Design and Performance Evaluation of a Reconfigurable Delta MIN for MPSOC

Yassine AYDI
CES- National School
of engineering of Sfax , Tunisia
LIFL and INRIA-Futurs, University
of Lille, France
Email: yassine.aydi@oous.rnu.tn

Samy MEFTALI
and Jean-Luc DEKEYSER
LIFL and INRIA-Futurs,
University of Lille, France
Email: samy.meftali@lifl.fr
dekeyser@lifl.fr

Mohamed ABID
CES- National School
of engineering of Sfax , Tunisia
Email: mohamed.abid@enis.rnu.tn

Abstract—**Multiprocessor system on chip is a concept that aims at integrating multiple hardware and software in a chip. Multistage interconnection network has been frequently proposed as connection means in classical multiprocessor systems. They are generally accepted concepts in the semiconductor industry for solving the problems related to an on-chip communication. This paper presents the design of reconfigurable Delta MINs in which the connections change dynamically at run time. Using SystemC timed simulations, performance evaluation of a Delta MINs are given and analyzed.**

Index Terms—**MPSOC, NOC, Delta MIN, Network design, Reconfigurable MIN.**

I. Introduction

MPSOC platforms that will be used in future generation will satisfy many critical requirements: they will be energy efficient, cheap, reliable and offer sufficient computing power for advanced and complexed applications. To satisfy all these requirements simultaneously, future MPSoCs will integrate various types of processors and data memory units, resulting in very heterogeneous platforms.

Intercommunication requirements of future MPSoCs will not be feasible using a single shared bus or a hierarchy of buses due to their poor scalability with chip size and their shared bandwidth between all the components. To overcome this problem, Network on Chip (NoC) has been proposed by academia and industry as a solution for the on-chip communication challenges for the future Multi processor system on chip [1]. NOCs can be implemented as reusable Intellectual Property (IP) which can be quickly reconfigured [2]. Also, they can use Globally Asynchronous locally Synchronous (GALS) paradigm for eliminating high wires delays and high clock skew [3].

Multistage Interconnection Networks (MINs) are used in multi processors systems. As an example, MINs are frequently used to connect the nodes of IBMSP [4] and CRAY Y-MP series [5]. Further on, MINs are applied for Network on Chip to connect processors to memory modules on MPSOC [6].

Many variations of MINs have been introduced. These architectures provide a maximum bandwidth to components (processors, DSP, IP..), and minimum delay access to memory modules. A MIN is defined by, its topology, switching strategy, routing algorithm, scheduling mechanism [7], and fault tolerance [8].

The communication platform of the MPSOCs architectures can be implemented with Multistage Interconnection Networks which must be reconfigured for different purposes.

Performance evaluation is a key step in any design, allowing for decisions and trade-offs in view of system optimization. It is determined by modeling, using simulation [9], formal methods [10] or mathematical methods [11].

A simulation of a model is good opportunities to evaluate the performance of a MIN. Its advantages are a more detailed network description and a shorter time development. A design and performance analysis of a reconfigurable Delta MINs for MPSOC Architectures is investigated in this paper. In section 2, the architecture of a multistage interconnection is introduced. Next, a design of a generic reconfigurable MIN and its components is presented. Section 3 gives a description for our evaluation methodology. Finally, simulation results and corresponding analyses are detailed.

II. Preliminaries

In this section we present the MIN architecture that is used to design the interconnection platform dedicated for MPSOC.

We proposed in figure 1 a topological classification of MINs. They can be defined as a network used to interconnect a group of N inputs to a group of M outputs using several stages of switches elements led by linking stages [12].

