Power Virus Generation Using Behavioral Models of Circuits

K. Najeeb, Vishnu Vardhan Reddy Konda,
Siva Kumar Sastry Hari, V. Kamakoti
Reconfigurable Intelligent Systems Engineering Group
Department of Computer Science and Engineering
Indian Institute of Technology Madras, India
kama@cs.iitm.ernet.in

Vivekananda M Vedula
Validation and Test Solutions
Intel Corporation, Austin, U.S.A.
vivekananda.vedula@intel.com

Abstract

The problem of peak power estimation in CMOS circuits is essential for analyzing the reliability and performance of circuits at extreme conditions. The dynamic power dissipated is directly proportional to the switching activity (number of gate outputs that toggles (changes state)) in the circuit. The Power Virus *problem involves finding input vectors that cause maximum dynamic power dissipation (maximum toggles) in circuits. As the power virus problem is* NP-complete *the gate-level techniques are less scalable with increasing design size and produce less optimal vectors. In this paper, an approach for power virus generation using* behavioral models *of digital circuits is presented. The proposed technique converts the given behavioral model* automatically *to an integer (word-level) constraint model and employs an integer constraint solver to generate the required power virus vectors. Experimenting the proposed technique on ISCAS behavioral level benchmark circuits and the standard DLX processor model show that the above technique is fast and yields higher-quality results than the known gate-level techniques. Interestingly, the paper attempts to* generate an assembly program that cause the maximum dynamic power dissipation *on the given DLX processor model. To the best of our knowledge the proposed technique is the first reported that considers power virus generation using behavioral level models.*

Keywords– *Behavioral Models, Dynamic power dissipation, Power virus, Integer Constraint Solvers, Hardware Description Languages (HDL).*

1 Introduction

The high transistor density, together with the growing importance of reliability as a design issue, has made early estimation of worst case power dissipation (peak power estimation) in the design cycle of logic circuits an important problem. The peak power consumption corresponds to the highest switching activity generated in the circuit under the test during one clock cycle.

In this work we made the assumption that the output capacitance for each gate is equal to the number of fanouts. Therefore, the total switching activity is the parameter that needs to be maximized for maximum power dissipation. Accurate estimation of maximum power consumption for a combinational circuit involves finding a pair of input vectors which when applied successively, maximize the number of toggles, among all possible input vector pairs. *These two vectors can be applied one after another in any order repeatedly to cause the estimated power dissipation for an indefinite time*. Given that the circuit has n primary inputs, there are 4^n possible two input vector sequences to be considered for an exhaustive search. It is easy to see that the power virus generation problem for sequential circuits reduces to solving multiple instances of the same problem on its underlying combinational part. Let S_i denote the state of a sequential circuit S at the beginning of the clock cycle i and I_i denote the primary inputs at the clock cycle i. The power virus problem on S is to find a k-cycle sequence $(S_1, I_1, S_2, I_2, \cdots, I_k, S_{k+1})$, such that, $S_{k+1} = S_1$, and at every cycle the input I_i is calculated so as to generate maximum toggle activity in the combinational portion of S. Two different power estimation metrics are defined in the context of sequential circuits, namely, the *peak single-cycle power* and the *peak sustainable power* [10]. The former denotes the maximum power dissipated during any one cycle of the k cycles while the latter denotes the average power measured across the k cycles. The values of I_0 and hence S_1, is got by *warming* up the circuit as discussed in [7].

2 Previous Work

This section presents the work reported in literature for the power virus generation and constraint model based test generation and proceeds to present the salient contributions of this paper.

Several approaches have been proposed to estimate the

IEEE COMPUTER SOCIETY

Authorized licensed use limited to: University of Illinois. Downloaded on September 20, 2009 at 02:24 from IEEE Xplore. Restrictions apply.

maximum power consumption for CMOS circuits in [7, 8, 10, 12, 13].In [8] Devadas *et al.* reduced the problem to a weighted max-satisfiability problem on a set of multi-output Boolean functions obtained from the circuit logic description. This approach took time exponentially proportional to the number of primary inputs (PIs) and hence applicable only to small circuits. Kriplani *et al.* [12] have presented a pattern independent algorithm to find an upper bound on the maximum instantaneous current through the power supply lines of CMOS circuits. An estimation of average switching in combinational circuit using symbolic simulation is discussed in [13]. An upper bound on the number of simultaneous switching gates using partitioning techniques is presented in [15]. On the sequential end, several techniques are proposed in [7, 10]. Of these, the genetic algorithm presented in [10], addresses both single-cycle power and sustainable power metrics. Four different heuristics are presented in [10] and it assumes a **variable-delay** model.

