Mapping and Scheduling Techniques for Network-on-Chip Architecture

Mohammad Behrouzian Nejad¹, Ebrahim Behrouzian Nejad², Aref Sayahi³, Sayed Mohsen Hashemi⁴, Javad Chaharlang⁵

¹Young Researchers and Elite Club, Dezful Branch, Islamic Azad University, Dezful, Iran
²Department of Computer Engineering, Shoushtar Branch, Islamic Azad University, Shoushtar, Iran
³Department of Computer Engineering, Soosangerd Branch, Islamic Azad University, Soosangerd, Iran
⁴Young Researchers and Elite Club, Mayboad Branch, Islamic Azad University, Mayboad, Iran
⁵Department of Computer Engineering, Masjed Soleyman Branch, Islamic Azad University, Masjed Soleyman, Iran

*Corresponding Author Email: mohamad.behrouzian@gmail.com

Abstract

Mapping and scheduling problems are important steps in the on-chip network design, which the goal is usually to optimize design and minimize power consumption and run time of an application. A good technique should be considered the design constraints. Indeed, the mapping quality varies with consideration of design constraints and optimization objectives such as minimizing energy consumption, communication delays. To improve performance of Network-on-Chip (NoC), many assignment algorithms are proposed to minimize the communication overhead and energy consumption. In this paper, we described the basic ideas of NoC architecture and review and analyse powerfully and efficient mapping and scheduling techniques, which proposed for NoC architecture. According to this problem that these techniques can have advantages and disadvantages, this paper focused on the review of these techniques. Review result of these techniques is described in the table form.

Keywords: NoC, System-on-Chip (SoC), Mapping, Scheduling.

Introduction

NoC is adopted for better data communication between PEs than SoCs (Jueping et al., 2010). Size and power consumption in the SoC level are very important (Chao et al., 2011). In the mapping problem, delay and bandwidth constraints and communication problem, the goal is to optimize and energy saving. So far, many algorithms have been presented to find solutions to optimizing (Verdipoor, 2009; Jang & Pan, 2010; Chou et al., 2008; Lee et al., 2010). To improve performance of NoC, assignment algorithms are proposed to minimize the communication overhead and energy consumption, for example algorithms such as, Fixed Center (FC), Best Neighbor (BN) searching algorithm, Simulated Annealing (SA). To reduce latency of transferring data between PEs, (Jueping et al., 2010) proposed central shared memory NoC architecture to improve communication throughput by direct transferring data from/to Shared Memory (SM) to increase efficient communication. In this paper, we reviewed some of the most popular and efficient mapping and scheduling techniques, which proposed for NoC architecture. According to this problem that these techniques can have advantages and disadvantages, review of these techniques in this paper can help to provide new efficient techniques. The reminder of this paper is organized as follows. In Section II, we review the NoC idea. Section III described the mapping and scheduling problems. In section IV, mapping and scheduling techniques in NoC are described. In Section V, review result and attributes of techniques are presented. VI Section contains the conclusion and future work of the paper.

Methodology

NOC Idea

NoCs are derived from computer networks and distributed computing in large-scale (Wu et al., 2006). Communication infrastructure between resources can be achieved by NoC. In the NoC PEs are connected together based on a communication infrastructure contains switch or router through communication links, which called interconnection network. Figure 1 shows a sample NoC includes 16 core. Each core is connected to a switch by a network interface. Communications between cores in NoC is
established by sending message by a path consisting of a series of switches and interswitch links (Hariharan, 2005; Bjerregaard & Mahadevan, 2006).

Routing in NoC

Routing in NoC is like to routing in other networks. A routing algorithm determines that how packets routed between sender and receiver nodes. According to constraints on memory and computing resources in NoC, the routing algorithms that proposed for NoC should be reasonably simple (Ni & McKinley, 1993; Behrouzian-nejad et al., 2011; Rantala et al., 2006). In the NoC routing algorithms can be dividing into two categories, deterministic and adaptive. Deterministic routing algorithms choose one path between source node and destination node. To this type of routing algorithms not important traffic conditions and congestion on the network when choose path for sending packets. When network congestion is low, these algorithms with less routing delay send packets. Other type of routing algorithms is adaptive routing, which distributes traffic dynamically in response to the network state. When network congestion is high, these algorithms can have better performance than deterministic algorithms.

