

AN ILP-BASED SCHEDULING SCHEME FOR ENERGY EFFICIENT HIGH PERFORMANCE DATAPATH SYNTHESIS

Saraju P. Mohanty, N. Ranganathan and Sunil K. Chappidi

Department of Computer Science & Engineering
Nanomaterial and Nanomanufacturing Research Center
University of South Florida, Tampa, FL 33620
{smohanty,ranganat,chappidi}@csee.usf.edu

ABSTRACT

In this paper, we describe an integer linear programming (ILP) based datapath scheduling algorithm which uses both multiple supply voltages and dynamic frequency clocking for power optimization. The scheduling technique assumes the number and type of different functional units as resource constraints and minimizes the energy delay product (EDP). The energy savings directly comes from the use of multiple supply voltages and the performance improvement from dynamic frequency clocking. The algorithm has been applied to various high level synthesis benchmark circuits under different resource constraints. The experimental results show that under various resource constraints using two supply voltage levels (5.0V, 3.3V), the average energy reduction is 55% and the average EDP reduction is 13%.

1. INTRODUCTION

The need for low power synthesis is increasing due to demand of portable systems, thermal considerations, environmental concerns and reliability issues. These factors affect the battery life, cooling and packaging costs, and use of natural resources. While the first three factors relate to average power dissipation, the last factor is mainly influenced by the peak power dissipation of CMOS circuit. For energy efficient (longer battery life) and high performance (smaller delay) circuits, the energy delay product (EDP) has to be minimized. During low power synthesis, several low power sub-tasks, such as, scheduling, allocation and binding are performed. In this work, we focus on low power scheduling.

Various low power datapath scheduling techniques have been reported in the literature. A scheduling algorithm using "shut-down" technique, multiplexor reordering and pipelining has been developed in [1]. This system minimizes switching capacitance for power minimization. Scheduling and resource binding algorithms are developed in [2], which reduces power by reducing activity of functional units by minimizing transition of operands. A multiple supply voltage

scheduling algorithm called MOVER is presented in [3], which uses ILP formulations. A dynamic programming technique for multiple supply voltage scheduling is discussed in [4]. A low power datapath synthesis system called SCALP is discussed in [5], which uses voltage scaling and capacitance reduction. A time constrained multiple voltage scheduling technique is proposed in [6]. A resource constrained scheduling algorithm with multiple supply voltages is given in [7] which helps in reducing power using multiple supply voltages. In [8], a resource and a latency constrained list-based scheduling algorithms with multiple supply voltages are discussed.

The above scheduling techniques are based on a single clock frequency and consider multiple supply voltages, voltage scaling, capacitance reduction, and switching activity reduction. Recently, use of dynamic frequency clocking or frequency scaling alongwith multiple supply voltages has been investigated in developing low power or high performance VLSI systems [9, 10, 11]. It is evident that by scaling frequency and voltage in a coordinated manner, both energy and power can be saved while maintaining performance. In [12], a heuristic algorithm is proposed for low power scheduling using multiple supply voltages and dynamic clocking. In this work we describe an ILP based algorithm that uses multiple supply voltages and dynamic clocking to minimize EDP.

2. ILP FORMULATION TO MINIMIZE EDP

For a data flow graph (DFG), let us assume : (i) NCC = total number of control steps, (ii) NR_c = number of resources active in control step c , (iii) f_c = cycle frequency, (iv) $\alpha_{i,c}$ = switching at resource i operating in control step c , (v) $C_{i,c}$ = load capacitance of resource i operating in control step c and (vi) $V_{i,c}$ = operating voltage of resource i operating in control step c . The total energy consumption of the whole DFG is given by Eqn. 1. The level converters are considered as resources operating in the control step in which it needs to step up signal. The dynamic clocking unit (DCU) is responsible

for generating dynamic clock is considered as a resource operating in all the control steps. The energy consumptions of the level converters and DCU are included.