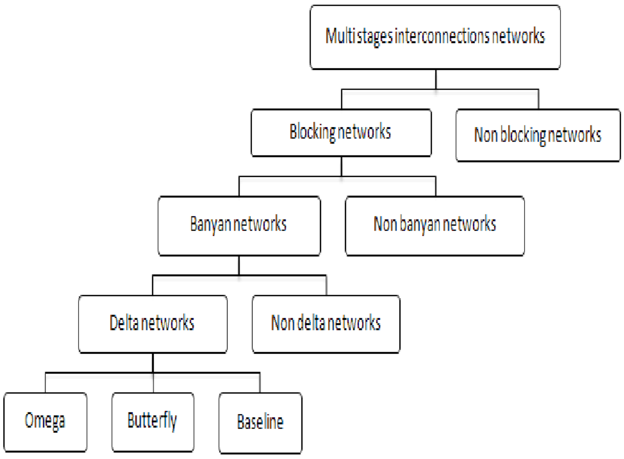

Fig.1: Classification of MINs

A. MINs with Banyan property

A banyan MIN is a multistage interconnection network having the property of the existence of one and only one path between each source and destination.

For banyan MIN of size NxN consists of CxC crossbars. Suppose that the switch input and output are presented to the base C (d_0 , d_1 , …..d_{c-1}) , and they have the same indexes, then digits d_0 of all inputs of a switch must be equal. A network having this characteristic is called to be having the Delta property [13].

B. MIN Architecture

The basic building blocks of the MINs are switches elements, connected by links. The multistage networks having N inputs and N outputs nodes and using C×C switches have N/C number of switches at each stage. The number of nodes for K stages is N*K/2. Each stage is associated with a $\log_2 N$ bit vector called "stage-mask". A path between source and target is obtained by operating each corresponding switch in stage i in straight mode if the i bits of the source and the target are equal, otherwise in exchange mode.

III. Proposed Model

In this section, we present the MINs architectures designed for Multiprocessor System on Chip platform. The model of simulation is fully implemented in SystemC. It is preferred over a HDL language because of SystemC's facility to model and simulate a large quantity of cycles and it has a high level descriptions as well as cycle-accurate precision. We define down the assumptions and the basic components under which our model is implemented and analyzed.

A. Network topology

The topology plays an important role in designing routing strategy, network latency, throughput and area.

We will restrict study to Delta MINs networks (figure 2). That is the MIN has N inputs and N outputs nodes using a switches elements which have C inputs and C outputs ports. There exist various popular MINs. The difference between each of these networks is the topology of interconnection links between the crossbar stages. A study of equivalence of various types of Delta MINs has been studied in [14].

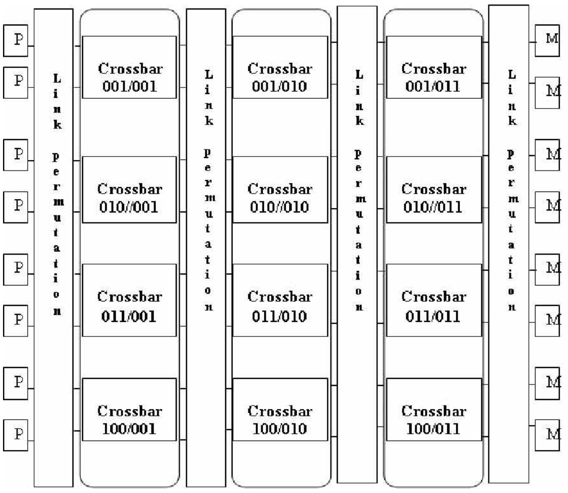

Fig.2: a generic model of a Delta MIN

We thus present the link permutation used in the multistage interconnection network using a 2x2 crossbar elements:

- The perfect shuffle denoted $\sigma n\text{-}1$ is a bit-shuffling permutation where: $\sigma_{n-1}(x_{n-1}\ x_{n-2} \ldots x_1\, x_0) = x_{n-2}\ \ x_1\, x_0\, x_{n-1}$.
- The butterfly permutation denoted β is a bit-shuffling permutation where: $\beta_i^k (x_{n-1}\ x_{n-2} \ldots\ x_{i+1}\ x_i\ x_{i-1}\ \ldots\ x_1\, x_0) = x_{n-1}\ x_{n-2} \ldots\ x_{i+1}\ x_0\ \ x_{i-1}\ \ldots\ x_1\, x_i$.
- The baseline permutation denoted δ is a bit-shuffling permutation where: $\delta_i^k (x_{n-1}\ x_{n-2} \ldots\ x_{i+1}\ x_i\ x_{i-1}\ \ldots\ x_1\, x_0) = x_{n-1}\ x_{n-2} \ldots\ x_{i+1}\ x_0\ x_i\ x_{i-1}\ \ldots\ x_1$.
- The identity permutation denoted I is a bit-shuffling permutation where: $I_{n-1}(x_{n-1}\ x_{n-2} \ldots x_1\, x_0) = x_{n-1}\ x_{n-2} \ldots x_1\, x_0$.

