

# Run-time Mapping Techniques for NoC-based Heterogeneous MPSoC Platforms

Amit Kumar Singh

School of Computer Engineering, Nanyang Technological University, Singapore  
amit0011@ntu.edu.sg; <http://www3.ntu.edu.sg/home2007/AMIT0011/>

**Advisor:** Thambipillai Srikanthan

**Expected PhD Defense:** December 2012

The reliance on Multi-Processor Systems-on-Chip (MP-SoCs) to satisfy the high performance requirement of complex embedded software applications is increasing. The Networks-on-Chip (NoCs) based interconnection infrastructure is fast becoming a preferred approach to facilitate communication among the processing elements (PEs) of MPSoCs. The heterogeneity of MPSoCs is also increasing by employing different types of PEs such as general purpose processor (GPP) for flexibility and custom accelerators for compute intensive tasks in order to meet the functional and non-functional requirements. However, application development and run-time management for different type of PEs is difficult as compared to their homogeneous counterpart.

Run-time mapping techniques need to load the application tasks into the system at run-time as and when the application needs to be supported. In this thesis, a number of efficient techniques have been proposed to realize run-time mapping algorithms for heterogeneous MPSoC platforms. The proposed techniques take computation and communication costs into account in order to improve the efficiency of run-time mapping. These techniques are well suited for on-the-fly assignment of new tasks into the system shown as *On-the-fly Mapping* in Fig. 1. Novel design-time analysis of applications has also been proposed to further improve the efficiency of run-time mapping by using the analysis results shown as *Mapping using DSE results* in Fig. 1.

MPSoC with single-task supported PEs are considered first. A new packing strategy to map the various tasks of an application in close proximity has been proposed to reduce the communication overhead [1]. The proposed strategy was further extended to devise a time-bounded method to minimize the overall computation time of the mapping process. This provides for a systematic approach to identify the best mapping configuration within a pre-assigned time constraint. Performance evaluations based on 20 random applications show that the proposed techniques outperform the existing techniques by up to 22%.

Subsequently, the proposed mapping process was extended to support an MPSoC platform in which each PE is capable of supporting multiple tasks [2]. The extended techniques facilitate in the mapping of a group of communicating tasks on the same PE, thereby resulting in a further reduction in

the communication overhead. This approach has shown to further reduce the time required to identify the best mapping configuration within a pre-assigned time constraint.

The proposed techniques do not consider communication between tasks during mapping. In order to further improve the performance, new mapping techniques have been proposed that map the tasks of applications in communication-aware manner [3] [4]. The communicating tasks are mapped on the same PE as far as possible in order to reduce the communication overhead. We show that the proposed techniques are capable of alleviating Network-on-Chip (NoC) congestion bottlenecks and thus optimize the overall performance. On an average, channel load and total energy consumption is reduced by 30% and 46% respectively when evaluated with varying number of applications.

The run-time mapping techniques were further enhanced by taking into account both the computation and communication costs so as to optimize the overall computation efficiency. The techniques first target homogeneous MPSoCs [5] and then heterogeneous MPSoCs [6]. They rely on the systematic elimination of the longest communication path at a time until the computation load on any PE impedes on the overall performance. The proposed techniques were tested using multiple scenarios of an MPEG-4 application to demonstrate that the total execution time and energy consumption can be reduced by 33% and 39% respectively when compared to an approach that rely solely on the communication-aware strategy.

A hybrid strategy has also been proposed to further accelerate the run-time mapping process when the applications to be supported on a platform are known at design-time. The hybrid strategy follows *Mapping using DSE results* path in Fig. 1. It relies on the design-time analysis to generate light-weight run-time mapping heuristics, which aid the communication and computation aware run-time mapping process. The overall computation complexity of the design-time analysis has been reduced notably. The strategy initially targets homogeneous MPSoCs [7] and later on has been extended for heterogeneous MPSoCs employing pruning-based [8] and communication-aware [9] analysis. Experiments based on models of real-life multimedia applications show that the proposed analysis strategy is faster by 83% and run-time mapping is accelerated by 93% when compared to state-of-the-art analysis and on the fly mapping approaches, respectively.

Further, techniques to design MPSoCs of given dimension such as  $4 \times 4$  grid of cores are presented for hardware platforms