NoC Platforms

NoC platforms are dividing into three categories: hard, firm and soft. In hard NoC platforms any change in the network is not acceptable which this attribute can be reduce performance of NoC. This platform is not flexible. In firm platform can done mapping, it can be done as you wish. Finally, soft platform have more flexibility than other platforms. Fig.2 shows different NoC Platforms (Hu & Marculescu, 2005-a).
Mapping and scheduling problem

Mapping and scheduling are important steps in the on-chip network design, which the goal is usually to optimize design and minimize power consumption and run time of an application. Mapping problem is one of the NoC design steps which in this an application can be implemented in a specific architecture. Points to be considered in the mapping are:

- Application model
- NoC architecture model
- Energy and performance constraints
- Objective functions being optimized
- A good mapping should be considered the design constraints. Indeed, the mapping quality varies with consideration of design constraints and optimization objectives such as minimizing energy consumption, communication delays and etc. Mapping can be done either static or dynamic. In static mapping, assigning tasks to the resources of a particular architecture is done before the run time of application and at the run time of the application is not performing any other mapping. In dynamic mapping, assigning tasks to the resources of a particular architecture is done before the run time of application and at the run time of the application to be performed tasks allocation of other applications. This mapping should lead to better results, but because this type of mapping algorithms need too many calculations, latency and power consuming of applications will increase in run-time. In addition, test and experiment of this type of mapping also is difficult (Verdipoor, 2009; Pop & Kumar, 2004).

Mapping and scheduling techniques in NOC

Raina and Muthukumar (2009) have presented a heuristic and traffic aware mapping and scheduling algorithm for hard NoCs. This algorithm acts static and NoC is assumed to be homogeneous, so, in this algorithm, the cost of any PEs is not considered in the scheduling and mapping. Goals of this algorithm are reduce traffic and NoC delay. Rajaei et al (2010) have presented an energy aware and heuristic scheduling and mapping algorithm. This algorithm is suitable for scheduling and mapping tasks to real-time applications on heterogeneous and hard NoC. This algorithm simultaneously considers the costs of processing of PEs and communication in chip. This algorithm can reduce energy consumption and traffic of NoC. Jueping et al. (2010) proposed a new energy and communication aware architecture for NoC with a Shared Memory (SM) on each node to reduce communication delay between PEs. In this architecture PE to PE communication are divide into two steps: PE to SM and SM to PE. With using Simulated Annealing (SA) algorithm authors can reduce power consumption and transferring cycles. Jang and Pan (2010) proposed a new algorithm for NoC which called Architecture-Aware Analytic Mapping (A3MAP). This algorithm proposed for homogeneous NoC with regular mesh topology, heterogeneous NoC with irregular mesh topology and custom NoC architecture. Authors use of two algorithms for development of itself algorithm: relaxation algorithm and genetic algorithm. Performance of this algorithm with genetic algorithm show that can be better than relaxation algorithm to reduce traffic in homogeneous NoC with regular Mesh, heterogeneous NoC with irregular mesh topology and custom NoC architecture.

Hu and Marculescu (2005-a) proposed a new Energy-Aware Scheduling (EAS) algorithm. This algorithm presented for statically scheduling of specific transactions and computing tasks in heterogeneous NoCs. This algorithm automatically distributes application tasks onto PEs and schedules under real-time constraints. This algorithm instead of considering the performance criteria as the objective optimization, try reducing energy consumption under hard performance limitations. In this algorithm communication scheduling and computing tasks scheduling can done parallel. Mehran et al (2007) proposed a new core mapping algorithm for NoC that called Spiral. To develop this algorithm used of heuristic method and presented for two dimension mesh topology. Authors used of XY routing algorithm to implement and compared this algorithm with genetic algorithm and random mapping algorithm. This algorithm reduced energy consumption than genetic and random algorithm. Authors show that this algorithm is fast and can use for fast dynamic core mapping. Chou and Marculescu (2008) first reviewed effect of different type of contention in the network on the application mapping and then proposed a contention-aware mapping technique for reduce contention and energy consumption in the NoC. Authors implement this technique with using XY routing algorithm in 2-D mesh topology. Authors show that with using of this technique packet latency can be reduced. Hu and Marculescu (2005-b) proposed a new energy and performance aware mapping algorithm. This algorithm presented for regular NoC. For develop this algorithm authors use of deterministic routing which is deadlock free. Authors show that this algorithm can be reduce communication energy than ad hoc implementation and can be use for different regular NoC topologies.

