Operand Isolation with Reduced Overhead for Low Power Datapath Design

Conference Paper *in* Proceedings of the IEEE International Conference on VLSI Design · January 2014

DOI: 10.1109/VLSID.2014.90

3 authors, including:

Lokesh Siddhu
Indian Institute of Technology Delhi
11 PUBLICATIONS **33** CITATIONS

2014 27th International Conference on VLSI Design and 2014 13th International Conference on Embedded Systems

Operand Isolation Circuits with Reduced Overhead for Low Power Data-Path Design

Lokesh Siddhu*, Amit Mishra*, and Virendra Singh†
* Indian Institute of Science, Bangalore 560012, INDIA
Email: {siddhulokesh, amitscmishra}@gmail.com
† Indian Institute of Technology Bombay, INDIA
Email: viren@ee.iitb.ac.in

***Abstract*—Dynamic power dissipation due to redundant switching is an important metric in data-path design. This paper focuses on the use of ingenious operand isolation circuits for low power design. Operand isolation attempts to reduce switching by clamping or latching the output of a first level of combinational circuit. This paper presents a novel method using power supply switching wherein both PMOS and NMOS stacks of a circuit are connected to the same power supply. Thus, the output gets clamped or latched to the power supply value with minimal leakage. The proposed circuits make use of only two transistors to clamp the entire Multiple Input Multiple Output (MIMO) block. Also, the latch-based designs have higher drive strength in comparison to the existing methods. Simulation results have shown considerable area reduction in comparison to the existing techniques without increasing timing overhead.**

***Index Terms*—Low Power Design, Operand Isolation, Power Supply Switching.**

I. Introduction

Power dissipation continues to grow as an important challenge in deep sub-micron chip design. Power management is crucial for reliability, packaging, cooling costs of high-performance systems, and battery life of portable devices [1]. Low-power techniques that give significant power savings with low overhead in area and timing are of practical interest.

Today's complex designs have multiple parallel data-paths, so the switching activity at the inputs of idle data-path modules causes redundant computations which are not processed by the downstream circuit; thus increasing power consumption significantly. Operand isolation is a technique for minimizing the energy overhead associated with redundant operations by selectively blocking the propagation of switching activity through the circuits [2]. It basically blocks the toggling of input vector in the downstream combinational circuit. Input vector toggling can be suppressed using several methods discussed later, however they result in area and timing overhead in the design which needs to be optimally reduced. It was first introduced in IBM PowerPC 4xx-based controllers [3]. In their proposed design, the isolation was done by controlling the select lines of the multiplexers. In an ALU, all the execution units (logical units, arithmetic units, shifters and rotators) share (a) common operand bus(es). If operand isolation is not used, all execution units execute concurrently and a single result is selected and propagated. Meanwhile, power has been needlessly expended in those non-selected units. By employing operand isolation, operands are steered only to those execution units which are going to be active in a given cycle [3]. This method can only be used where evaluation has to be performed each clock cycle.

Pre-computation based methods have been proposed by M. Alidina et al. [4]. The primary optimization step is the synthesis of the pre-computation logic, which computes the output values for the subset of input conditions. If output values can be pre-computed, the original circuit can be "turned OFF" in the next clock cycle, hence avoids switching activity. Since the saving in the power dissipation of the original circuit is offset by the power dissipated in the pre-computation phase, the selection of the subset of input conditions for which the output is pre-computed is critical [4]. However this method may duplicate significant pre-existing logic of the modules especially, in case of multipliers and adders. Also pre-computation logic can result in increased clock period.

In [5], "Guarded evaluation" has been proposed as a method to achieve operand isolation. This technique involves identifying sub-circuits (within a larger circuit) whose inputs can be held constant (guarded) at specific times during circuit operation. Sub-circuit identification is done using a property that certain inputs are not observable at specific times during circuit operation. Latches are used to isolate the sub-circuit. However, implementing logic for automatic selection of latches for large designs becomes difficult.

