INTEGRATION, the VLSI journal 52 (2016) 347–354

Contents lists available at ScienceDirect

INTEGRATION, the VLSI journal

journal homepage: www.elsevier.com/locate/vlsi

Challenges of cell selection algorithms in industrial high performance microprocessor designs

Tiago Reimann [a,*], Cliff C.N. Sze [b], Ricardo Reis [c]

[a] PGMicro - UFRGS, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brazil
[b] IBM Research, Austin, TX, United States
[c] PGMicro/PPGC - UFRGS, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brazil

ARTICLE INFO

Article history:
Received 21 December 2014
Received in revised form
26 August 2015
Accepted 11 September 2015
Available online 25 September 2015

Keywords:
Lagrangian relaxation
Gate sizing
Cell selection
VLSI

ABSTRACT

Timing-constrained power-driven gate sizing has aroused lot of research interest after the recent discrete gate sizing contests organized by International Symposium on Physical Design. Since then, there are plenty of research papers published and new algorithms are proposed based on the contest formulation. However, almost all (new and old) papers in the literature ignore the details of how power-driven gate sizing fits in industrial physical synthesis flows, which limits their practical usage. This paper aims at filling this knowledge gap. We explain our approach to integrate a state-of-the-art Lagrangian Relaxation-based gate sizing into our actual physical synthesis framework, and explain the challenges and issues we observed from the point of view of VLSI design flows.

© 2015 Elsevier Ltd. All rights reserved.

1. Introduction

As frequency scaling slows down when we enter the nanometer era, and most (if not all) high performance microprocessors focus on multi-core architectures, there is no doubt that power dissipation has become one of the most important criteria in the design flow. Power-driven physical design has attracted a lot of attention for decades. Out of all physical design techniques, the problem of gate sizing and threshold voltage assignment[1] is studied extensively because it directly defines the dynamic and leakage power of the transistors.

In the literature, lots of algorithms are proposed for the gate sizing problem. Methods using Lagrangian Relaxation [5,17], linear programming [4,6,12], network-flow [21] are applied to the continuous sizing problem. Also, the discrete cell selection problem is solved by sensitivity-based methods [10,11], constraint relaxation [7], slew targeting [9] and branch-and-bound-based discretization method [19]. The NP-hard property of the discrete problem [14] offers continuum space for new methodologies to be proposed in the seek of narrowing the gap between current approaches quality and the optimal solution for the problem. However, as will be explained in the next section, most of the prior arts in gate sizing and threshold voltage assignment fail to consider their integration into a realistic physical synthesis flow, what is still an open problem for industry.

In industrial design flows, different cell selection algorithms have to be utilized in different stages along the flow [2,3,22], such as logic synthesis, technology mapping, timing-driven placement iterations, electrical correction, critical path optimization, histogram compression, and area/power reduction. Typically, gate-by-gate heuristics are used early in the flow because:

1. Other algorithms such as integer programming, simulated annealing and Lagrangian Relaxation are computationally prohibitive because early optimization iterations are often performed thousands of times.
2. During early stages of physical synthesis flows, it is very common to use simple timing models. Therefore, the advantages in the quality of results provided by those more sophisticated algorithms are usually offset by the error range of the timing model used.
3. Even when one wants to use signoff timer early in the flow, the timing analysis can be highly inaccurate because of (a) slew/capacitance/fanout violations, (b) missing parasitics extraction information, etc.
4. Designers may hide a subset of library cells with particular threshold voltages in order to control the leakage power of the synthesis run until timing has converged to a certain level.

* Corresponding author.
E-mail addresses: tjreimann@inf.ufrgs.br (T. Reimann), csze@us.ibm.com (C.C.N. Sze), reis@inf.ufrgs.br (R. Reis).

[1] Throughout the paper, the term "cell selection" is used to include both gate sizing and threshold voltage assignment. All discussions and algorithms in the paper can be applied for the general library cell selection problem for each gate, which may also include different device width and length, beta ratio, threshold voltage, V_{dd}, and oxide thickness.

http://dx.doi.org/10.1016/j.vlsi.2015.09.001
0167-9260/© 2015 Elsevier Ltd. All rights reserved.

Those cell options may then get exposed during late timing optimization stages (e.g., post-placement or post-routing) so that only the most critical paths can use the low-V_t cell options.

This paper[2] focuses on the power-driven cell selection problem for high frequency microprocessor designs while maintaining the timing quality of result, which is invoked late in the physical synthesis flow after extensive timing optimization. In the literature, [17,18] are the two most relevant works because they discussed practical issues in applying gate sizing, and the experimental results are with real circuits from high performance microprocessor designs.

Reference [18] presents a sensitivity-based method that allows fast incremental and concurrent gate sizing and V_{th} optimization. The method starts with a design meeting the timing requirements and then minimizes the leakage power without creating any timing violation. A graph-based iterative approach is used to look for an optimal set of gates to modify in each iteration.

