Chapter 7
Dynamic Reconfigurable NoCs: Characteristics and Performance Issues

Vincenzo Rana
Politecnico di Milano, Italy

Marco Domenico Santambrogio
Politecnico di Milano, Italy

Simone Corbetta
Politecnico di Milano, Italy

ABSTRACT

The aim of this chapter is the definition of the main issues that arise when dealing with the design of a NoC-based reconfigurable system. In particular, after the definition of the target architecture, several factors, requirements and constraints that have to be taken into account during the design of reconfigurable NoCs will be described and analyzed. The second part of this chapter will focus on the main issues in dynamic reconfigurable NoCs design, such as the definition of a layered approach, of a packet-switched communication infrastructure, of a proper routing mechanism and of a communication protocol support. Finally, the last part of this chapter will deal with the description of the most relevant implementation details, such as the placement of the bus-macros, the design of the network switches and the physical implementation of the routing mechanism.

INTRODUCTION

In traditional System-on-Chip design it is possible to know in advance the actual communication requirements, the application needs and all the components (modules) needed to realize the desired architecture. They can be understood a priori, at synthesis-time by the analysis of the application specification. For this reason, once all the components of the system have been defined, they will remain unchanged throughout the life cycle of the system. In dynamically changing environments, on the contrary, the design factors are likely to change, in that the dynamic nature of the target system does not allow to be fully understood at design time. This chapter describes and analyzes several different issues that arise when designing a reconfigurable NoC. In this particular scenario it is not possible to utilize a generic NoC infrastructure on a generic reconfigurable system, since the interactions between the issues of the two dif-
different scenarios creates a new set of difficulties that need to be explicitly faced by the designer.

The main relevant design factors in NoC design are: latency, parallelism, resource usage. These factors are, in general, subject to a design trade-off, since independent optimization does not guarantee system-wide performance. Latency corresponds to the amount of time required for an output to be generated, given the inputs to that particular functional unit performing the computation. This value relies upon the complexity of the circuitry of the component, and on physical characteristics of the interconnects. The degree of parallelism captures the simultaneous communications that are performed between pairs of modules. This can be used to express and (qualitatively) define the communication performance. The resource usage is a hard constraint, since the bounded physical resources within the device have not to be exceeded. Cost-driven designs take care of the silicon cost when realizing SoC, and employing as few resources as possible reduces the design costs. Reconfiguration capabilities, however, can be used to (partially) solve this problem: the same chip area can be used in time division fashion, by allocating different modules in different time instants on the same area. A flexible communication infrastructure ensures several modules to be connected in different ways, and it guarantees the required level of connectivity. In this way, fault-tolerance could be achieved, resulting in a higher availability of the system.

Dynamic reconfiguration at the communication level increases the adaptability level to the communication infrastructure resulting in the system to be able to scale the communication as required. Last, but not least, by exploiting reusability it is possible to decrease the (mean) time required to design the desired target system. Within this new context the designers have to face new issues, ad previously hinted. Among this new set of issues, the most relevant ones concern the definition of a novel approach to implement a reliable routing mechanism and reliable communication channels, both during and after a reconfiguration process. When the NoC has to be configured at run-time, it is necessary to extend the routing protocol with a support that makes it possible to dynamically change the routing of the packets accordingly to the new topology of the underlying communication infrastructure. This means that also the number of switches that a single packet has to cross is not known at design-time, and also this issue has to be solved within the communication protocol.

From the communication channels point of view, it is necessary to find a way in which it is possible to connect different modules that can be placed in different reconfigurable regions and that can be characterized by a different implementation. In order to solve this problem it is necessary to define a shared interface and to employ particular components, the bus-macros, which can be placed on the boundaries between two reconfigurable regions, creating a sort of bridge between them that will not be affected by the reconfiguration processes. This makes it possible to anchor the signals of a reconfigurable module on predefined positions, which will not be changed during the evolution of the system, since bus-macros are never involved in reconfiguration processes. The placement of these bus-macros affects the definition of the reconfigurable architecture and has to be faced by taking into account several factors that influence the overall performance of the whole system.

BACKGROUND

Target Reconfigurable Architecture

The target reconfigurable architecture is based on the model presented in (Ferrandi, F.; Santambrogio, M.D.; Sciuto, D. (2005)), consisting of a static part and a reconfigurable one, that can be further split into several reconfigurable regions. The chip (FPGA device) is divided into two (or more) regions, the first, the static part, containing