Constraint propagation techniques across different domains, that is, (both arithmetic and boolean domains) have been explored to generate functional tests and high level ATPG vectors on HDL descriptions [6, 9]. The ideas discussed in [3, 14] use constraint solvers to generate tests for functional verification of higher-level architectural features.

3 Contribution of this paper

This paper proposes a methodology for power virus generation using a behavioral description of a circuit. The salient features of the proposed technique are outlined below:

1. *The input to the proposed technique is a behavioral level HDL model*: The behavioral models exploits the word-level parallelism that exist in modern processors to speed-up the computations involved in the generation of the power-virus. *The word-level variables grow less dramatically than the bit-level variables with increasing design functionality*. This accounts for the better scalability of the proposed technique with increasing design size than the gate-level techniques.

2. *Automatic generation of the constraint model from the behavioral model*: The proposed methodology converts the given behavioral description into an A^2M graph based representation [4] and further converts the same into a set of integer constraints. The constraints are solved using an integer solver to generate the required power virus vectors. All the above steps are *fully automated*.

3. *Correlation to Gate level representation*: The power virus computed by the proposed technique shall maximize the toggles on the inputs and outputs of a circuit. The interesting question is that will this maximize the number of toggles in the underlying gate level representation too, as this is the required *real estimate* of the power. The answer to the above question is predominantly positive. This paper discusses the basic intuition behind this positive answer. This is further supported by experimental validation.

4. *Handling sequential designs with dynamic unrolling*: The behavior of a sequential circuit S over k time frames can be modeled as a combinational circuit using the conventional time frame expansion approach, which unrolls the combinational part of S, k times [2].

4 Constraint Generation

The proposed methodology may be divided into the following three phases, namely, (1) Generation of the A^2M graph from the input behavioral model, (2) Generation of constraints from the A^2M graph representation and (3) Solving the constraints using a constraint solver. The following subsections explain the first two phases.

4.1 A^2M graph generation

The synthesizable part of the Verilog code that describes the behavior of a circuit has three major structures, namely, the **A**`lways block`, the **A**`ssign (continuous) statements` and the **M**`odule instantiations`. Hence, the acronym A^2M.The A^2M Graph essentially has 4 entities, namely, `Module`, `Component`, `Signal` and `ADD Node`. Every *module definition* in the Verilog code maps on to a `Module` entity of the A^2M graph. Every *always*, *assign* and *module instantiation* statements in the Verilog code map on to a `Component` entity of the A^2M graph.

The `Module` entity contains the various fields necessary to encapsulate a module definition in Verilog.

The `Component` entity contains the various fields required to encapsulate the three types of structures, namely, a module instantiation, an assign statement and an always block. The various fields inside a `Component` are `name` and `type` of the Component, and the `ADDNode` corresponding to the Component. The `ADDNode` captures the functionality of the `Component` and is defined only for the `Components` that corresponds to an `always` block or an `assign` statement. The ADD Framework is an internal representation of the HDL Description and has been shown to be *complete and efficient*. More details on ADD are available in [5]

4.2 Constraints Generation

This section deals with different Verilog constructs, their corresponding A^2M representations and their equivalent constraint models. This in turn, shall explain the *automatic* generation of the *model constraints* from the given behavioral model. The standard ILOG [11] constraint solver was used

IEEE COMPUTER SOCIETY

Authorized licensed use limited to: University of Illinois. Downloaded on September 20, 2009 at 02:24 from IEEE Xplore. Restrictions apply.

for experimentation. Hence, ILOG-type syntax is used to illustrate the example model constraints in this paper.

4.2.1 The `reg` and `wire` variables

Two types of variables are commonly used in Verilog models,namely, the `reg` and `wire` variables. These variables can either be *bits* or integers (*bit-vectors*).

Any *bit-vector* of size n in the input behavioral model is mapped on to a `Signal` entity in the A^2M graph with $size = n$. This is treated as an integer variable ($IloIntVar$ in ILOG) whose value ranges from 0 to $2^n - 1$. The *bit* variables are also treated as *bit-vector*s of size $n = 1$.

