P_{peak}, DP, and DP_{peak}. The absolute function in the numerator also contributes to the non-linearity. The complex behavior of the function is also contributed by the denominator parameters, P_{peak} and DP_{peak}.

In this work, we aim at developing ILP-based model for minimization of *CPF-DFC*. We alter *CPF-DFC* in order to simplify the ILP-based model. It is known that the denominator parameters, P_{peak} equals to $max(P_c)_{\forall c}$ and the DP_{peak} equals to $max(|P - P_c|)_{\forall c}$. It is evident that $|P - P_c|$ is upper bounded P_c for all control steps c, since $|P - P_c|$ is a measure of mean difference error of P_c. Thus, we conclude that DP_{peak} is upper bounded by P_{peak}. We modify the *CPF-DFC* by substituting DP_{peak} with P_{peak} and define modified *CPF-DFC* (denoted as *CPF-DFC**) as follows.

$$CPF\text{-}DFC^* \quad = \quad \frac{\frac{1}{N}\sum_{c=1}^{N} P_c + \frac{1}{N}\sum_{c=1}^{N} |P - P_c|}{P_{peak}} \quad (10)$$

Unlike *CPF-DFC*, *CPF-DFC** is dependent on three factors, P, P_{peak} and DP, which helps in reducing the complexity of the ILP formulations.

3 ILP Formulations

We first address the transformations required to derive linear models of nonlinear functions. The general form of programming with absolute non-linearity can be represented as below [8, 9].

$$\begin{array}{ll} \text{Minimize :} & \sum_i |y_i| \\ \text{Subject to :} & y_i + \sum_j a_{ij} * x_j \leq b_i,\ \forall i,\ x_j \geq 0,\ \forall j \end{array} \quad (11)$$

Where, y_i is the deviation between the prediction and observation. Let, y_i be represented as the difference of two non-negative variables, $y_i = y_i^1 - y_i^2$. The formulations in Eqn. 11 and 12 are equivalent.

$$\begin{array}{ll} \text{Minimize :} & \sum_i y_i^1 + y_i^2 \\ \text{Subject to :} & y_i^1 - y_i^2 + \sum_j a_{ij} * x_j \leq b_i, \quad \forall i \\ & x_j \geq 0, \quad \forall j, \quad y_i^1, y_i^2 \geq 0, \quad \forall i \end{array} \quad (12)$$

The general expression [9] for the LP formulation involving fractions is considered below.

$$\begin{array}{ll} \text{Minimize :} & \frac{\sum_j c_j * x_j}{\sum_j d_j * x_j} \\ \text{Subject to :} & \sum_j a_{ij} * x_j \leq b_i, \quad \forall i, \quad x_j \geq 0, \quad \forall j \end{array} \quad (13)$$

Where, c_j and d_j are known constants and the denominator $\sum_j d_j * x_j$ is strictly positive. Let us assume new variables,

$$z_0 = \left|d_0 + \sum_j d_j * x_j\right|^{-1} \quad \text{and} \quad x_j = \frac{z_j}{z_0} \quad (14)$$

Using the above transformation, the original formulation in Eqn. 13 can be modified to the following.

$$\begin{array}{ll} \text{Minimize :} & c_0 * z_0 + \sum_j c_j * z_j \\ \text{Subject to :} & \sum_j a_{ij} * z_j - b_i * z_0 \leq b_i,\ \forall i \\ & \sum_j d_j * z_j + d_0 * z_0 = 1,\ z_0, z_j \geq 0,\ \forall j \end{array} \quad (15)$$

The problems defined in Eqn. 13 and 15 are equivalent. On solving the problem in Eqn. 15, we substitute, $z_j = x_j * z_0$ to get the results for x_j.

We now discuss the ILP models for minimization of *CPF-DFC** using MVDFC [10, 11]. The following notations are used in ILP formulations of Eqn. 10.