$$E_D = \sum_{c=1}^{NCC} \sum_{i=1}^{NR_c} \alpha_{i,c} C_{i,c} V_{i,c}^2 \quad (1)$$

The critical path delay of the DFG is given by the summation of the inverse of the clock frequencies. The energy delay product can be calculated as the product of the total energy consumption and the critical path delay as shown in the following equation.

$$EDP_D = \sum_{c=1}^{NCC} \sum_{i=1}^{NR_c} \frac{\alpha_{i,c} C_{i,c} V_{i,c}^2}{f_c} \quad (2)$$

This would serve as an objective function for the scheduling algorithm. It may be noted that for single frequency and single frequency mode of operation, $V_{i,c}$ and f_c are the same for any clock cycle (c) and operation (i).

In order to formulate an ILP based model for Eqn. 2 and hence a scheduling scheme for the DFG, we use the following notations : (i) N = total number of operations in the DFG excluding the source and sink nodes (NO-OPs), (ii) o_i = is any operation i , $1 \leq i \leq N$, (iii) $FU_{k,v}$ = functional unit of type k operating at voltage level v , (iv) $M_{k,v}$ = maximum number of functional unit of type k operating at voltage level v , (v) S_i = ASAP time stamp for the operation o_i , (vi) E_i = ALAP time stamp for the operation o_i , (vii) $EDP(i, v, f)$ = energy delay product of operation o_i at voltage level v and operating frequency f , (viii) $x_{i,c,v,f}$ = decision variable which takes the value of 1 is operation o_i is scheduled in control step c using the functional unit $FU_{k,v}$ and c has frequency f_c .

(a) *Objective Function* : The objective function minimizes the total energy delay product of the whole DFG represented in Eqn. 2. Using the decision variable $x_{i,c,v,f}$, we write the objective function as follows. Minimize :

$$EDP_D = \sum_i \sum_c \sum_v \sum_f x_{i,c,v,f} * EDP(i, v, f) \quad (3)$$

(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, $\forall i, 1 \leq i \leq N$,

$$\sum_c \sum_v \sum_f x_{i,c,v,f} = 1 \quad (4)$$

(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. These are modelled as, $\forall i, j, o_i \in Pred_{o_j}$,

$$\sum_v \sum_f \sum_{d=S_i}^{E_i} d * x_{i,d,v,f} - \sum_v \sum_f \sum_{e=S_j}^{E_j} e * x_{j,e,v,f} \leq -1 \quad (5)$$

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

$$\sum_{i \in FU_{k,v}} \sum_f x_{i,c,v,f} \leq M_{k,v} \quad (6)$$

(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, $\forall i, 1 \leq i \leq N, \forall c, 1 \leq c \leq NCC$, if $f < v$, then $x_{i,c,v,f} = 0$.

3. ILP SOLUTION

In this section, we will discuss the solution of ILP formulations obtained in the previous section. We are using the same target architecture and same characterised datapath components as used in [12]. The ILP based scheduler which minimizes EDP is outlined in Fig. 1. The first step is to determine the ASAP and ALAP time stamp of each operation. The ASAP time stamp is the start time and ALAP time stamp is the finish time of each operation. These two times provide the mobility of a operation and the operation must be scheduled in this mobile range. Then the scheduler finds the ILP formulations based on the models described in section 2. The scheduler determines the cycle frequencies in step 6, which are basically the smallest frequencies of all operations scheduled in a particular cycle. Finally, we estimate the energy delay product and the energy consumptions of the whole DFG.

- Step 1:** Find ASAP and ALAP schedule of the UDFG.
- Step 2:** Determine the mobility graph of each node.
- Step 3:** Construct the ILP formulations for the DFG.
- Step 4:** Solve the ILP formulations using LP-Solve.
- Step 5:** Find the scheduled DFG.
- Step 6:** Determine the cycle frequencies.
- Step 7:** Find the energy and EDP estimates of the DFG.