The `reg` and `wire` variables in a Verilog expression map on to one of the two types of nodes on the A^2M graph, namely, the *read nodes* and the *write nodes*. As the names suggest, the variables on the right hand side (left hand side) of a Verilog expression map on to read (write) nodes. The read and write nodes in the A^2M graph have two attributes attached to it, namely, the $range$ and the $index$. The need for these attributes arises from the fact that a vector can be referred in the following three ways in a Verilog code.

- *Bit Select*: For example, $ir[3]$, that selects the third bit of the bit-vector ir. The respective node in the A^2M graph has $index = 3$ and $range = [-1 : -1]$. Since, the integer solver do not directly deal with bits, the corresponding ILOG representation of $ir[3]$ is

  $$2^3 * (IloDiv(ir, 2^3) - 2 * IloDiv(ir, 2^4)),$$

  where, `IloDiv` stands for integer division in ILOG.

- *Part select*: For example, $ir[16 : 13]$, that selects the four bits starting from the bit 13 thru 16 of the bit-vector ir. The respective node in A^2M graph has $range = [16 : 13]$ and $index = -1$. The corresponding ILOG representation is

  $$2^{13} * IloDiv(ir, 2^{13}) - 2^{17} * IloDiv(ir, 2^{17})$$

The variables defined inside a module M carry different values for different instantiations of M. To encapsulate this, every variable in M is declared as an array of size equal to the number of instantiations of M.

4.2.2 Modeling sequential circuits

Every sequential circuit can be represented by the conventional Huffman model [2]. The combinational and the sequential parts are clearly distinguished in this representation. To model the circuit in the integer domain the two basic principles used are (1)Each sequential element is a variable in the integer domain; and, (2)Each combinational element produces a constraint on its inputs and outputs in the integer domain. The behavior of a sequential circuit S over k time frames can be modeled as a combinational circuit using the conventional time frame expansion approach, which unrolls the combinational part of S, k times [2]. The above mentioned approach for modeling in the integer domain is illustrated by using an example of a counter. The following Verilog code models a counter:

```
reg [4:0] counter;
always @(posedge clk)
   counter = counter + 1;
```

Without loss of generality, let there be only one instantiation of the module containing the above Verilog code. Therefore, the variable shall be denoted by $counter[0]$ in the corresponding A^2M graph. To unroll a circuit over time frames, another dimension is added to the variables to represent the time frame. Thus, the variable $counter[0][j]$, denotes the variable $counter$ in the jth time-frame. The combinational part of the Huffman Model is the one that sets constraints on the variable $counter[\,][\,]$ across time-frames. The Verilog code and the underlying A^2M graph imply the following constraint:

$$counter[0][j] == counter[0][j-1] + 1$$

Thus, the clock in Verilog is realized as a time frame in the corresponding constraint model.

4.2.3 Assign Statement

When a circuit is unrolled for MAX_TFS time frames, an assign statement leads to a constraint in each time frame. For example, the assign statement

$$assign\ sum = input_1 + input_2$$

inside the kth instantiation of a module M shall lead to the following set of constraints.

$$\begin{array}{l} for(tf = 0; tf < MAX_TFS; tf{+}{+}) \\ \quad sum[k][tf] == input_1[k][tf] + input_2[k][tf]; \end{array}$$

The other operators that include *bitwise, comparison, concatenation* etc. can be modeled similarly.

4.2.4 Always Statement

The constraint model for an always construct in Verilog depends on the *event* in its *sensitivity list*. The event can be either *clocked* or *non-clocked* and is represented by an event node in the corresponding A^2M graph. The following code is a clocked always structure:

$$\begin{array}{l} always\ @(posedge\ clk) \\ \quad opcode2 <= opcode; \end{array}$$

The A^2M graph for the above case is shown in figure 1. The model constraint for the above A^2M graph for the kth *instantiation* of the module to which it belongs to will be:

Authorized licensed use limited to: University of Illinois. Downloaded on September 20, 2009 at 02:24 from IEEE Xplore. Restrictions apply.

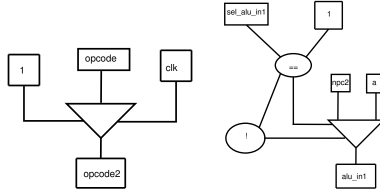

Figure 1. Always statement with and without clk

$$opocde2[k][tf+1] == opocde[k][tf]$$

The following code is a non-clocked always structure.