Table I shows the link permutation used in the popular types of Delta MINs which have been designed and analyzed.

Link permutation	stage 0	stage k ∈[1…n-1]	stage n
Omega	σ_{n-1}	σ_{n-1}	I
Baseline	I	δ_i^k	I
Butterfly	δ_i^0	β_i^k	I

Table I: Link permutation of Delta MINs

B. Data exchanged

Traffic passed through the network is composed of fixed size packets. The switching strategy implemented requires dividing packets into flits, so the number of flits per packet is a parameter for simulation analysis.

The traffic is produced in sources, the bytes produced are driven by a random number generator, various offered load are introduced and analyzed for simulation. The targets represent the destination output, they are distributed uniformly, and they are in charge to remove the packets immediately upon arrival.

C. Buffer Sizing

Each one is connected to an input port of a switch, it stores packets, it has a serious impact on the overall area. Various size of FIFO buffers are analyzed to evaluate the amount of buffering resources in the network architecture.

D. Router

Figure 3 shows the router designed with SystemC. It is composed of 2x2 crossbars, a control component (arbiter), a couple of input and output ports and a couple of FIFO connected to each input ports. The connection between the couple change dynamically according the destination in the packet header. The round robin algorithm stored in the arbiter is activated when multiple inputs contain messages routed to the same output. The switching processes also occur in every stage in pipeline.

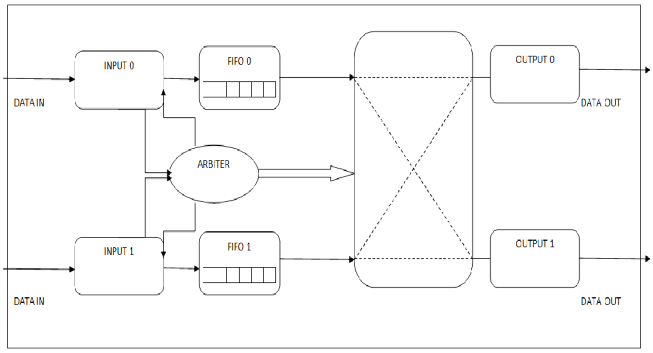

Fig.3: router's architecture

IV. A Dynamic reconfigurable Delta MIN

The goal is to have a communication infrastructure in which the reachability of packets is ensured, independent of the changing topology which occurs when connections are changed on the MPSOC. For enabling reconfiguration in Delta MIN architecture, a link permutation manager is used to change the connections between stages during simulation time. The software manager ensures the reproducibility of the design topology and it is fully implemented using SystemC.

For evaluation performance, a communication scenario is defined as traffic generated at source uniformly. Packets are generated at source, the packet size is 32 bits. Sources and targets are selected randomly, and are injected at the packet header.

The routing between nodes is performed according to method proposed in the precedent section. The buffer size is a crucial parameter and it interacts with the performance parameters, the buffer size in the range of 1 to 5 packets are evaluated in the simulation.

We compare the performance of MINs having the same size of the switch elements to evaluate the dynamicity of the link permutation. Tested MINs are Omega and butterfly. First, we begin with simulation of Omega network for 100 clock cycles, second we remove packets from stages for 3 cycles. Finally, connections change dynamically between the stages to simulate a butterfly network for the same period. We will limit our analysis to three parameters for performance evaluation: latency, FIFO occupation, and communication load.

V. Simulation results

Simulations are run for Delta MINs have 8 inputs and 8 outputs nodes using 2x2 switches elements. The following figures identify the results derived from timing simulation. First, traffic generators and crossbar elements are running at 1 GHz clock frequency. In the second phase, we introduce a second clock to synchronize crossbar activities running three times more than the first clock frequency, only traffic generators follow the first clock tick (figure 4) .