Bandwidth can be an effective role on energy consumption in the NoC architecture. To reduce bandwidth requirement of NoC, Wang et al (2011) proposed a new mapping technique to task scheduling of an application. In this technique use of Ant Colony Optimization technique to task mapping. Authors show that this technique can reduce bandwidth requirement and cost of NoC implementation. Wu et al (2011) has proposed a new Energy-aware Mapping Algorithm which called GA-MMAS. This algorithm used from Genetic Algorithm (GA) and MAX-MIN Ant System Algorithm (MMAS) to reduce energy consumption of NoC. Authors for this algorithm first improve MMAS, and then with combination of MMAS and GA, make compensation to lack of pheromones in the initial stage of MMAS. Increase the accuracy of the optimal solution, which will reduce energy consumption.
Results

Review result of mapping and scheduling techniques

Table I. show the some attributes of mapping and scheduling techniques that reviewed in this paper.

<table>
<thead>
<tr>
<th>Techniques Names</th>
<th>References</th>
<th>Attributes of techniques</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traffic aware scheduling</td>
<td>Raina and Muthukumar (2009)</td>
<td>• Reduce traffic and NoC delay</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Not considered cost of any PEs in the scheduling and mapping</td>
</tr>
<tr>
<td>Energy aware algorithm</td>
<td>Rajaee et al (2010)</td>
<td>• Reduce energy consumption and traffic of NoC.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Considered costs of processing of PEs</td>
</tr>
<tr>
<td>Energy and communication aware</td>
<td>Jueping et al (2010)</td>
<td>• Using of new architecture with SM</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Reduce power consumption and transferring cycles</td>
</tr>
<tr>
<td>A3MAP</td>
<td>Jang and Pan (2010)</td>
<td>• Reduce traffic</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Useful for regular and irregular Mesh, heterogeneous NoC and custom NoC</td>
</tr>
<tr>
<td>EAS</td>
<td>Hu and Marculescu (2005-a)</td>
<td>• Reducing energy consumption</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Parallelism communication scheduling and computing tasks</td>
</tr>
<tr>
<td>Spiral</td>
<td>Mehran et al (2007)</td>
<td>• Reduced energy consumption</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Fast algorithm</td>
</tr>
<tr>
<td>Contention-aware mapping</td>
<td>Chou and Marculescu (2008)</td>
<td>• Reduce contention and energy consumption</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Reduce packet latency</td>
</tr>
<tr>
<td>Energy and performance aware</td>
<td>Hu and Marculescu (2005-b)</td>
<td>• Reduce communication energy</td>
</tr>
<tr>
<td>mapping</td>
<td></td>
<td>• Useful for different regular NoC topologies</td>
</tr>
<tr>
<td>Bandwidth aware application</td>
<td>Wang et al (2001)</td>
<td>• Reduce bandwidth requirement</td>
</tr>
<tr>
<td>mapping</td>
<td></td>
<td>• Reduce cost of NoC implementation.</td>
</tr>
<tr>
<td>GA-MMAS</td>
<td>Wu et al (2011)</td>
<td>• Reduce energy consumption</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• More computing time</td>
</tr>
</tbody>
</table>

Conclusion

To improve performance of NoC, many assignment algorithms are proposed to minimize the communication overhead and energy consumption. In this paper some of the most popular and efficient mapping and scheduling techniques which proposed for NoC, introduced and examined. According to this problem that these techniques can have advantages and disadvantages, review of these techniques in this paper can help to authors to provide new efficient techniques. As future work we propose more use from heuristic and optimization algorithms such as ant colony, PSO, simulated annulling and etc to improving mapping and scheduling techniques.

References