$always\ @(sel_alu_in1\ or\ npc2\ or\ a)$
$if(sel_alu_in1 == 1)$
$alu_in1 <= npc2;$
$else$
$alu_in1 <= a;$

The A^2M graph for the above case is shown in figure 1. The model constraint for the above A^2M graph for the kth *instantiation* of the module to which it belongs to will be:

$$alu_in1[k][tf] == (sel_alu_in1[k][tf] == 1) * npc2[k][tf] + (sel_alu_in1[k][tf] == 0) * a[k][tf]$$

4.2.5 Module Instantiation

The module instantiations lead to generation of *interface constraints* that establish a connection between the input-output variables of the instantiating module and the instantiated module. Consider the module instantiation:

```
module alu(out,in1, in2)
```

Let the following be its kth instantiation.

```
alu myalu (o1,i1, i2)
```

The following are the interface constraints assuming that the circuit is unrolled for MAX_TFS time frames. Note that the following ensures the connectivity of the variables in every time frame.

$for(j = 0; j < MAX_TFS; j{+}{+})$ {
$alu_out[k][j] == myalu_o1[k][j];$
$alu_in1[k][j] == myalu_i1[k][j];$
$alu_in2[k][j] == myalu_i2[k][j];$
}

4.2.6 Power Virus Constraints

As mentioned earlier, the power dissipated is directly proportional to the number of toggles. It is straightforward to see that every variable in the Verilog model corresponds to a signal in the corresponding A^2M graph. Every signal in the A^2M graph corresponds to an input or output signal in the underlying gate level netlist. These signals also get modeled as variables in the constraint model. Given this, it is easy to infer that by toggling the variables in the constraint model between successive time-frames shall lead to toggling gate outputs in the underlying gate level netlist. For every variable V in all its k instantiations in the constraint model, the following constraints are added.

$for(j = 0; j < MAX_TFS - 1; j{+}{+})$
Max.$\sum(Hamming_Distance(V[k][tf+1], V[k][tf]))$
over all variables V and all instantiations k of V. The `Hamming_Distance` operator captures the notion of toggles between successive time frames (tf) of variable V.

5 Experimental Results

The constraints generated as described in the previous section is input to a constraint solver that *automatically* generates the required vectors. In this section we shall describe details of the experiments conducted for the ISCAS combinational circuits and the 16-bit 5-stage pipelined DLX processor.

5.1 Correlation between Behavioral and Gate level representations

As mentioned earlier, the proposed approach maximizes the number of toggles on the inputs and outputs of the different behavioral level A^2M structures and does not take into account the gate level representation of the same. The ultimate objective is to generate maximum toggles on the underlying gate level netlist. The graph shown in figure 2 plots for different behavioral level constructs, that includes, an 8-bit adder, an 8-to-1 multiplexer and a 4-bit multiplier, the number of toggles in the input and output signals against the total number of toggles in the corresponding gate level implementation. A similar trend is seen for most of the other A^2M constructs. The Magma Blast Fusion tool using the TSMC 0.13 micron standard cell library was used to synthesize the behavioral representations to the corresponding gate level netlists. The graph shows an increase in the total number of toggles in the circuit with increase in the number of toggles in the corresponding inputs and outputs. This suggests that maximizing the toggles on the inputs and outputs for these structures shall in turn maximize the toggle in the underlying gate level netlist. Experimental results shown in the subsequent sections further strengthens our claim.

5.2 Power Virus for Combinational Circuits

As mentioned earlier the power virus generation for combinational circuits involves generating a pair of input vectors

IEEE COMPUTER SOCIETY

Authorized licensed use limited to: University of Illinois. Downloaded on September 20, 2009 at 02:24 from IEEE Xplore. Restrictions apply.

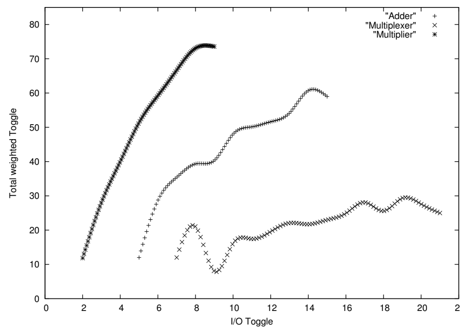

Figure 2. Behavioral Vs Gate Level Correlation

which when applied successively shall cause the maximum toggle in the circuit. Given a n-input circuit the number of such pairs shall be $O(4^n)$ resulting in a huge search space. Even for moderately large circuits the above problem becomes difficult to solve. A simpler version of the above is as follows: Given the first of the two input vectors, generate the other vector such that successive application of both produces maximum toggle in the circuit. This reduces the search space to $O(2^n)$. The proposed approach follows the simplified version.