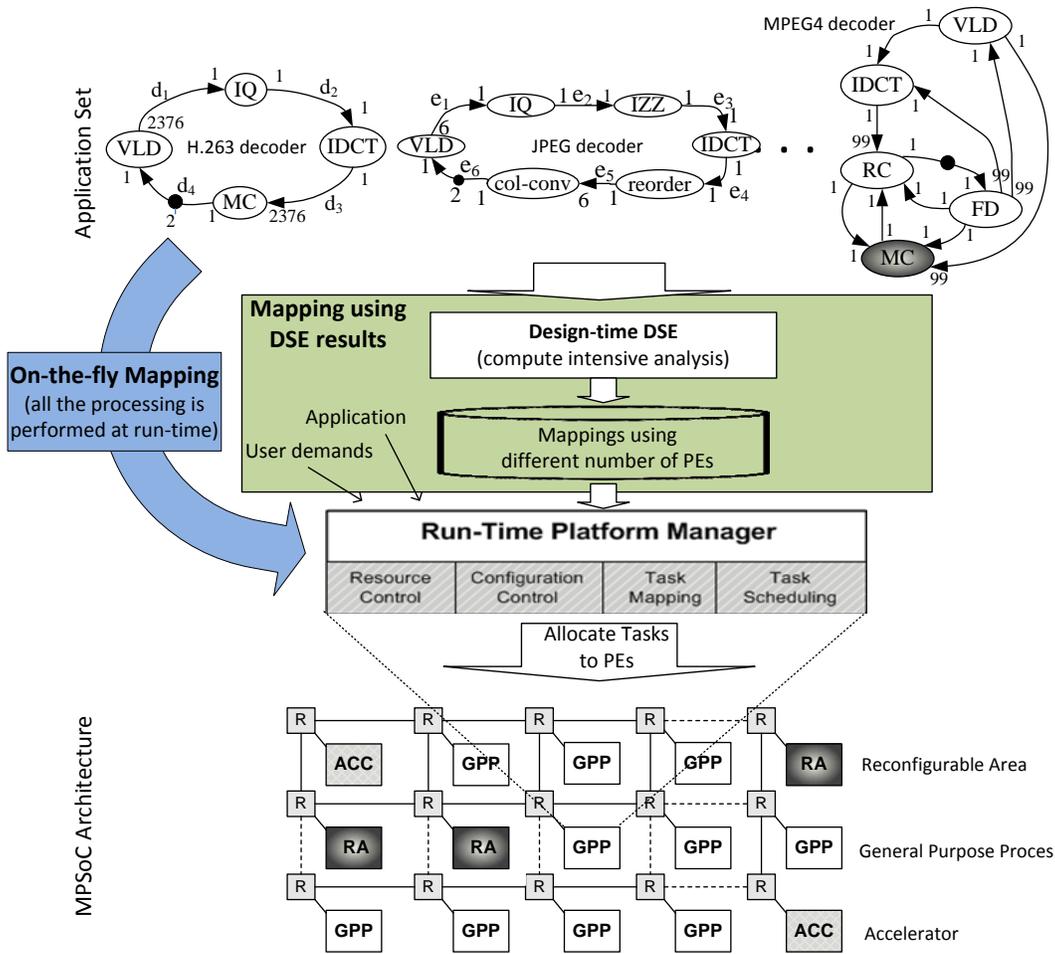


Fig. 1. On-the-fly and Hybrid Mapping Flow.

on FPGAs [10]. Results from the run-time mapping techniques are used to map real-life applications on the hardware MPSoC platforms in order to validate accuracy of the techniques.

Finally, the proposed run-time mapping process relies on an efficient computation and communication aware strategy, which is complemented by light-weight heuristics to implement embedded computing applications that demand high performance.

## REFERENCES

- [1] A. K. Singh, W. Jigang, A. Prakash, and T. Srikanthan, "Efficient heuristics for minimizing communication overhead in noc-based heterogeneous mpsoC platforms," in *IEEE/IFIP International Symposium on Rapid System Prototyping*, 2009, pp. 55–60.
- [2] —, "Mapping algorithms for noc-based heterogeneous mpsoC platforms," in *Proceedings of the 2009 12th Euromicro Conference on Digital System Design, Architectures, Methods and Tools*, ser. DSD '09, 2009, pp. 133–140.
- [3] A. K. Singh, T. Srikanthan, A. Kumar, and W. Jigang, "Communication-aware heuristics for run-time task mapping on noc-based mpsoC platforms," *J. Syst. Archit.*, vol. 56, pp. 242–255, July 2010.
- [4] A. K. Singh, W. Jigang, A. Kumar, and T. Srikanthan, "Run-time mapping of multiple communicating tasks on mpsoC platforms," *Procedia Computer Science*, vol. 1, no. 1, pp. 1019 – 1026, 2010.
- [5] S. Kaushik, A. K. Singh, and T. Srikanthan, "Computation and communication aware run-time mapping for noc-based mpsoC platforms," in *IEEE International SOC Conference*, 2011, pp. 185–190.
- [6] S. Kaushik, A. K. Singh, W. Jigang, and T. Srikanthan, "Run-time computation and communication aware mapping heuristic for noc-based heterogeneous mpsoC platforms," in *IEEE International Symposium on Parallel Architectures, Algorithms and Programming*, 2011, pp. 203–207.
- [7] A. K. Singh, A. Kumar, and T. Srikanthan, "A hybrid strategy for mapping multiple throughput-constrained applications on mpsoCs," in *International Conference on Compilers, Architectures and Synthesis of Embedded Systems*, 2011, pp. 175–184, (Enclosed paper).
- [8] —, "Accelerating throughput-aware run-time mapping for heterogeneous mpsoCs," *ACM Transactions on Design Automation of Electronic Systems*. (Accepted for Publication), 2012.
- [9] A. K. Singh, A. Kumar, W. Jigang, and T. Srikanthan, "Communication-aware design space exploration for efficient run-time mpsoC management," in *IEEE International Symposium on Parallel Architectures, Algorithms and Programming*, 2011, pp. 72–76.
- [10] A. Singh, A. Kumar, T. Srikanthan, and Y. Ha, "Mapping real-life applications on run-time reconfigurable noc-based mpsoC on fpga," in *International Conference on Field-Programmable Technology*, dec. 2010, pp. 365–368.