Kapadia et al. [1] proposed a scheme for saving power dissipation in large data-path buses by preventing switching activity in bus driver modules. In this scheme, insertion of extra latches to block transition activity was avoided by utilizing the enable signals of steering modules [registers, multiplexers (MUX), tri-state buffers] as isolation signals. However, this method is unable to provide optimal isolation in multiple fan-out steering modules. Munch et al. [6] proposed adding latches at the inputs of every module and using the activation signal for that module to decide whether to retain the previous input or not. N. Banerjee et al. [2] have proposed power gating-based isolation circuitry. The circuit comprises of power gating transistor connected to the first level of combinational circuit. Upon receiving the power down signal, power gating transistor will remove the power supply connected to the gates and hence resulting in power saving. In order to clamp the output to logical value, a weak keeper transistor is added at every

1063-9667/14 $31.00 © 2014 IEEE
DOI 10.1109/VLSID.2014.90

output of these gates. Power down signal will turn ON the weak keeper transistor and will clamp the output to a specific value. In case of latch-based circuit, a complete latch is added in series at the output of the gates. Latch comprises of two power gated inverters connected back to back. The power down signal activates the latch and they become functional. However, both the circuit implementation techniques consume larger area. Also, we may have some contention issues in case of latch-based circuit when it tries to come out of power down mode. It has been discussed in detail later in this paper.

The following are the contributions of this paper:

- Novel Power Supply switching based isolation circuitry that significantly lowers design overhead compared to existing implementations.
- Applying operand isolation at circuit-level granularity to prevent redundant switching inside the data-path.

The rest of the paper is organised as follows: Section II describes our ideas presented in the paper. It explains the proposed clamp-based and latch-based designs and describes their functionality. Section III presents the experimental results in terms of area and timing delay for clamp-based and latch-based circuits. Section IV is divided in two subsections. First subsection discusses the experimental results and compares and contrasts the data with the existing implementation. The second subsection presents the advantages of our novel design against the existing techniques. Section V mentions several heuristics which can be used along with proposed technique to reduce power. It also presents the criteria to select the input vector in power down mode. Section VI concludes the paper and briefly mentions about the future scope.

II. NOVEL REDUCED-OVERHEAD ISOLATION CIRCUITRY

Operand Isolation is a technique by which output of the first level of combinational circuits is driven to certain specific logical value. It is achieved either by clamping the output to certain value or to latch the output of first level of combinational circuit when it is not getting used. This helps in isolating the rest of the circuit from the redundant switching of the inputs.

It has been observed that operand isolation results in significant power reduction. Simulation results in [2] show that the operand isolation techniques achieve at least 40% reduction in power consumption compared to original circuit with minimal area overhead upto 5% and delay penalty upto 0.15%.

Operand Isolation technique makes use of low power (LP) signal to data-gate the logic block. It is not a timing critical signal as it is asserted/deasserted several cycles before the actual data path operation.

There are three basic categories of isolation cell: those that clamp the output to '0', those that clamp it to '1', and those that latch it to the current value. We propose to use a new method of power supply switching for the design of isolation circuits as explained in further sections.

A. CLAMPING CIRCUITS

The output swing of the combinational circuit is determined by the power supply swing connected to it. The basic principle of power supply switching circuits is that, if the PMOS and NMOS stacks are connected to the same power supply then the output logic value will be same as that of the power supply irrespective of transitions at the input. Therefore, in order to clamp to '1', we need to connect the power supply of the NMOS stack to VDD. And in order to clamp to '0', we need to connect the PMOS stack to ground (GND) to clamp the output to '0'. In both these methods, we are effectively reducing the power supply swing to zero and thus clamping the output to the specific value.

1) Clamp-to-1 Circuit: Figure 1 depicts the clamp-to-1 circuit. In the normal operation, the NMOS stack is connected to GND via transistor Q1. This transistor is similar to the power gate transistor. The difference here is the presence of transistor Q2. In power down mode, the NMOS stack is connected to VDD via transistor Q2. Thus, in the low power mode the total output swing of circuit becomes zero and the output gets clamped to logic '1'.