$M_{k,v}$	: maximum number of functional units of type k operating at voltage level v ($FU_{k,v}$)
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 FU_i at voltage v and operating frequency f used by o_i for its execution
$x_{i,c,v,f}$	: decision variable which takes the value of 1 if operation o_i is scheduled in step c using $F_{k,v}$ and c has frequency f_c

(a) *Objective Function* : The objective is to minimize the modified cycle power function *CPF-DFC**. Using Eqn. 10, this can be restated as :

$$\text{Minimize :} \quad \frac{\frac{1}{N}\sum_{c=1}^{N} P_c + \frac{1}{N}\sum_{c=1}^{N} |P - P_c|}{P_{peak}} \tag{16}$$

We first remove the non-linearity because of the fraction by expressing the denominator as a constraint and transform Eqn. 16 to the following.

$$\begin{array}{ll} \text{Minimize :} & \frac{1}{N}\sum_{c=1}^{N} P_c + \frac{1}{N}\sum_{c=1}^{N} |P - P_c| \\ \text{Subject to :} & \text{Peak power constraints} \end{array} \tag{17}$$

However, this problem still has the non-linearity in it because of the absolute function. This can be converted to an equivalent problem using the transformation discussed before.

$$\begin{array}{ll} \text{Minimize :} & \frac{1}{N}\sum_{c=1}^{N} P_c + \frac{1}{N}\sum_{c=1}^{N} (P + P_c) \\ \text{Subject to :} & \text{Modified peak power constraints} \end{array} \tag{18}$$

The "peak power constraint" in Eqn. 17 and the "modified peak power constraint" in Eqn. 18 will be discussed in later part of the section. Using Eqn. 4 in Eqn. 18, using decision variables, and assuming that $P^*(C_{swi}, v, f) = P(C_{swi}, v, f) * \left(\frac{3}{N}\right)$, we get the following.

$$\begin{array}{ll} \text{Minimize :} & \sum_c \sum_{i \in F_{k,v}} \sum_v \sum_f x_{i,c,v,f} P^*(C_{swi}, v, f) \\ \text{Subject to :} & \text{Modified peak power constraints} \end{array} \tag{19}$$

(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 supply voltage and operating frequency. We represent them as, $\forall i$, $1 \leq i \leq O$,

$$\sum_c \sum_v \sum_f x_{i,c,v,f} \quad = \quad 1 \tag{20}$$

(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 a later control step. These are modeled as, $\forall i, j, o_i \in Pred_{o_j}$,

$$\begin{array}{l} \sum_v \sum_f \sum_{d=S_i}^{E_i} d * x_{i,d,v,f} \\ -\sum_v \sum_f \sum_{e=S_j}^{E_j} e * x_{j,e,v,f} \leq -1 \end{array} \tag{21}$$

Input : UDFG,resource constraints, delays, voltage and frequency levels
Output : scheduled DFG, base frequency, N, cfi_c, power
Step 1 : Construct switching capacitance look up table.
Step 2 : Calculate the switching activities at each node.
Step 3 : Find ASAP and ALAP schedules for UDFG.
Step 5 : Determine the mobility graph of each node.
Step 6 : Get the ILP formulations using AMPL [12].
Step 7 : Solve the ILP formulations using LP-Solve.
Step 8 : Find the scheduled DFG.
Step 9 : Determine f_{base} and cfi_c, and estimate power.

Figure 1: The Scheduling Algorithm

(d) *Resource Constraints* : These constraints make sure that no control step contains more than $F_{k,v}$ operations of type k operating at voltage v. These can be enforced as, $\forall c$, $1 \leq c \leq N$ and $\forall v$,

$$\sum_{i \in F_{k,v}} \sum_f x_{i,c,v,f} \quad \leq \quad M_{k,v} \tag{22}$$

(e) *Frequency Constraints* : This set ensures that if a functional unit is operating at a 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, $\forall i, 1 \leq i \leq O$, $\forall c, 1 \leq c \leq N$, if $f < v$, then $x_{i,c,v,f} = 0$.