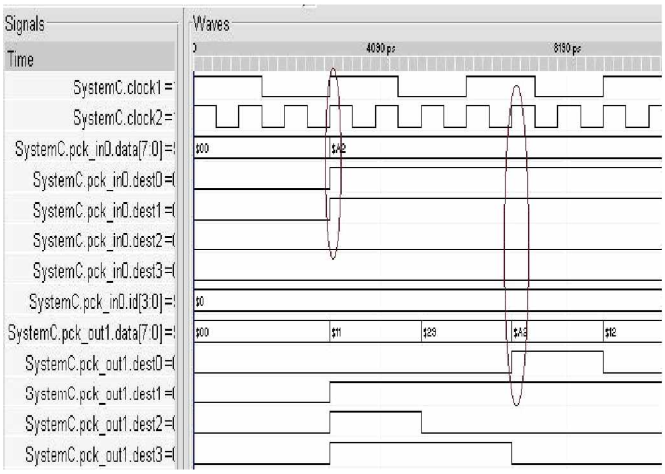

Fig.4: Latency measure with two clocks

We deduce for the first implementation (only one clock) an average latency of network of 14ns, a rate of packets dropped of 48%. In the second model (two clocks), we measure an average latency of network of 5ns, a rate of packets dropped of 0% for the same duration of simulation. Measurements obtained stimulated us for the second model in which each subsystem (processors and network) follows its own clock tick.

A. *Communication load*

We can conclude from figure 5 that communication load is slightly sensitive to the change of connection at run time. The average value of the communication load is 80% for Omega network against 84 % for the butterfly network. Figure 5 also shows a maximum value of load for the two networks after 3 clock cycles of simulation. This graphic can lead to a conclusion that the communication load is directly related to the traffic generated at sources and it also depends on the packet size.

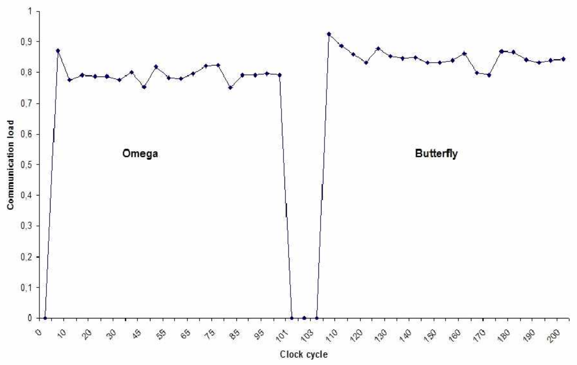

Fig.5: Communication load measures

B. *Latency*

Figure 6 gives a direct perception that latency is more sensitive than communication load when we change topology at run time. For omega network latency increasing is permanent, but it almost stabilizes from 20 to 90 clock cycles. For the second network, we observe a destabilization of the value of latency. It is obvious that latency of a message is sensitive to buffer size, and it depends on the time necessary to transmit the packets when traffic load increases.

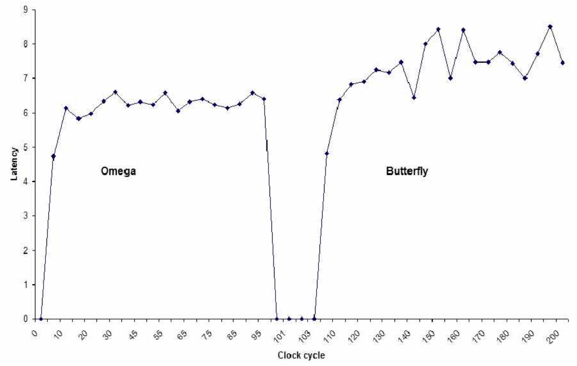

Fig.6: latency vs communication load

C. *FIFO occupation*

Evaluating buffer occupation is a crucial work in switch design and has a serious impact on the overall area of the network. Figure 7 shows a maximum fifo occupation for the two topologies in the second stage. In stage 3, we observe a better FIFO occupation for the omega network.

We deduce that the buffer size depends slightly on the network topology, and it has a serious impact on latency. In other words when traffic load increases, FIFO occupation increases and systematically latency increases.