Reference [17] discussed in depth about the realistic concerns and challenges of the cell selection problem in modern industrial designs. Some of them are: bigger and more sophisticated cell library, cell timing models which no longer can be formulated by convex functions, complex timing constraints, and runtime scalability. Then the paper proposes a Lagrangian relaxation (LR) based method to solve the problem, and deals with the relaxed Lagrangian subproblem by graph models and dynamic programming algorithms. The work uses a linear approximation model calculated based on the cell library to evaluate the cost of changes in delay/slew with respect to load variations when performing critical tree extraction-based DP to solve the *LRS* optimization problem.

However, the paper fails to handle two important issues when the cell selection algorithm is actually used for power reduction in the late physical synthesis stage: (1) it does not provide incremental optimization capability; (2) it does not work with different negative-slack constraints. As discussed in the next section, it is critical for post-timing optimization power-driven gate sizing to handle these two issues before the algorithm can be integrated into real design flows.

Regarding LR-based gate sizing problem, the most well-known paper is probably [5], which proves convergence and optimality for the continuous sizing problem with convex delay models. It was after the publication of [17] and the subsequent ISPD 2013 contest [16] that people pay much more attention to LR-based gate sizing algorithms because the technique is also used by the winning team [8].

The ISPD contest presents a much more realistic infrastructure when compared with usual experimental setups used in previous literature. Several industrial challenges are incorporated into the problem formulation to provide a definition closer to real problems in current industrial design flows and process technology. However, several simplifications present in the contest still compromise the application of those algorithms to real industrial problems.

This paper focuses on these practical challenges of applying LR-based algorithm for power-reduction at the late stage of physical synthesis where timing optimization has converged. The objective is to minimize both the leakage and dynamic power while making sure that timing is not degraded. Formal problem formulation is as follows:

Problem. *Given a placed design, a timing analysis engine, a power calculator, and a set of library cells for each gate minimize the total power dissipation and maintain the timing quality of results.*

Results in this work are evaluated using the same signoff timing analysis engine and power calculator that are available to the cell selection algorithm.

2. Physical synthesis flows and power-driven cell selection

In the literature, most papers addressing discrete gate sizing and threshold voltage assignment never mentioned how to apply the algorithm in a realistic physical design flow. They usually focus on the typical problem formulation, for example, maximization of worst slack, and/or total negative slack (TNS), minimization of total power dissipation, area, and/or electrical violations. These simplified problem formulations are not adequate for industrial design flows as explained later in this section.

Before diving deep into the details of the power-driven cell selection problem, Fig. 1(a) depicts a sample industrial physical design flow, as described in [2] and Chapter 39 of [3]. The box "Optimization" mainly refers to a set of timing optimization steps, but they are usually aware of other constraints such as power, area, routability, and design for manufacturability. The stages before "optimization" aim at finding the best placement of each gate, so algorithms in "optimization" usually do not move gates too far away. Fig. 1(b) shows a newer synthesis flow which pays more attention to routability. Our following discussion does not rely on any particular physical synthesis flow because we only focus on the relationship between "optimization" stage (which means the core timing optimization algorithms) and gate sizing for power minimization.

It is not surprising that multiple iterations of gate sizing and threshold voltage assignment algorithms are applied in steps prior to "optimization". As mentioned in previous section, gate-by-gate heuristic methods are preferable at these stages due to the aforementioned reasons. However, LR-based cell selection algorithms for power reduction have to be invoked later after "optimization" or "post-routing optimization" stages for the following reasons.

- As shown in [17], LR-based cell selection algorithms require signoff timing engine. Therefore, it makes more sense to wait until later after "optimization" when we have accurate timing and power information.
- LR-based gate sizing algorithm using a signoff timing engine usually takes hours for hundreds of thousands of sizable cells in medium-sized industrial designs, which does not fit in the runtime budget of global and timing-driven placement/optimization steps.
- It is true that almost all optimization sub-steps have to be power-aware in a physical synthesis flow nowadays. However, timing optimization has a higher priority earlier in the flow, and normally power-driven (and other design objectives such as area, thermal) optimization algorithms are applied after timing optimization has converged.
- Last but not least, physical synthesis flows are invoked by tool users and designers in different design stages. One example is that iterations of physical synthesis runs are required to evaluate logic designs, floorplans, pin assignments, etc., before timing closure and tapeout. In the middle of design flow iterations, one would rely on power-driven algorithms to compress power dissipation as much as possible after the best-so-far timing is achieved.

[2] The preliminary version of this paper was presented at the 20th Asia and South Pacific Design Automation Conference ASP-DAC in January 2015 [20].

Download English Version:

https://daneshyari.com/en/article/542672

Download Persian Version:

https://daneshyari.com/article/542672

Daneshyari.com