A New Fault Tolerant Routing Algorithm for Networks on Chip

Chakib Nehnouh, University of Oran1 Ahmed Ben Bella, Oran, Algeria
Mohamed Senouci, University of Oran1 Ahmed Ben Bella, Oran, Algeria

ABSTRACT

To provide correct data transmission and to handle the communication requirements, the routing algorithm should find a new path to steer packets from the source to the destination in a faulty network. Many solutions have been proposed to overcome faults in network-on-chips (NoCs). This article introduces a new fault-tolerant routing algorithm, to tolerate permanent and transient faults in NoCs. This solution called DINRA can satisfy simultaneously congestion avoidance and fault tolerance. In this work, a novel approach inspired by Catnap is proposed for NoCs using local and global congestion detection mechanisms with a hierarchical sub-network architecture. The evaluation (on reliability, latency and throughput) shows the effectiveness of this approach to improve the NoC performances compared to state of art. In addition, with the test module and fault register integrated in the basic architecture, the routers are able to detect faults dynamically and re-route packets to fault-free and congestion-free zones.

KEYWORDS
Congestion, Fault Tolerance, Network on Chip, Reliability, Routing Algorithm, Sub-Network

1. INTRODUCTION

Network-on-chip (NoC) has emerged as an efficient architecture to manage communication in a system on chip (SoC), where a large number of components and storage blocks are integrated on a single chip. This intensification of communications leads to important questions such as performance and energy consumption. Decreasing the transistor size has made semiconductors more sensitive to faults. Thus, the challenge is, to maintain the system functionality during its operational lifetime and ensure that the system performance is preserved. For this reason, researchers have attached a great deal of importance to the reliability in networks on chip. Faults which may occur in networks on chip can be divided into two main categories: permanent faults (or hard faults), and temporary faults (or soft faults) (Radetzki et al., 2013). The soft faults are classified into transient and intermittent. These three types of faults (permanent, transient and intermittent) can be caused by several internal or external factors. The majority of failures (80%) are caused by transient faults, whilst the rest of them originate mainly in permanent and intermittent faults (Lehtonen et al., 2007).

Faults in different components of the NoC have different causes, however, all can result in serious consequences such as loss of packet data, misrouting, deadlocks, and malfunctions. It follows that, the reliability of communication becomes an attracting challenge when designing the NoC. However, communication has a huge impact on the performance of the network on chip and its more desirable to design efficient algorithms to ensure that. Fault in routers or links is the major problem that causes the failure on transmitting the packet in a NoC. Communication performance of a NoC is depended

DOI: 10.4018/IJERTCS.2019070105

Copyright © 2019, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.
highly on the routing algorithm. Routing algorithm determines the path that each packet follows between source and destination node.

The fault tolerance routing algorithm is the process of finding a new path to steer packets from sources to destinations in a faulty network, by choosing an optimal path, the routing algorithm can efficiently increase the performance of the network. Congestion is another key factor which leads to increase the transmission delay and power consumption. For this, routing algorithms can improve performance by re-routing packets through less congested regions and distributing traffic over the network. Finally, failures and congestions should be managed in an effective way to ensure availability and robustness into the network on the chip.

In this context, many fault tolerance routing algorithms have been proposed for critical applications. It is, therefore, essential to handle failure, ensure correct and continuous operation of the circuit in its environment, even when the failure rate is high.

To achieve these objectives, we propose a new approach in the domain of fault- tolerant NoC with two main contributions. Firstly, we propose a unified fault model that includes transient faults, permanent faults, and the congestion considered as a fault. Secondly, we present a new architecture based on sub-nets. This architecture is able to achieve low latency and increase the network bandwidth. Additionally, our architecture is capable of handling multiple link and router failures up to 40% and neither utilizes any VC (Virtual Channels). Finally, the solution is deadlock-free and congestion-aware.

We describe in this paper an interesting solution jointly to the congestion management and fault tolerance in NoC called DINRA. The new architecture offers many advantages like reducing latency and the use of alternative paths to route packets in case of faulty links or/and routers. The rest of the paper is organized as follows. Section 2 gives a brief overview of the work. The new architecture is presented in the third section. Implementation details of the proposed solution are given in section 4. In Section 5, DINRA-FTNoC is evaluated. Our conclusions are drawn in the final section.

2. BACKGROUND

Many solutions have been proposed to tackle fault-tolerance in NoC (Chatterjee et al., 2014; Bishnoi et al., 2015; Ren et al., 2015; Poluri & Louri, 2016), we can say that the most effective ones can be grouped into two classes based on their ability to manage the congestion (if they can or cannot). Other classification is based on the fault’s type, and the mechanism to deal with it using routing algorithms or architectural solutions. The other class can tolerate an arbitrary number of faults without taking into account the network congestion state. To deal with that, it is highly recommended to use routing tables and logic to update any fault occurrence (Fick et al., 2009; Fu et al., 2014). When a link or switch is faulty, the routing algorithm should choose another fault-free route for transferring the packets and avoiding the corruption of data. The main challenge to increase the reliability and provide a good performance is to deal with the principal problems in routing algorithms such as deadlock, congestion and failures. In the literature, there are many various approaches which have been suggested to solve this issue. Figure 1 summarizes some significant works about fault-tolerant techniques in NoC.

The disadvantages of these methods are that they require an additional computation step to configure the routing logic and suffer from the high reconfiguration overhead. Another drawback of this class is that the routing tables limit the silicon area overhead (Wachter et al., 2013). The work presented in (Vitkovskiy et al., 2013) exploits bypass rules to reroute around faults using the local connectivity information and they can only support few faults before the network disconnects. The congestion can decrease significantly throughput and the latency. Moreover, the selection of the best path to route the packets in the NoC minimizes the traffic on the network and maximizes the performances. Despite that, few works in literature have integrated the congestion and fault awareness.

The Gradient (Pratomo & Pillement, 2012) suggested an adaptive routing algorithm to tolerate faulty node. The proposed routing algorithm divides the entire system in eight zones by gradient line. Each zone is provided with one main path and two alternative paths for the packet delivery when the
An Analysis of Traffic and Throughput for UMTS Packet Core Networks
Ye Ouyang and M. Hosein Fallah (2012). Research, Practice, and Educational Advancements in Telecommunications and Networking (pp. 82-107).
www.igi-global.com/chapter/analysis-traffic-throughput-umts-packet/62760?camid=4v1a

Visualization of Real-Time Radar Data by Integration of X-Band Software
www.igi-global.com/article/visualization-of-real-time-radar-data-by-integration-of-x-band-software/130892?camid=4v1a