Architectural Synthesis Frameworks on Distributed Register-File Microarchitecture Family

Chia-I Chen and Juinn-Dar Huang
Department of Electronics Engineering and Institute of Electronics
National Chiao Tung University, Hsinchu, Taiwan
cichen.ee94g@nctu.edu.tw, jdhuang@mail.nctu.edu.tw

I. Introduction

As advancing into the deep-submicron (DSM) era, interconnect is becoming one of the most crucial issues for electronic circuit and system designs. The system performance, power, and area are all deeply affected by interconnects, especially for global ones [1]. It is reported that interconnects are responsible for over 50% of the overall dynamic power for a microprocessor [2]. Previous studies have also shown that interconnect is dominating the total area and power in FPGA applications [3].

There are several approaches proposed to deal with the critical issue arisen from long interconnects. Globally-asynchronous locally-synchronous (GALS) designs adopt handshaking protocols for communication over long interconnects. In a synchronous latency-insensitive system (LIS), special pipelining elements, named relay stations, are inserted to break a long interconnect into shorter wire segments for sustaining high operating clock frequency. Furthermore, several types of distributed register (DR) architectures, in which the whole system is divided into several logic clusters, are also broadly studied [4]–[7]. In general, all DR-based architectures try to keep most of interconnects local within a cluster and thus minimize the number of required inter-cluster interconnects for better area and performance result.

The distributed register-file microarchitecture (DRFM) is one of the DR-based architectures and is recently proposed in [4], which takes full advantage of those platforms with a rich set of distributed embedded memory blocks. The overall DRFM architecture and the island architecture are shown in Fig. 1. A DRFM instance is composed of multiple clusters, named *islands* [4]. Each island contains input routing logic, local register file, and functional units (FUs). The local register file is used to store computation results produced by internal FUs. It is also responsible for feeding data operands to internal FUs and external FUs located in other islands. While utilizing DRFM, one should be aware that how to map operations of a target system into islands can have a significant impact on the final outcome in terms of area and performance [4]. Hence, developing an intelligent synthesis algorithm targeting DRFM is important and needs extensive studies further. In this work, we propose a new resource-constrained communication synthesis algorithm for minimizing both inter-island connections and latency targeting on DRFM [5]. More details are given in Section II.

Meanwhile, the inter-island delay is ignored in the original DRFM [4][5]; hence its delay model appears over-simplified. To be a bit more practical, here we propose a new microarchitecture – distributed register-file microarchitecture with inter-island delay (DRFM-IID), which adopts the unit inter-cluster delay model. Nevertheless, this interconnect delay model makes synthesis task more complicated than the one targeting the original DRFM with no inter-island delay. Therefore, we also develop a performance-driven architectural synthesis framework targeting DRFM-IID [6]. More details, including the architecture and dedicated synthesis framework, are given in Section III.

II. Communication Synthesis for Interconnect Minimization Targeting DRFM

In DRFM, the notion of *inter-island connection* (IIC) is presented to better estimate the actual cost of global interconnects. It also demonstrates that the number of IICs after synthesis is highly correlated with the final area size and system performance. Hence the number of IICs can be regarded as an evaluating metric for quality of result (QoR) at early design phases.

In [5], we propose a new resource-constrained resource binding algorithm for both IIC and performance optimization targeting DRFM, as shown in Fig. 2(a). Given a resource constraint (i.e., number of available islands), the proposed algorithm applies an *iterative binding-then-rescheduling* process first, and then invokes an *access conflict removal* procedure. At each control step (cstep), operation nodes scheduled at the current cstep are appropriately assigned to islands first, and then rescheduling is applied to expand the solution space so that a better synthesis output can be produced. The rescheduling and rebinding process also tries to minimize data access conflicts due to limited read ports at the same time. Finally, an access conflict removal procedure is invoked to ensure that no data access conflicts are left at the end of the proposed algorithm.

The experimental results confirm that our algorithm does produce better outcomes with 21% ~ 25% fewer IICs and

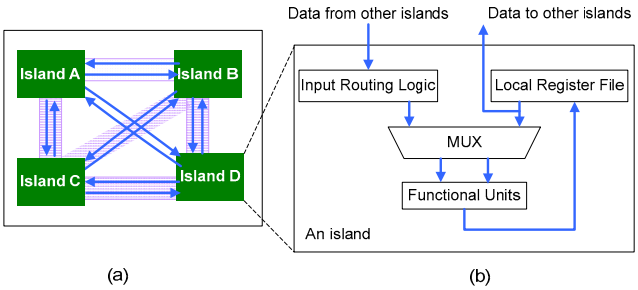

Fig. 1. (a) The DRFM architecture, and (b) the island structure of DRFM.

978-0-7695-4447-2/11 $26.00 © 2011 IEEE
DOI 10.1109/ISVLSI.2011.19

about 12% fewer control steps than the prior art [4].

III. Performance-Driven Architectural Synthesis Flow for DRFM-IID

A. Proposed Architecture

Here we propose a new microarchitecture – distributed register-file microarchitecture with inter-island delay (DRFM-IID). Shown in Fig. 3(a), as one of the DR-based architecture family, DRFM-IID is also composed of multiple islands and the structure of an island is depicted in Fig. 3(b). Within an island, the inputs of the local FU solely come from the local RF, whereas the data from other islands (i.e., inter-island transfers) should be stored in the local RF first and become available only from the next control step (cstep). That is, an inter-island transfer (IIT) takes one whole cstep for data delivery. This interconnect delay model makes synthesis task more complicated than the one targeting the original DRFM [4][5] with no inter-island delay since uncertain interconnect latency is very likely to make a serious impact on whole system performance. Therefore, we also develop a performance-driven architectural synthesis framework targeting DRFM-IID.