For generating the power virus the constraint model of the combinational circuit is unrolled for two time frames. The constraint variables corresponding to the input of the behavioral model in the first time frame are initialized with the value of the random vector. Solution to the above constraint model that includes the power constraints yields the second vector that maximizes the toggle. Table 1 presents the results obtained by applying the above technique for the ISCAS High-Level Models [1].

Circuit	# Toggles at beh. level	# Toggles at gate level	# Wires at gate level	# Toggles (Random)
74181	103	123	201	77
74182	20	25	70	17
74283	33	51	104	33

Table 1. Comparison of Behavioral Toggles with Random Gate Level Toggles

From table 1 it is seen that the number of toggles generated at the gate level netlist by applying the vectors output by our method (column 3) is much better than what is generated by applying random pairs (column 5) of vectors. In addition, the total toggle count is closer to the number of wires (upper bound on number of possible toggles) in the circuit (column 4). Column 2 gives the number of toggles generated by the behavioral level representations of the corresponding circuits.

5.3 Power Virus for the DLX processor

The DLX processor is a pipelined sequential circuit. The number of gates in Two types of experiments were carried out on the model. The first experiment was to generate instructions that will maximize the toggle when input to the processor given its initial state S. This enables automatic generation of assembly instruction sequences from any given initial state S that can maximize the toggle count and hence cause *peak single-cycle power* across different cycles. For experimental purpose this initial state S was arrived by applying a sequence of random instructions. Table 2 presents the percentage improvement in toggle count at behavioral level got by application of the instructions generated by the proposed technique in contrast to application of randomly generated instruction sequences of different lengths. It also presents the time and memory requirements which are measured on the HP workstation $xw4200$.

Length of Sequence	% improvement over random	Memory (MB)	Time (sec.)
2	247.89	21	1.2
3	197.51	34	2.1
4	200.35	46	0.9

Table 2. Generating instruction sequences of specified length

A sequence of length 4 *automatically output* by the method assuming an initial state S is as follows:

```
NOP
ADD R1, R0 R2
ADD R1, R0 R2
ADD R1, R16, R13
```

The above was generated by unrolling the circuit for four time frames. An interesting point that is revealed during the experimentation is that an introduction of a `NOP` between two non-`NOP` instructions causes large number of toggles in the system. This is something which may be taken note by the compiler developers who tend to generate NOPs in executable code for various reasons, that include data and control hazard management in modern superscalar and VLIW based architectures. The results in table 2 indicate that our proposed approach is much better than the random one. Note that the memory requirement linearly scales with the length of sequence.

The second experiment is to find a loop of instructions which when applied continuously should maximize the *peak sustained power*. As in the first experiment the processor is driven to an initial state S by applying a random sequence of instructions. From this state S, a loop of instructions were generated that achieves the objective mentioned above. This needs some additional constraints for the following reason.

0-7695-2812-0/07 $20.00 © 2007 IEEE
IEEE COMPUTER SOCIETY

Authorized licensed use limited to: University of Illinois. Downloaded on September 20, 2009 at 02:24 from IEEE Xplore. Restrictions apply.

To ensure same sustained power dissipation during every execution of the loop with say, k instructions, the system should be in the same state at the start of every loop. This implies that for every variable v, $v[i][0] = v[i][k]$. This ensures that at the end of each k cycle the system will be in the same state. Table 3 presents the results of this experiment.

Length of the Loop	Normalized per cycle toggle count	Length of the Loop	Normalized per cycle toggle count
5	50	9	55
6	70	10	49
7	67	11	46
8	62		

Table 3. Loop generation for Maximizing Toggles

From table 3 it is seen that the *per cycle toggle* count is the maximum for loops of length 6 and decreases with subsequent increase of the loop length. The reason for the same may be that the DLX has a five stage pipeline. The following is a code of length 6 output by our method that generates the maximum per cycle toggle.

```
NOP
ADD R1, R0, R2
ADD R1, R0, R2
LD R1, [R2]
NOP
NOP
```

Several interesting points are exhibited by the above sequence. The first one is that the `NOP`s are predominant. The second is that even though the value of register $R1$ is changed by the ADD instructions, the `LD` (LOAD) instruction initializes the register $R1$ back with a value stored in a memory location whose address and content does not change across loops. Given that the above instructions were generated assuming the initial state of the processor to be S, the state of the processor at the end of every execution of the above six instructions should also be equal to S. This also explains why `ST` (STORE) instructions that write into data memory are not generated, as they change the state of the processor.