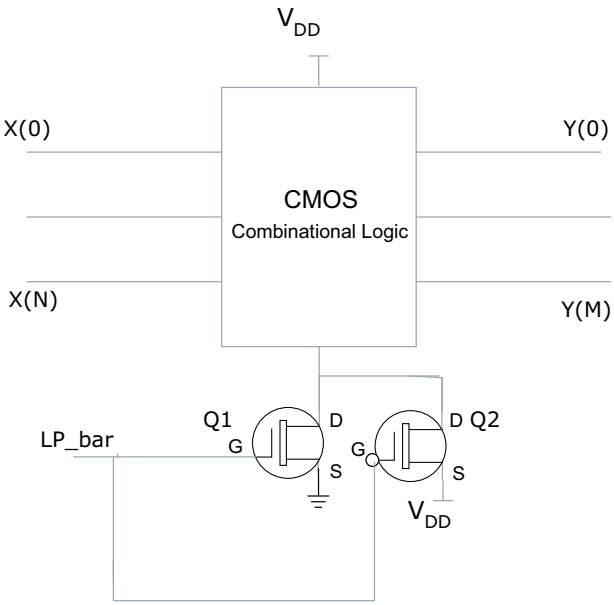

Fig. 1. Clamp-to-1

2) Clamp-to-0: Figure 2 depicts the clamp-to-0 circuit. In the normal operation, the PMOS stack is connected to VDD via transistor Q1. Here, in power down mode, the PMOS stack is connected to GND via transistor Q2. Thus, the total output swing of circuit becomes zero and the output gets clamped to logic '0'.

These circuits have an advantage that the transistors Q1 and Q2 also help in reducing the leakage of the circuit. Also, only one set of transistors can be used for the entire first level of combinational circuit or MIMO structure as shown in Figure 1 and Figure 2.

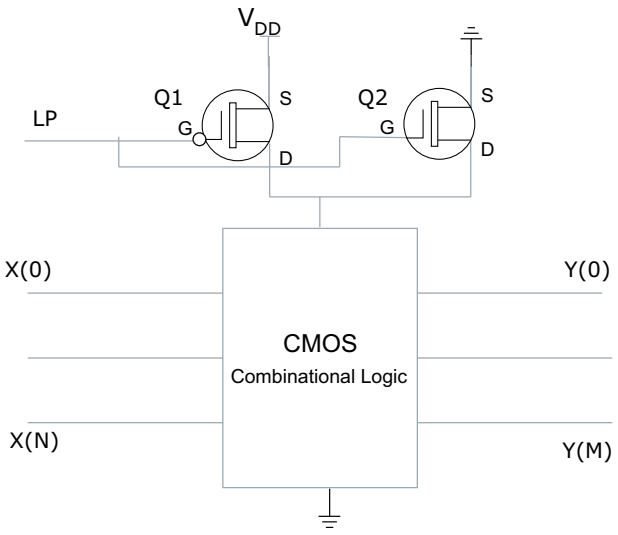

Fig. 2. Clamp-to-0

It is also observed that the overhead area is less when adding logic to the NMOS stack, so it is favourable to clamp to '1'. Also, in MIMO circuit configuration, it is desirable to have large sized isolation transistor (Q1 and Q2) and keep the first level NMOS/PMOS stack size unchanged. This will reduce the area overhead drastically in comparison to upsizing all the first level gates.

B. LATCH CIRCUITS

1) Power Latch: Figure 3 shows the power latch. In order to achieve the latch operation, the power supply connected to the NMOS and PMOS stacks is determined by the current output. Incase the current output is logic '1' then NMOS/PMOS stack power supply is connected to VDD and if the current output is logic '0' then NMOS/PMOS stack power supply is connected to GND. Detailed transistor level operation of the power latch is explained below.