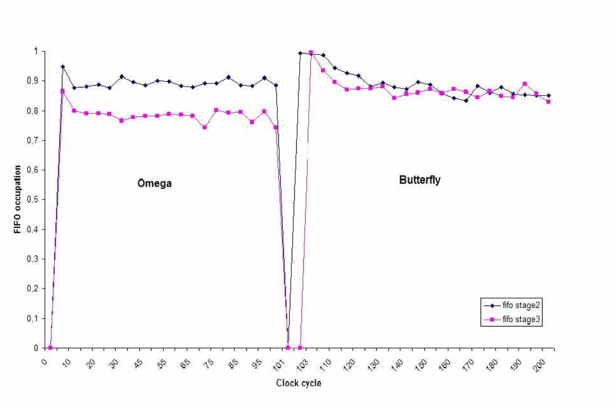

Fig.7: FIFO occupation

We conclude that changing topology at run time slightly affects the performance of the communication platform. This result stimulates us to search for others reconfigurable parameters to know routing and commutation strategies, traffic, components placement.

VI. Conclusion and future work

MINs have been used as interconnection platform in multiprocessor systems. There exist various popular MINs. The difference between each of these networks is the link permutation between the crossbar stages. We proposed in this paper a model of simulation of a generic Delta MIN, in which the reachability of packets is ensured, independent of the changing topology which occurs when connections are changed on the MPSOC. For enabling reconfiguration in Delta MIN architecture, a link permutation manager is used to change the connections between stages during simulation time. With SystemC timed simulations, performance evaluation of ours NOC are given and analyzed.

We plan to extend this work to employ formal methods aimed at representing our communication architecture at a high level of abstraction. This design will identify the components of our MIN, and their properties, from which the dynamicity of the link permutation of Delta MINs is checked.

REFERENCES

[1] W.O. Cesario, "Multi-processor SoC platforms: a component-based design approach", IEEE Design and Test of Computers 19 (6) (2002), pp52–63.

[2] C. Stunkel, "The SP2 High-Performance Switch", IBM Systems Journal 34 (2) (1995), pp. 185–204.

[3] M. Keating, P. Bricaud, "Reuse methodology manual for system on chip design", Kluwer Academic, Boston (2002).

[4] R. Lauwereins, "Creating a world of smart reconfigurable devices", In Field Programmable Logic FPL (2002), pp790-794.

[5] T. Cheung, "A simulation study of the cray x-mp memory system", IEEE Transactions on Computers 35 (7) (1986), pp613–622.

[6] S. Duquennoy, S. Le Beux , P. Marquet , S. Meftali , J. Dekeyser, "MpNOC Design: modeling and simulation", In 15th IP based SOC Design Conference (2006), pp229-232.

[7] C.P. Kruskal, "The performance of multistage interconnection networks for multiprocessors", IEEE Transactions on Computers 37 (11) (1986), pp1091–1098.

[8] C.C. Fan, J. Bruck, "Tolerating Multiple faults in Multistage interconnection Networks with Minimal Extra Stages", IEEE Transactions on Computers 49 (9) (2000), pp998–1004.

[9] D. Tutsch, M. Brenner, "MINSimulate . A Multistage Interconnection Network Simulator",. In 17th European Simulation Multiconference: Foundations for Successful Modeling & Simulation (ESM'03), Nottingham, pp 211.216.

[10] K. Goossens, "Formal Methods for Networks on Chips", In Fifth International Conference on Application of Concurrency to System Design :ACSD (2005), pp188-189.

[11] D. Tutsch, G. Hommel, "Generating Systems of Equations for Performance Evaluation of Buffered Multistage Interconnection Networks", Journal of Parallel and Distributed Computing 62 (2) (2002), pp228–240.

[12] T. Szymanski, V. Hamacher, " On the universality of multistage interconnection networks", *IEEE Computer Society Press* (1994) , pp. 73-101.

[13] C. Kruskal, M. Snir, "A unified theory of interconnection network ", Theoretical Computer Science 48 (1986), pp75-94.

[14] M. Collier, "A systematic analysis of Equivalence in Multistage Networks", In Journal of Light wave Technology 20 (9) (2002), pp228–240.