

# 行政院國家科學委員會專題研究計畫 成果報告

## 子計劃一：可重組化運算之系統分析與設計(I)

計畫類別：整合型計畫

計畫編號：NSC91-2215-E-002-043-

執行期間：91年08月01日至92年07月31日

執行單位：國立臺灣大學資訊工程學系暨研究所

計畫主持人：楊佳玲

計畫參與人員：陳依蓉 喻秉鴻 施澤聰

報告類型：精簡報告

處理方式：本計畫可公開查詢

中 華 民 國 92 年 10 月 30 日

# 行政院國家科學委員會補助專題研究計畫成果報告

多媒體通訊系統中可重組化運算計術之研究  
子計畫 一可重組化運算之系統分析與設計

計畫類別：個別型計畫 x 整合型計畫

計畫編號：NSC 91-2215-E-002-043-

執行期間： 91年 8月 1 日至 92年 7 月 31日

計畫主持人：楊佳玲

共同主持人：

計畫參與人員：陳依蓉 喻秉鴻 施澤聰

本成果報告包括以下應繳交之附件：

赴國外出差或研習心得報告一份

赴大陸地區出差或研習心得報告一份

出席國際學術會議心得報告及發表之論文各一份

國際合作研究計畫國外研究報告書一份

執行單位：台灣大學資訓工程學系

中 華 民 國 92 年 10 月 29 日

# 行政院國家科學委員會專題研究計畫成果報告

## 可重組化運算之系統分析與設計

計畫編號：NSC 91-2215-E-002-043-

執行期限：91年8月1日至92年7月31日

主持人：楊佳玲 台灣大學資訊工程系

e-mail: yangc@csie.ntu.edu.tw

http://www.csie.ntu.edu.tw/~yangc

### 一、中英文摘要

關鍵字：可重組結構、系統效能評估、功率消耗、工作排程

可重組結構 (Reconfigurable Architecture) 已逐漸成為發展多媒體通訊系統之重要平台，因其能有效利用應用程式中潛在的運算平行度 (Inherent Parallelism)，且藉由可重組態硬體 (Reconfigurable Hardware) 可重新組態之特色，達到廣泛的可應用性。可重組硬體可於程式執行中被重新組態去執行不同的工作，不會導致系統運算停滯。此動態特色，帶給晶片系統設計者一些新挑戰，包括系統效能評估的方法、工作排程以及功率消耗問題。

此計畫著重於發展高效能、功率的可重組結構之系統層次設計問題。我們將建立一個完整的系統效能與功率消耗評估架構，此架構可讓系統設計者在晶片發展的前期，對於晶片功率消耗與效能進行評估，以達到功率/效能完善折衷 (Trade-Off) 之設計。我們也將對於一些工作排程上的問題進行研討，包括考量耗電量的工作排程及組態內容記憶體 (Context Memory) 的管理。因為多數多媒體與通訊應用程式都是以資料為主，在這個計畫中，我們將利用可重組硬體來改善記憶體系統效能。

#### Abstract

Keywords: reconfigurable architecture, system performance evaluation, power consumption, task scheduling

Reconfigurable architectures are becoming

a viable design alternative for implementing multimedia/communication systems because of the flexibility and ability to exploit inherent parallelisms in applications. A reconfigurable hardware can be programmed to execute different tasks during program execution without causing a computational stall. The dynamic feature of reconfigurable architectures brings new challenges to chip designers, including system performance evaluation methodology, task scheduling and power consumption issues.

The goal of this project is to tackle system level issues in delivering a power-efficient reconfigurable architecture. We will build a complete system performance and power consumption evaluation framework to allow power/performance trade-offs to be examined in the early stage of chip development. We will also study the task scheduling issues, including power-aware task scheduling and context memory management. Since most applications in the multimedia and communication domains are data dominated, we will explore the potential of using reconfigurable logics to make application-specific improvements to memory behavior.

### 二、Introduction & Objective

Traditionally, computing devices are implemented using Application-Specific Integrated Circuits (ASICs) or programmable processors, such as microprocessors or DSPs. ASICs contain circuitry implemented for a

specific task, which has the advantage of low power dissipation and high clock rate, but is lack of flexibility. It cannot adapt to the fast-changing standards quickly [1]. Programmable processors have the advantage of flexibility over ASICs design, however, it might not be able to deliver satisfactory performance if the native processor operations are not well suited to the task. Timing is particularly important for multimedia and communication applications because they usually require real-time performance. The goal of the joint project is to develop a reconfigurable platform considering both the flexibility and performance requirement at the same. The target platform contains a microprocessor, caches and reconfigurable IP cores. This subproject focuses on three objectives listed below:

(1) **Building System Level Performance/Power Evaluation Framework**

The computer architectural development is a repeated process of evaluation of existing design, invention and evaluation of new architectural features through simulation and implementations. Building a simulation framework that allows an architect to explore the design space efficiently before the implementation stage is important to the success of a new architectural design. In this project, we shall develop a system-level simulation framework that integrates a RISC-based instruction-set simulator with IP cores. We shall also build a visualization environment that allows a user to step through the program execution and observe the system performance and energy consumption dynamically and produces graphics representation of simulation results (such as execution time, IPC, and cache miss ratio, etc).