6 Conclusion and Future Work

This paper presented a technique that used behavioral descriptions to generate power virus test vectors. To the best of our knowledge, this is the first technique reported for the behavioral level power virus generation. Experimenting the proposed technique on ISCAS behavioral benchmark circuits and the 16-bit DLX processor yielded encouraging results. Interestingly, the paper addressed the problem of *generating an assembly program that cause the maximum dynamic power dissipation* on the given DLX processor model. The technique assumed zero delay model. Future work shall involve modifying the technique to account for unit-delay and variable-delay models.

Acknowledgement: This work is supported by the IIT Madras - Intel joint research project on Processor verification.

References

[1] Iscas-85 high level models web site. In *http://www.eecs.umich.edu/ jhayes/iscas.restore/benchmark.html*.

[2] M. Abromovici, M. A. Breuer, and A. D. Friedman. Digital systems testing and testable design. IEEE Press, 2001.

[3] A. Adir, E. Almog, L. Fournier, and E. Marcus. Genesys-pro: Innovations in test program generation for functional processor verification. pages 84–93, April 1994.

[4] K. U. Bhaskar, M. Prasanth, V. Kamakoti, and K. Maneparambil. A framework for automatic assembly program generator (a^2pg) for verification and testing of processor cores. In *In Proc. Asian Test Symp. (ATS)*, pages 40–45, Dec 2005.

[5] V. Chaiyakul and D. D. Gajski. Assignment decision diagram for high-level synthesis. In *Technical Report 92-103 University of California, Irvine, CA*, December 1992.

[6] A. K. Chandra and V. S. Iyengar. Constraint solving for test case generation: a technique for high-level design verification. In *Proc. of Int. Conf. on Computer Design: VLSI in Computers and Processors*, pages 245–248, 1992.

[7] T. Chou and K. Roy. Statistical estimation of sequential circuit activity. In *In Proc. ACM IEEE Intl. Conf. on CAD*, pages 34–37, 1995.

[8] S. Devadas, K. Keutzer, and J. White. Estimation of power dissipation in CMOS combinational circuits using boolean function manipulation. *IEEE Trans. on Computer-aided Design*, 11(3):373–383, March 1992.

[9] F. Fallah, S. Devadas, and K. Keutzer. Functional vector generation for hdl models using linear programming and boolean satisfiability. In *IEEE Trans. on Computer-Aided Design of Integrated Circuits and Systems*, volume 20(8), pages 994–1002, August 2001.

[10] M. S. Hsiao, E. M. Rudnick, and J. H. Patel. Peak power estimation of vlsi circuits: New peak power measures. *IEEE Trans. on VLSI Systems*, 8(4):435–439, Aug 2000.

[11] ILOG. Technology web site. In *http://www.ilog.com*, 2006.

[12] H. Kriplani, F. Najm, P. Yang, and I. Hajj. Pattern independent maximum current estimation in power and ground buses of CMOS vlsi circuits. *IEEE Trans. on CAD*, 14(8):998–1012, Aug 1995.

[13] J. Monteiro, S. Devadas, A.Ghosh, K. Keutzer, and J. White. Estimation of average switching activity in combinational logic circuits using symbolic simulation. *IEEE Trans. on CAD*, 16(1):121–127, Jan. 1978.

[14] J. Yuan, C. Pixley, A. Aziz, and K. Albin. A framework for constrained functional verification. In *ICCAD '03: Proc. of the 2003 IEEE/ACM international conference on Computer-aided design*, page 142, Washington, DC, USA, 2003. IEEE Computer Society.

[15] K. Zhang, T. Shinogi, H. Takase, and T. Hayashi. A method for evaluating upper bound of simultaneous switching gates using circuit partition. In *Asia and South Pacific Design Automation Conf. (ASPDAC)*, page 291, 1999.

0-7695-2812-0/07 $20.00 © 2007 IEEE

Authorized licensed use limited to: University of Illinois. Downloaded on September 20, 2009 at 02:24 from IEEE Xplore. Restrictions apply.