Figure 1: ILP-based scheduling for low EDP

We illustrate solution of the ILP formulation with the help of DFG shown in Fig.2. The ASAP schedule is shown in

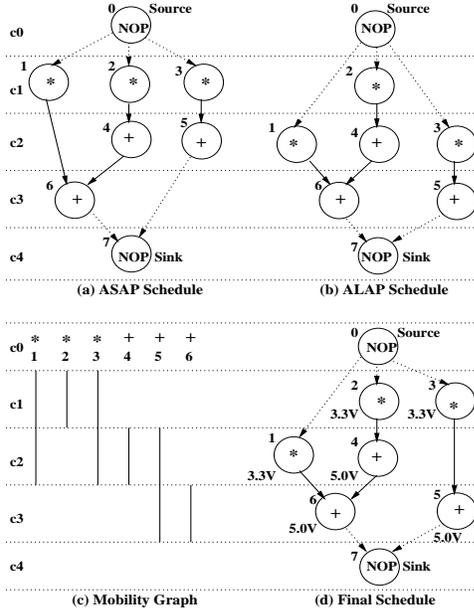


Figure 2: Example data flow graph

Fig.2(a) and the ALAP schedule is shown in Fig.2(b). From the ASAP and ALAP we obtained the mobility graph which is Fig.2(c). Using this mobility graph, for the resource constraint, two multipliers at 3.3V and two ALUs operating at 5.0V we have the ILP formulations shown in Fig.3. We solved the formulation using LP-solve and based on the results we obtained the scheduled DFG shown is Fig.2(d).

4. RESULTS AND CONCLUSIONS

We tested the ILP scheduler with selected benchmark circuits, such as, (1) Example circuit, (2) FIR filter, (3) IIR filter, (4) HAL differential equation solver and (5) Auto regressive filter. The FUs used in the processor model are ALUs and MULTs. The characterised datapath cells are borrowed from [12]. The following notations are used to express results : (i) E_S , E_M are the total energy consumption (in pJ) for single supply voltage and multiple supply voltage operations respectively, (ii) EDP_S , EDP_M and EDP_D are the energy-delay-products (in $10^{-18}Js$) for single supply voltage and single frequency, for multiple supply voltage and single frequency and for multiple supply voltage and dynamic clocking operations, respectively, (iii) the percentage energy savings is calculated as, $S_E = \frac{(E_S - E_M)}{E_S} * 100$ and (iv) the percentage EDP reduction S_{EDP} is calculated as, $S_{EDP} = \frac{(EDP_S - EDP_D)}{EDP_S} * 100$.