(2) **Task Scheduling**

The second objective of this project is to develop a task scheduler for a dynamically reconfigurable FPGA device [2]. Because of the space constraint on a FPGA device, we

cannot load all the tasks into the device at the same time. The goal of this project is to design a static task scheduling method for simultaneous energy and performance optimization without violating the precedence constraints of scheduled tasks.

(3) **Energy-Efficient Reconfigurable Cache Architecture**

Most applications in the multimedia and communication domains are data dominated. The memory subsystem contributes a significant portion of overall performance and energy consumption for this type of applications. Therefore, an energy-efficient cache architecture is important for delivering an energy-efficient reconfigurable system [5][6][7]. Different applications usually present different memory system behaviors. Even within an application, the optimal architectural parameters are not fixed but time-dependent. In the traditional cache design, we only choose a set of parameters that achieve performance and energy optimization in an average sense. In this project, we propose a cache resource allocation framework that allocates the minimum amount of resources to an application provided that the performance is not degraded and power down the unused cache sections to save energy.

**三、Results Summary**

(1) **System Level Performance/Power Evaluation Framework**

We have modified the PowerAnalyzer simulator [4] to take hardware/software partition as part of architectural configuration. We use MPEG2 decoder to perform our first experiment. IDCT is an important kernel in MPEG2 decoder (contributing to 30% of execution time). Assume that an FPGA hardware can execute IDCT function in  $x$  CPU cycles and the configuration time is  $y$  cycle. Figure 1 shows the speedup of using FPGA to accelerate IDCT with various  $x$  and  $y$  values (ranging from 2 to 1024). We have also added MMX instruction set to the PowerAnalyzer. The performance advantage of using MMX has been well studied. However, the energy advantage of using MMX is not well quantified yet. We have performed a simple test of

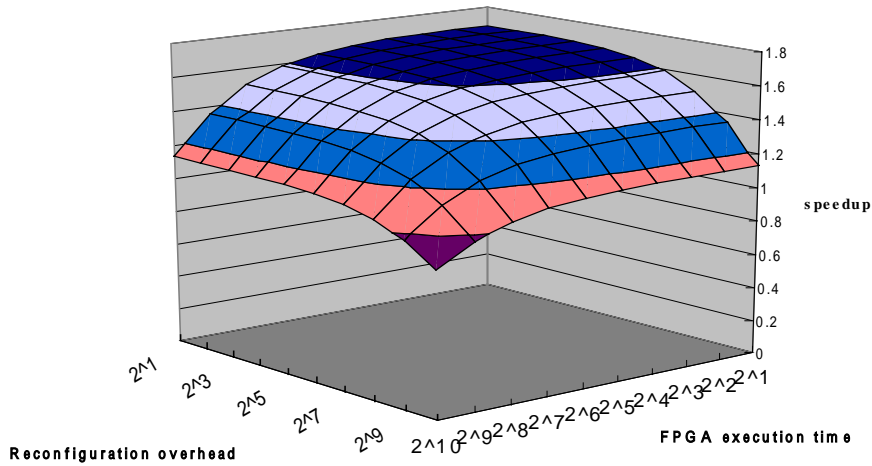


Figure 1: Speedup of Using FPGA to accelerate IDCT

MMX using IDCT. Simulation results show that using MMX reduces the execution time by 42% and energy by 47%. The higher energy reduction percentage than execution time implies that the energy advantage of MMX does not merely come from smaller execution time. We shall look into other factors in details.