B. Proposed Algorithm

The problem formulation of this work is described as: *Given a data flow graph* (*DFG*) *and a resource constraint* (*the number of available islands*), *obtain a scheduled and bound DFG with the minimized latency targeting DRFM-IID*.

As shown in Fig. 2(b), in this framework, list-scheduling is first employed to obtain an initial scheduling result. Then

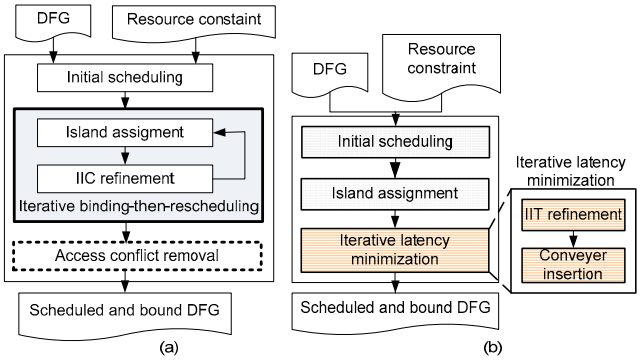

Fig. 2. (a) The overall flow in [5], and (b) the overall flow in [6].

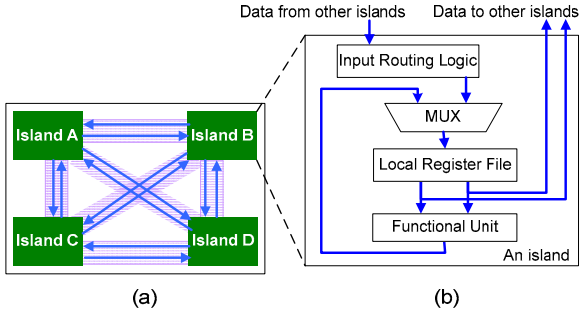

Fig. 3. (a) The DRFM-IID architecture, and (b) the island structure of DRFM-IID.

island assignment and *iterative latency minimization* are applied consecutively to get final solutions. Island assignment is performed to bind operations into islands considering both the number of IITs and the performance criticality of IIT. Then, it tries to iteratively reduce the number of IITs, balance island utilization, and perform necessary conveyer insertion to preserve data dependency.

The experimental results demonstrate that the proposed approach can produce synthesis results with higher performance (27%) and lower power (38%) consumption than the prior art [4] since the number of IITs is commonly regarded as a good metric for on-chip communication power estimation.

IV. Conclusion

In this work, we first develop a communication synthesis framework targeting the original DRFM. The algorithm aims to optimize both interconnect resources (IIC) and system performance (latency). Then we propose a new distributed register-file based platform – DRFM-IID, where the delay model is one step toward reality. Furthermore, a dedicated synthesis framework for DRFM-IID, which focuses on minimizing the system latency and power consumption simultaneously, is also proposed. We thoroughly investigate all existing woks about distributed register-file microarchitecture family with variant inter-island delay models, and the experimental results indicate that our work does provide better synthesis outcome than the prior art.

Acknowledgment

This work was supported in part by the National Science Council of Taiwan under Grant NSC 99-2220-E-009-008.

References

[1] International Technology Roadmap for Semiconductors. Semiconductor Industry Association, 2007.

[2] N. Magen, A. Kolodny, U. Weiser, and N. Shamir, "Interconnect-power dissipation in a microprocessor," *Proc. Int'l Workshop System Level Interconnect Prediction*, 2004, pp. 7–13.

[3] A. Singh, G. Parthasarathy, and M. Marek-Sadowska, "Efficient circuit clustering for area and power reduction in FPGAs," *ACM Trans. Design Automation of Electronics Systems*, vol. 7, no. 4, pp. 643–663, Oct. 2002.

[4] J. Cong, Y. Fan, and J. Xu, "Simultaneous resource binding and interconnection optimization based on a distributed register-file microarchitecture," *ACM Trans. Design Automation Electronics Systems*, vol. 14, no. 3, pp. 1–31, May 2009.

[5] J.-D. Huang, C.-I Chen, Y.-T. Lin, and W.-L. Hsu, "Communication synthesis for interconnect minimization targeting distributed register-file microarchitecture," to appear, *IEICE Trans. on Fundamentals*, vol. E94-A, no. 4, pp. 1151–1155, Apr. 2011.

[6] J.-D. Huang, C.-I Chen, W.-L. Hsu, Y.-T. Lin, and J.-Y. Jou, "Performance-driven architectural synthesis for distributed register-file microarchitecture considering inter-island delay," *Proc. IEEE Int'l Symp. VLSI Design, Automation and Test*, Apr. 2010, pp. 169–172.

[7] C.-I Chen and J.-D. Huang, "A hierarchical criticality-aware architectural synthesis framework for multicycle communication," *IEICE Trans. on Fundamentals*, vol. E93-A, no. 7, pp. 1300–1308, Jul. 2010.