The transistors Q1 and Q2 are connected to the NMOS stack of the logic gate. Q1 and Q2 connects the NMOS stack power supply to GND and VDD respectively. Only one transistor Q1 or Q2 is ON at any time in the circuit operation. Similarly transistor Q3 and Q4 connects the PMOS stack power supply to VDD and GND respectively and only one of them is ON at any time.

Additional control logic for latch operation includes a tri-state inverter (INV), transistors Q5, Q6, and Q7. It helps in connecting the correct power supply to the PMOS and NMOS stacks. The INV and Q5 are only operational in the low power mode. INV is also power gated in the normal mode. Q6 and Q7 are only operational in the normal mode.

In normal operation, *LP* signal is low and *LP_Bar* signal is high. Transistors Q1, Q3, Q6, and Q7 are turned ON while the rest of the additional circuitry is turned OFF. Thus, PMOS stack power gets connected to VDD and NMOS stack power gets connected to GND.

In the low power mode, *LP* signal is high and *LP_Bar* is low. The circuit behaves as a latch and the power supply connected to the NMOS and PMOS stacks is determined by the current output of the gate. The control logic is designed so that the power supply reinforces the output and forms a positive feedback loop.

If current output is '1' and low power mode gets enabled, transistor Q5 is turned ON and Q6 and Q7 are turned OFF. The INV output is logic '0' and this turns ON transistor Q2 and Q3. Thus PMOS and NMOS power supply gets connected to VDD. So the output is always latched to VDD independent of the input transitions.

Similarly, if the current output is '0' and low power mode gets enabled, the INV output is logic '1' and this turns ON transistor Q1 and Q4. Thus PMOS and NMOS power supply gets connected to GND.

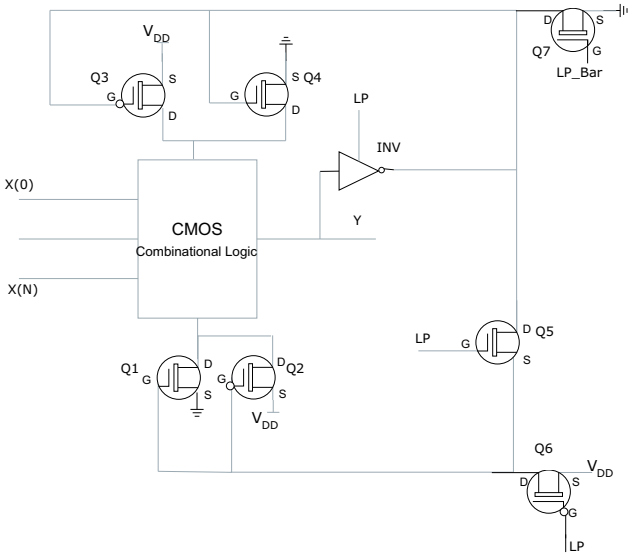

Fig. 3. Power Latch

2) Fused Gate and Latch Logic: Primary motivation of designing this gate is to optimize the area taken by the cell and latch together. In normal latch-based operand isolation, we put the latches after the first stage of the combinational circuit. As shown in Figure 4, we share an inverter between the first stage and the latch to improve the area efficiency. In Figure 4, transistors Q1 and Q2 form a logical AND gate using pass transistor logic. Transistors Q3 and Q4 form a transmission gate. This transmission gate helps in isolating the latch operation from the logic operation. The final output is taken from a static CMOS inverter INV1. Second inverter INV2 and pass transistor Q5 form the loop to latch the output in low power mode.

In normal mode, the transmission gate is turned ON while the pass transistor Q5 is turned off. Thus, INV1 is driven by the logical AND gate which is formed by Q1 and Q2.