(f) *Peak Power Constraints* : As discussed before, with reference to the Eqn. 16 and 17, these constraints are introduced to eliminate the fractional non-linearity of the objective function and are enforced as, $\forall c$, $1 \leq c \leq N$,

$$\sum_{i \in F_{k,v}} \sum_v \sum_f x_{i,c,v,f} * P(C_{swi}, v, f) \leq P_{peak} \tag{23}$$

(g) *Modified Peak Power Constraints* : To eliminate the non-linearity introduced due to the absolute function, we modify the above constraints (Eqn. 18), $\forall c$, $1 \leq c \leq N$,

$$\begin{array}{l} \frac{1}{N}\sum_c \sum_{i \in F_{k,v}} \sum_v \sum_f x_{i,c,v,f} * P(C_{swi}, v, f) \\ -\sum_{i \in F_{k,v}} \sum_v \sum_f x_{i,c,v,f} * P(C_{swi}, v, f) \leq P^*_{peak} \end{array} \tag{24}$$

The P^*_{peak} is a modified peak constraint which is added to the objective function and minimized alongwith it.

4 Scheduling Results and Conclusions

The scheduler which minimizes *CPF-DFC** is outlined in Fig. 1. The target architecture model assumed for the scheduling schemes is from [6]. Level converters are used when a low-voltage functional unit drives a high-voltage functional unit. 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. The controller has a storage unit to store cycle frequency index (cfi_c) values obtained from scheduling. This serves as the clock dividing factor for the dynamic clocking unit. The cycle frequency f_c

Table 1: Power, energy and energy delay product estimates for benchmarks

	RC	P_{pS}	P_{pD}	ΔP_p	P_{mS}	P_{mD}	ΔDP	P_S	P_D	ΔP	E_S	E_D	ΔE	EDP_S	EDP_D	ΔEDP
		mW	mW	%	mW	mW	%	mW	mW	%	nJ	nJ	%	$10^{-15}Js$	$10^{-15}Js$	%
E	1	17.28	4.56	73.61	0.46	0.35	74.97	8.87	2.42	72.72	2.96	1.57	46.8	0.99	0.87	11.34
X	2	17.28	4.56	73.61	0.46	0.35	74.97	8.87	2.42	72.72	2.96	1.57	46.8	0.99	0.87	11.34
P	3	17.28	4.56	73.61	0.46	0.9	78.24	8.87	2.61	70.57	2.96	1.6	46.0	0.99	0.8	18.98
F	1	17.51	4.62	73.62	0.23	0.12	73.96	8.82	2.35	73.36	4.9	2.6	47.20	2.7	2.3	15.52
I	2	25.92	6.84	73.61	0.23	0.12	73.84	8.82	2.36	73.24	4.9	2.6	47.20	2.7	2.0	26.09
R	3	17.51	4.67	73.33	0.23	0.45	75.58	8.82	2.5	71.66	4.9	2.6	46.22	2.7	2.0	24.71
H	1	17.51	4.62	73.62	0.46	0.35	74.96	13.25	3.55	73.21	5.9	3.12	47.0	2.62	2.43	7.25
A	2	26.15	6.90	73.61	0.46	0.35	74.50	13.25	3.55	73.21	5.9	3.12	47.0	2.62	2.43	7.25
L	3	17.74	4.78	73.05	0.46	0.9	76.97	13.25	3.73	71.85	5.9	3.17	46.2	2.62	2.23	12.55
I	1	25.92	8.88	65.74	0.23	0.12	65.9	11.03	3.5	68.36	4.9	3.05	37.7	2.18	2.04	6.57
I	2	25.92	6.84	73.61	0.23	0.12	73.84	11.03	2.98	72.98	4.9	2.6	47.96	2.18	1.73	20.44
R	3	17.51	4.67	73.34	0.23	0.45	75.58	8.82	2.57	70.86	4.9	2.64	46.22	2.72	2.05	24.71
A	1	8.87	2.34	73.62	0.23	0.12	74.1	4.5	1.22	72.9	5.0	2.64	47.2	5.56	4.4	20.83
R	2	8.87	2.34	73.62	0.23	0.12	74.1	4.5	1.22	72.9	5.0	2.64	47.2	5.56	4.4	20.83
F	3	8.87	2.39	73.05	0.23	0.45	77.6	4.5	1.4	68.9	5.0	2.74	45.3	5.56	3.8	31.63

is generated dynamically and a functional unit operating at one of the supply voltages is activated. The scheduling algorithm was tested with five benchmark circuits : EXP, FIR, HAL, IIR, and ARF [5]. The following sets of resource constraints are used in the experiment.