The scheduling scheme was simulated using the different sets of resource constraints : (1) multipliers (2 at 3.3V and 1 at 5.0V) and ALUs (1 at 3.3V and 1 at 5.0V), (2) multipliers (3 at 3.3V) and ALUs (1 at 3.3V and 1 at 5.0V), (3)

```

/* Objective Function */
min: 122.3x1111 + 244.7x1112 + 53.3x1121 + 106.7x1122 +
122.3x1211 + 244.7x1212 + 53.3x1221 + 106.7x1222 +
122.3x2111 + 244.7x2112 + 53.3x2121 + 106.7x2122 +
122.3x3111 + 244.7x3112 + 53.3x3121 + 106.7x3122 +
122.3x3211 + 244.7x3212 + 53.3x3221 + 106.7x3222 +
3.2x4211 + 6.3x4212 + 1.4x4221 + 2.8x4222 + 3.2x5211 +
6.3x5212 + 1.4x5221 + 2.8x5222 + 3.2x5311 + 6.3x5312 +
1.4x5321 + 2.8x5322 + 3.2x6311 + 6.3x6312 + 1.4x6321 +
2.8x6322;

/* Uniqueness Constraints */
x1111 + x1112 + x1121 + x1122 + x1211 + x1212 + x1221 +
x1222 = 1;
x2111 + x2112 + x2121 + x2122 = 1;
x3111 + x3112 + x3121 + x3122 + x3211 + x3212 + x3221 +
x3222 = 1;
x4211 + x4212 + x4221 + x4222 = 1;
x5211 + x5212 + x5221 + x5222 + x5311 + x5312 + x5321 +
x5322 = 1;
x6311 + x6312 + x6321 + x6322 = 1;

/* Precedence Constraints */
3x6311 + 3x6312 + 3x6321 + 3x6322 - 2x1211 - 2x1212 -
2x1221 - 2x1222 - x1111 - x1112 - x1121 - x1122 ≥ 1;
2x4211 + 2x4212 + 2x4221 + 2x4222 - x2111 - x2112 -
x2121 - x2122 ≥ 1;
3x6311 + 3x6312 + 3x6321 + 3x6322 - x4211 - x4212 -
x4221 - x4222 ≥ 1;
3x5311 + 3x5312 + 3x5321 + 3x5322 + 2x5211 + 2x5212 +
2x5221 + 2x5222 - 2x3211 - 2x3212 - 2x3221 - 2x3222 -
x3111 - x3112 - x3121 - x3122 ≥ 1;

/* Resource Constraints */
x1111 + x2111 + x3111 + x1112 + x2112 + x3112 ≤ 0;
/* mult1 */
x1121 + x2121 + x3121 + x1122 + x2122 + x3122 ≤ 2;
/* mult2 */
x1211 + x3211 + x1212 + x3212 ≤ 0; /* mult1 */
x1221 + x3221 + x1222 + x3222 ≤ 2; /* mult2 */
x4211 + x5211 + x4212 + x5212 ≤ 2; /* alu1 */
x4221 + x5221 + x4222 + x5222 ≤ 0; /* alu2 */
x5311 + x6311 + x5312 + x6312 ≤ 2; /* alu1 */
x5321 + x6321 + x5322 + x6322 ≤ 0; /* alu2 */

/* Frequency Constraints */
x1121 = 0; x1221 = 0; x2121 = 0; x3121 = 0; x3221 =
0; x4221 = 0; x5221 = 0; x5321 = 0; x6321 = 0;

/* Zero-One Type Cast */
INT
x1111, x1112, x1121, x1122, x1211, x1212, x1221, x1222,
x2111, x2112, x2121, x2122, x3111, x3112, x3121, x3122,
x3211, x3212, x3221, x3222, x4211, x4212, x4221, x4222,
x5211, x5212, x5221, x5222, x5311, x5312, x5321, x5322,
x6311, x6312, x6321, x6322;

```

Figure 3: ILP formulations for example DFG

5. REFERENCES

Table 1: Energy and EDP estimates for benchmarks

	R C	Energy Delay Products ($10^{-18} Js$)				Energy Estimates (pJ)		
		edp_S	edp_M	edp_D	S_{edp}	E_S	E_M	S_E
1	2	3	4	5	6	7	8	9
(1) e x p	1	376	332	327	13	6777	2987	55
	2	376	332	325	14	6777	2987	55
	3	376	339	329	13	6777	3051	55
	4	376	339	329	13	6777	3051	55
(2) f i r	1	624	544	541	13	11238	4900	56
	2	624	544	540	14	11238	4900	56
	3	624	559	549	12	11238	5028	55
	4	624	559	549	12	11238	5028	55
(3) i i r	1	624	686	685	No	11238	6174	45
	2	624	548	546	13	11238	4932	56
	3	624	558	549	12	11238	5028	55
	4	624	558	549	12	11238	5028	55
(4) h a l	1	750	657	653	13	13497	5917	56
	2	750	654	650	13	13497	5885	56
	3	750	654	662	12	13497	6045	55
	4	750	672	662	12	13497	6045	55
(5) a r f	1	637	556	549	14	11466	5000	56
	2	637	556	549	14	11466	5000	56
	3	637	584	560	12	11466	5256	54
	4	637	584	560	12	11466	5256	54