In low power mode, the transmission gate is turned OFF

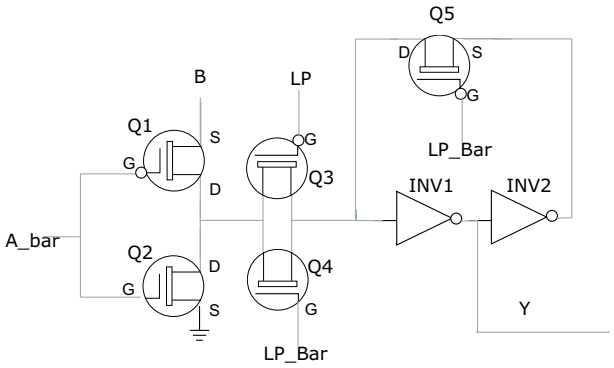

Fig. 4. Fused Gate and Latch Logic

while the pass transistor Q5 is turned ON. Here, the INV1 is driven by the INV2 via Q5, thus it latches the output.

Since this circuit does not use the power stack, the area overhead reduces drastically. Also, we have complementary outputs available because of INV2.

III. RESULTS

The circuits are simulated using 130nm library. NMOS transistor is assumed to have twice the drive strength of PMOS transistor. We have used a constant load of 100fF to characterize the clamp and latch circuits for timing and area overhead. The choice was made such that the output capacitance of the overhead circuit does not affect the timing responses. Only the isolation transistors were sized so as to maintain uniformity in comparison. Delay numbers and area numbers for proposed design and existing implementation are mentioned in the Table I. [1]

DESIGN	Rise Time	Fall Time	Area
CLAMPING CIRCUITS			
3-Input NAND[1]	1.1ns	827ps	1
Clamp-to-1	1.1ns	844ps	1
Clamp-to-0	1.3ns	684ps	1
LATCHING CIRCUITS			
Power Gated NAND and Cascaded Latch (FLH)[1]	1.4ns	898ps	1
Power Latch	1.4ns	798ps	1.25
Fused Latch	1.3ns	918ps	0.75

TABLE I
SIMULATION RESULTS USING 130NM LIBRARY

The 'Area' number shown in Table I takes into account the number of transistors used in the circuit. The area numbers in the Table I are normalised with respect to existing implementation. 3-input NAND gate has total of 6 transistor. Three transistors in parallel form the PMOS stack while 3 transistors in series form the NMOS stack. First Level Hold (FLH) circuit has 12 transistors in total. Four transistors form the NAND gate, and 2 transistors are used to power gate this circuit. Each power gated inverter is comprised of 3 transistors and we need two such inverters to form the latch.

Delay numbers and area numbers are compared against existing implementation [2]. Clamp-based circuits are compared against the 3-input NAND gate for reference while latch-based circuits are compared against the Power Gated NAND and Cascaded Latch (FLH).

- In clamp-to-1 circuit, we do not see any timing degradation in the output rise transition while there is 2% of timing degradation in the output fall transition. The area numbers for the design are same as that of 3-input NAND gate.
- In clamp-to-0 circuit, we see approximately 18% timing degradation in output rise transition while there is similar improvement of about 17% in the output fall transition. The area number in this case also remains same as that of 3-input NAND gate.
- In Power latch circuit, we do not see any timing degradation in the output rise transition while there is improvement of about 11% in the output fall transition. However, this design has an area overhead of 25% against the existing latch-based implementation.
- In Fused latch circuit, we see an improvement of about 7% in the output rise transition while there is 1% timing degradation in output fall transition. However, this design has an better area efficiency of 25% against the existing latch-based implementation

IV. DISCUSSION

A. CLAMP BASED CIRCUITS

Since the NMOS transistor has higher drive capability than the PMOS transistor, it is observed that clamping to '1' is better than clamping to '0' for the same area. However, the simulation was done keeping the same drive strength of both PMOS and NMOS stack. We also observed that the output rise transition because of transition in input A or input B is critical in the NAND gate simulated.