**(2) Task Scheduling**

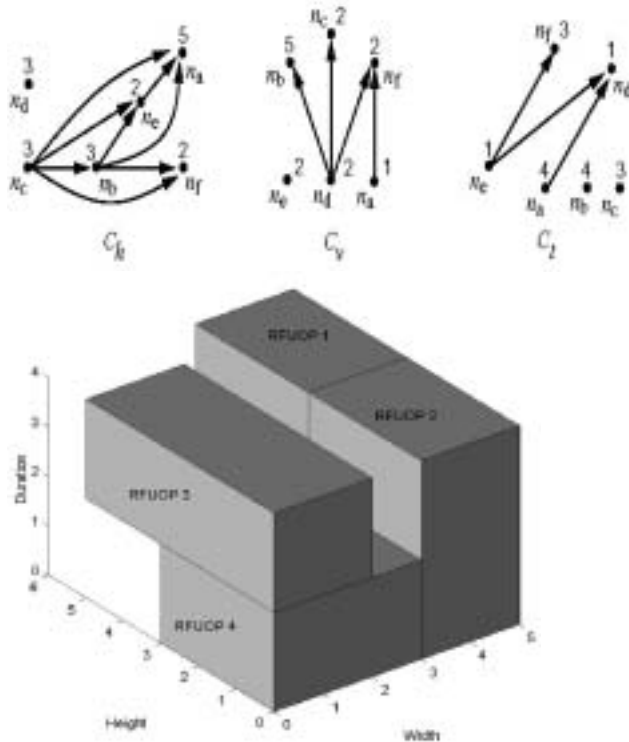


Figure 2:A 3D Placement and corresponding 3D sub-TCG representation

fifth project), we have proposed a 3D-subTCG representation to model the temporal relation between modules. A 3D-sub TCG contains three graphics  $C_v$ ,  $C_h$  and  $C_t$ . Similar to TCG,  $C_v$  ( $C_h$ ) represents vertical (horizontal) relation between modules. 3D-subTCG uses an additional graph,  $C_t$ , to represent the temporal relation. Figure 2 shows a temporal placement and the corresponding 3D-subTCG representations. We have also derived a set of feasibility detection formulas for temporal constraints. We have implemented a SA (Simulated Annealing) method to solve the temporal floorplanning problem (optimizing both the area and execution time). The results have been published in ASP-DAC 2003[8].

**(3) Energy-Efficient Reconfigurable Cache Architecture**

We have first evaluated the cache size requirement for a set of multimedia and communication applications. We found that the working set size for the tested suite ranges from 2K to 58K. We have also observed that an individual application experiences phase change during execution. Figure 3 shows how the working set size changes for JPEG during program execution. Note that we define the working set size as the smallest cache size (from 2K to 64K) achieving less than 1% miss rate. These experimental results indicate that the cache size requirement varies among/within applications. The previously proposed working set detection algorithms fails to

In cooperation with 張耀文教授(PI of the

detect phase changes for gsm and tend to overestimate the required cache size. We have developed a loop-based algorithm that determine the reconfiguration point based on profiling information. We are currently evaluating the accuracy of the proposed algorithm on a set of multimedia applications.

#### 四、Project Evaluation & Conclusion

In this project, we target at the system-level design for a reconfigurable multimedia system. We focus on three issues: system-level perform/power evaluation framework, task scheduling on FPGAs and reconfigurable memory hierarchy. Since this project is originally proposed for a three-year period, all of the research works mentioned above are still in progress. For the first objective, building a system level simulator, we have successfully modified the PowerAnalyzer to take hardware/software partition as part of architectural configuration. We are currently building a simulation framework implemented in SystemC that integrates instruction set simulator with other IP models. For the second objective, task scheduling on FPGA devices, we have derived a 3D placement algorithm based on a topological representation. The results have been accepted to publish in ASP-DAC 2004. The on-going work is to develop a 3D placement algorithm based on binary tree representation. We will also consider the fixed outline constraint. For the third objective, building a reconfigurable cache architecture, we are preparing a paper on the proposed loop-based I-cache size estimation algorithm for DAC 2004. The follow-up work is to build a cache resource allocation mechanism supporting multi-programming system.

#### 五、Acknowledge

Students who join this project are 陳依蓉, 喻秉鴻 and 施澤聰.

#### 六、Bibliography

[1] R. Tessier and W. Burlison. Reconfigurable Computing for Digital Signal Processing: A Survey. *Journal of VLSI Signal Processing* 20, 7-27, 2001

- [2] Xilinx Inc. XC6200 Field Programming Gate Arrays, Data Sheets. Available : <http://www.xilinx.com/>
- [3] D. Brooks, V. Tiwari, and M. Martonosi. Wattch: A Framework for Architectural-Level Power Analysis and optimizations. In *Proceedings of the 27th International Symposium on Computer Architecture (ISCA)*, Vancouver, British Columbia, June 2000.
- [4] <http://www.simplescalar.com/v4test.html>
- [5] D. H. Albonesi. Selective Cache Ways: On-Demand Cache Resource Allocation. In *Proceedings of the 32<sup>nd</sup> Annual IEEE/ACM International Symposium on Microarchitecture (MICRO 32)*, pages 248-259, Nov. 1999.
- [6] Se-Hyun Yang, Michael D. Powell, Babak Falsafi, Kaushik Roy, and T. N. Vijaykumar. An Integrated Circuit/Architecture Approach to Reducing Leakage in Deep-Submicron High-Performance I-Caches. In *Proceedings of the 7<sup>th</sup> International Symposium on High-Performance Computer Architecture (HPCA)*, Jan. 2001.
- [7] Se-Hyun Yang, Michael D. Powell, Babak Falsafi, and T. N. Vijaykumar. Exploiting Choice in Resizable Cache Design to Optimize Deep-Submicron Processor Energy-Delay. In *Proceedings of the 8<sup>th</sup> International Symposium on High-Performance Computer Architecture*, Feb. 2002.
- [8] Ping-Hung Yuh, Chia-Lin Yang and Yao-Wen Chang. Temporal Floorplanning Using 3D-SubTCG. In *Proceedings of the IEEE ASP-DAC*, January, Japan