multipliers (2 at 3.3V) and ALUs (2 at 5.0V), and (4) multipliers (2 at 3.3V) and ALUs (1 at 5.0V). The experimental results for various benchmark circuits are reported in Table 1. The energy estimation includes the energy consumption of the overheads. It is assumed that each resource has equal switching activity ($\alpha_{i,c}$). The results are reported for two supply voltages and for switching = 0.5. It is observed that the energy consumption is increased for higher switching and decreased for lower switching activity. The energy savings for the proposed algorithm is listed along with other multiple voltage scheduling algorithms in Table 2.

Table 2: Savings for various scheduling schemes

Bench- mark ckts	% Average energy savings (S_E)					
	This work	Shiue [8]	Sarrafzadeh [6]	Johnson [3]	Chang [4]	Mohanty [12]
(2)fir	56	-	23	53	-	46
(3)jir	53	-	-	36	-	-
(4)hal	55	24	-	-	36	40
(5)arf	55	12	18	39	29	39

It is observed that using two supply voltage levels an average energy reduction of 55% and average EDP reduction was 13%. If in the critical path there are proportionate number of multipliers and ALUs such that the net performance degradation due to the low frequency operation of multipliers can be overcome by high frequency operation of ALUs then the reduction was significant. With such a scheduler incorporated into a low power datapath synthesis tool will greatly benefit low power processor design especially for compute intensive applications.

- [1] Jose Monteiro, Srinivas Devadas, Pranav Ashar, and Ashutosh Mauskar, "Scheduling techniques to enable power management," in *Design Automation Conference*, 1996, pp. 349–352.
- [2] E. Musoll and J. Cortadella, "Scheduling and resource binding for low power," in *Proc. of 8th International Symposium on System Synthesis*, 1995, pp. 104–109.
- [3] M. Johnson and K. Roy, "Datapath scheduling with multiple supply voltages and level converters," *ACM Trans. on Design Automation of Electronic Systems*, vol. 2, no. 3, pp. 227–248, July 1997.
- [4] J. M. Chang and M. Pedram, "Energy minimization using multiple supply voltages," *IEEE Trans. on VLSI Systems*, vol. 5, no. 4, pp. 436–443, Dec 1997.
- [5] A. Raghunathan and N. K. Jha, "SCALP: An iterative-improvement-based low-power data path synthesis system," *IEEE Trans. on CAD of Integrated Circuits and Systems*, vol. 16, no. 11, pp. 1260–1277, Nov 1997.
- [6] M. Sarrafzadeh and S. Raje, "Scheduling with multiple voltages under resource constraints," in *Proc. of ISCAS'99*, 1999, pp. 350–353.
- [7] A. Kumar and M. Bayoumi, "Multiple voltage-based scheduling methodology for low power in the high level synthesis," in *Proc. of International Symposium on Circuits and Systems, 1999, ISCAS'99*, 1999, pp. 371–379.
- [8] W. T. Shiue and C. Chakrabarti, "Low-power scheduling with resources operating at multiple voltages," *IEEE Trans. on Circuits and Systems-II : Analog and Digital Signal Processing*, vol. 47, no. 6, pp. 536–543, June 2000.
- [9] I. Brynjolfson and Z. Zilic, "Dynamic clock management for low power applications in fpgas," in *Proc. of IEEE Custom Integrated Circuits Conference*, 2000, pp. 139–142.
- [10] T. Burd and R. W. Brodersen, "Energy efficient cmos microprocessor design," in *Proc. of the 28th Hawaii Intl. Conf. on System Sciences*, 1995, pp. 288–297.
- [11] L. Benini, G. De Micheli, A. Liyo, E. Macii, G. Odasso, and M. Poncino, "Automatic synthesis of large telescopic units based on near-minimum timed supersampling," *IEEE Trans. on Computers*, vol. 48, no. 8, pp. 769–779, Aug 1999.
- [12] 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.