In case of clamp-to-1 we notice better worst case transition than clamp-to-0. In clamp-to-1, the NMOS stack increases whereas, in clamp-to-0, the PMOS stack increases. An increase in the NMOS stack size causes timing degradation in the output fall transitions. Similarly, an increase in the PMOS stack size causes timing degradation in the output rise transitions. In simulation, we choose higher load capacitance in order to eliminate the effect of capacitance added by the NMOS and PMOS stack.

In clamp-to-0, we find that PMOS stack increases while NMOS stack decreases in comparison with 3-input NAND (as we have 1 less transistor in NMOS stack). Thus, we see timing degradation in rise transition while there is improvement in fall transition. While in clamp-to-1, we see that NMOS stack increases while PMOS stack decreases. Though the PMOS stack decreases, but the worst case rise transition is calculated when only one transistor is ON while others are OFF. Thus, we see timing degradation in fall transition while the rise time remains same. In simulation, it has been observed that the

[1]Existing Isolation Methods

timing degradation in the output rise transition due to increase in the PMOS stack was more than the timing degradation in the output fall transition due to increase in the NMOS stack.

B. LATCH BASED CIRCUITS

It has been observed that power latch and FLH [2] perform similarly. The critical transition is still rise in output due to fall transition in either A or B. However, FLH has an inherent disadvantage. Consider a case when we are in low power mode and the latch output is '1'. Second inverter in the latch (INV2) has much higher drive strength as it has to the drive the entire combinational circuit after the first level of gates. Thus, it is bigger in area to avoid the degradation in the timing path. Now when *LP* is removed, the INV2 is still driving logic '1' as there is no tri-state transistor in PMOS stack of INV2, while the logic gate is also driving the output. Now, when the logic gate drives '0', it results in a contention with INV2 leading to short circuit currents flowing between them. Also, it will deteriorate the timing for that fall transition. Eventually, the logic gate predominates and the output logic of INV1 changes to logic '1'. This will cut off INV2.

Additional tri-state transistor can be provided in the PMOS stack of INV2 to avoid these short circuit currents but that will result in increase in area overhead of FLH.

In case of fused latch, critical transition is rise in output due to fall transition in B. The input A is connected to the gate input of the transistor whereas the input B is connected to the drain of the transistor. The input B thus sees two transmission gates in series. Also, transistor Q1 passes weak '0' and this will reduce the input drive to inverter connected at the output (INV1). Hence, fall transition in input B becomes critical and we see the timing degradation in rise in output transition. To eliminate such effect, the design of output inverter is skewed to prefer rise transition.

C. ADVANTAGES

1) Advantage of clamp-based designs: The primary advantage of such configurations over the 3-input NAND based isolation is that only 2 transistors are required to power down the complete combinational block switching. In case of 3-input NAND based isolation circuit, each input gate requires an operand isolation circuitry. Also, the parasitic capacitance of transistor Q2 in our circuit does not add to the output capacitance of the first level combinational logic gate, and both transistors (Q1 and Q2) use the low power signal of same polarity and thus saves routing of power down signal.

In low power mode, the PMOS and NMOS stacks are connected to same power supply, thus the logic gate has no leakage. Circuit leakage minimization during normal operation is also enhanced by the additional isolation circuitry.

2) Advantage of latch-based designs:

- Power Latch: The advantage of power latch is that it inherently has higher drive strength than existing method [2]. In the FLH scheme [2], the logic gate drives the output in normal operation, whereas, the inverter chain drives the output in case of latched operation. So, in case of high load, inverter chain has to be upsized. In the power latch scheme, the logic gate itself drives the output in all cases. So the overhead isolation circuitry can be made smaller. The load output capacitance in case of FLH is a sum of the gate capacitance of the first inverter and the drain capacitance of the second inverter. In case of power latch, load output capacitance is the gate capacitance of the inverter itself. This reduction in the output capacitance helps timing critical path.
- Fused Gate and Latch Logic: The proposed static implementation achieves the functionality of the gate and latch with minimal number of transistors. The area reduction achieved here can be significantly used in the low power design of non-critical paths.