RC	Multipliers	ALUs
(RC1)	2 at $2.4V$, 1 at $3.3V$	1 at $2.4V$, 1 at $3.3V$
(RC2)	3 at $2.4V$	1 at $2.4V$, 1 at $3.3V$
(RC3)	2 at $2.4V$	2 at $3.3V$

We perform the characterization of the physical implementations of the library modules available in [14] by applying the input patterns generated using ARMA model [13] for some values of $(\alpha_i{}^1, \alpha_i{}^2)$ pairs. The above generated signals are propagated through different operators in the DFG and the average switching activities are calculated as described in [13]. The experimental results for various benchmark circuits are reported in Table 1. The notations used to express the results are : S = single voltage operation, D = multiple voltages and dynamic clocking, m = minimum value, Δ = % reduction, E = energy, p = peak, EDP = energy delay product, and ΔDP = $\left(\frac{(P_{pS}-P_{mS})-(P_{pD}-P_{mD})}{(P_{pS}-P_{mS})} * 100\right)$. The power/energy estimation include the power consumption of the overheads, such as level converters (data taken from [14]). The results are reported for two supply voltages and the frequencies are found out to be $(4.5MHz, 9MHz, 18MHz)$.

The datapath scheduling algorithm described in this paper is particularly useful for synthesizing data intensive application specific integrated circuits. The algorithm attempts to optimize energy and power while maintaining performance. The scheduling algorithm assumes number of different types of resources at each voltage levels and the number of allowable frequencies as resource constraints. The energy delay product was estimated to keep track of the effect of scheduling algorithms on circuit performance. The scheduling resulted in reduction of EDP for all benchmarks and all resource constraints, which shows its effectiveness. The effectiveness of the scheduling schemes in the context of pipelined datapath and control intensive applications, needs to be investigated.

References

[1] R. S. Martin and J. P. Knight, "Optimizing Power in ASIC Behavioral Synthesis," *IEEE Design and Test of Computers*, vol. 13, no. 2, pp. 58-70, Summer 1996.

[2] W. T. Shiue and C. Chakrabarti, "ILP Based Scheme for Low Power Scheduling and Resource Binding," *Proc. of ISCAS*, 2000, pp. 279-282.

[3] V. Raghunathan, et. al., "Transient Power Management through High Level Synthesis," in *Proc. of ICCAD*, 2001, pp. 545-552.

[4] 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*, Jan 2003, pp. 539-545.

[5] 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*, Apr. 2003, pp. 215-220.

[6] M. Johnson and K. Roy, "Datapath Scheduling with Multiple Supply Voltages and Level Converters," *ACM TODAES*, vol. 2, no. 3, pp. 227-248, July 1997.

[7] J. M. Chang and M. Pedram, "Energy Minimization using Multiple Supply Voltages," *IEEE TVLSI*, vol. 5, no. 4, pp. 436-443, Dec 1997.

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

[9] 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.

[10] N. Ranganathan, et. al., "A Linear Array Processor with Dynamic Frequency Clocking for Image Processing Applications," *IEEE Trans. on CSVT*, vol. 8, no. 4, pp. 435-445, August 1998.

[11] I. Brynjolfson and Z. Zilic, "Dynamic Clock Management for Low Power Applications in FPGAs," in *Proc. of IEEE Custom Integrated Circuits Conf.*, 2000, pp. 139-142.

[12] R. Fourer, D. Gay, and B. Kernighan, *AMPL: A Modeling Language for Mathematical Programming*, Thomson Brooks Cole, 2003.

[13] S. Ramprasad, N. R. Shanbhag, and I. N. Hajj, "Analytical Estimation of Signal Transition Activity from Word-Level Statistics," *IEEE Trans. on CAD*, vol. 16, no. 7, pp. 718-733, Jul 1997.

[14] S. P. Mohanty and N. Ranganathan, "Energy Efficient Scheduling for Datapath Synthesis," in *Proc. of Intl. Conf. on VLSI Design*, Jan 2003, pp. 446-451.