V. HEURISTICS TO SELECT CLAMPING VALUE OF THE CURCUIT

A. LEAKAGE REDUCTION IN OPERAND-ISOLATED MODE

In the gated mode, the functional block does not switch; however, it still dissipates power due to standby leakage which becomes significant in scaled technologies [7]. Leakage of a combinational circuit is a strong function of the state of its inputs [8]. Therefore, by selecting the best input vector for a combinational circuit in standby mode, its leakage power can be significantly reduced.

B. MINIMAL SWITCHING POWER DISSIPATION WHEN CLAMPING

The output of a n-input NAND gate is '1' (for 2^n-1 combination of inputs) and '0' only when all inputs are '1', so NAND gates can be camped to '1' and NOR gates can be clamped to '0'. Similarly for other complex gates heuristics may be used to determine the clamping value so that the power dissipation due to switching to the power down state may be minimized.

C. WEIGHTED APPROACH

In the first method mentioned above, we try to reduce leakage power dissipation in the power down mode whereas in the second method, our attempt is to reduce power dissipation when clamping. So a combination of these two can be used to find an optimal result. For instance, if the unit is to be disabled frequently the weight of the second method should be more whereas if the unit is disabled for long cycles the weight of the first method should be more.

VI. CONCLUSION

In this paper, we presented novel operand isolation circuits for clamping and latching the output of the combinational circuit to a specific logic value. The clamping circuits presented in paper are applied at block level instead of bit level granularity thus, resulting in saving area and increasing

routing efficiency. Proposed Fused-Latch design has shown considerable reduction in the area overhead in comparison with the current implementation. Also, we have been able to reduce the timing overhead in some cases.

In future, we aim to develop an integrated synthesis methodology to automate the application of the proposed operand isolation techniques at the RT-level. Also, we would like to develop some heuristic to select the best input vector for the given combinational circuit.

REFERENCES

[1] H. Kapadia, L. Benini, and G. De Micheli. Reducing switching activity on datapath buses with control-signal gating. *Solid-State Circuits, IEEE Journal of*, 34(3):405 –414, March 1999.

[2] N. Banerjee, A. Raychowdhury, K. Roy, S. Bhunia, and H. Mahmoodi. Novel low-overhead operand isolation techniques for low-power datapath synthesis. *Very Large Scale Integration (VLSI) Systems, IEEE Transactions on*, 14(9):1034 –1039, September 2006.

[3] Jr Anthony Correale. Overview of the power minimization techniques employed in the ibm powerpc 4xx embedded controllers. *ISLPED '95 Proceedings of the 1995 international symposium on Low power design*, page 7580, 1995.

[4] M. Alidina, J. Monteiro, S. Devadas, A. Ghosh, and M. Papaefthymiou. Precomputation-based sequential logic optimization for low power. *Very Large Scale Integration (VLSI) Systems, IEEE Transactions on*, 2(4):426 –436, December 1994.

[5] V. Tiwari, S. Malik, and P. Ashar. Guarded evaluation: pushing power management to logic synthesis/design. *Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions on*, 17(10):1051 –1060, October 1998.

[6] M. Munch, B. Wurth, R. Mehra, J. Sproch, and N. Wehn. Automating rt-level operand isolation to minimize power consumption in datapaths. In *Design, Automation and Test in Europe Conference and Exhibition 2000. Proceedings*, pages 624 –631, 2000.

[7] K. Roy, S. Mukhopadhyay, and H. Mahmoodi-Meimand. Leakage current mechanisms and leakage reduction techniques in deep-submicrometer cmos circuits. *Proceedings of the IEEE*, 91(2):305 – 327, February 2003.

[8] M.C. Johnson, D. Somasekhar, and K. Roy. Models and algorithms for bounds on leakage in cmos circuits. *Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions on*, 18(6):714 –725, June 1999.