



Contents lists available at ScienceDirect

Journal of Systems Architecture

journal homepage: www.elsevier.com/locate/sysarc

An efficient task mapping algorithm with power-aware optimization for network on chip

Wei Hu^{a,b,*}, Qingsong Shi^c, Yonghao Wang^d, Kai Zhang^{a,b}, Jun Liu^{a,b}, Xiaoming Liu^{a,b}, Hong Guo^{a,b}

^a College of Computer Science and Technology, Wuhan University of Science and Technology, Wuhan, Hubei, China, 430065

^b Hubei Province Key Laboratory of Intelligent Information Processing and Real-time Industrial System, Wuhan, Hubei, China, 430065

^c College of Computer Science, Zhejiang University, Hangzhou, Zhejiang, China, 310027

^d Centre of Digital Media Technology, Faculty of Computing, Engineering and the Built Environment, Birmingham City University, United Kingdom

ARTICLE INFO

Article history:

Received 9 September 2015

Revised 6 February 2016

Accepted 14 April 2016

Available online xxx

Keywords:

Power consumption

Mapping algorithm

Communication

Network on chip

ABSTRACT

More and more cores are integrated onto a single chip to improve the performance and reduce the power consumption of CPU without the increased frequency. The cores are connected by lines and organized as a network, which is called network on chip (NOC) as the promising paradigm of the processor design. However, it is still a challenge to enhance performance with lower power consumption. The core issue is how to map the tasks to the different cores to take full advantages of the on-chip network. In this paper, we proposed a novel mapping algorithm with power-aware optimization for NOC. The traffic of the tasks will be analyzed. The tasks of the same application with high communication with the others will be mapped to the on-chip network as neighborhoods. And then the tasks of different applications are mapped to the cores step by step. The mapping of the tasks and the cores is computed at run-time dynamically and implement online. The experimental results showed that this proposed algorithm can reduce the power consumption in communication and the performance enhanced.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

With the development of semiconductor technology, more and more transistors can be integrated onto a single chip. However, it also has a big problem to increase the frequency of the single processor core for the faster growth of power-consuming [1]. Hardware manufacturers have begun to focus on the development of on-chip systems with more than one core. More processor cores are integrated onto the CPU, which is called Chip Multi Processor (CMP). The multiple processor cores enhance the performance of the system without increasing the frequency of CPU. Bus-based communication is the traditional architecture to connect the on-chip devices, which is faster but needs more on-chip size. If more cores have the communication requirements, the bus itself will be the bottleneck of the CPU performance both in power-consuming and the transmission speed. It results in system performance degradation. Network on chip (NOC) is proposed as the promising paradigm to solve this problem [2–4]. Typical NOC has an on-chip network to connect the processor cores, the processing

units, memory blocks or the other on-chip devices [5]. The processor cores on NOC are distributed on the chip via the lines not the traditional buses [6,7]. The on-chip network makes NOC more scalable in communication. On-chip cores can communicate with each other through the network via on-chip routers for high efficiency. The design of NOC is communication-centric, not computation-centric [8]. When more and more processor cores are integrated onto a single chip, the communication will be more important than the computation.

The applications are divided into a plurality of tasks in the system for running simultaneously on multiple cores in NOC. The tasks from the partitioning are very important to improve the efficiency of the multi-core processors. They have to communicate with each other on the basis of their relationship to complete the target of the applications. Different applications have different traffic characteristics, which are the necessary requirements for NOC. When the traffic is heavy, the long latency from ends to ends will affect the system performance seriously. The power consumed by such traffic will be the main portion of the system power consumption [9]. When there are many tasks, the on-chip distribution of these tasks has great impact on the performance and power consumption. It plays an important role in performance enhancing and power saving to schedule the tasks to the cores for less traffic.

* Corresponding author. Tel.: +8618607100789.

E-mail address: huwei@wust.edu.cn (W. Hu).

In this paper, a novel online mapping approach is proposed to provide the task mapping with high efficiency for NOC. The communication between the tasks is analyzed. And the mapping algorithm re-maps the tasks according to the analysis results. The tasks of the same application with more communication with others are mapped as neighborhoods to reduce the power consumption.

This paper is organized as the follows. Section 2 describes the related works. Section 3 describes the system model of NOC. Section 4 presents the design of the algorithm. The experiments and results are described and discussed in Section 5. And at last, we give the conclusions and future work in Section 6.

2. Related work

On-chip network is the basic infrastructure provided by NOC. According to the requirements of the tasks, the communication is distributed on the network. Structures of NOC itself are optimized to provide better support for the on-chip network [10–13]. Compared with the traditional bus structures, the communication on lines will consume more time. The delay on lines may be the bottleneck of the performance. How to map the tasks to the cores is one of the key approaches to solve the above problem. There are existing works on this problem to provide better solutions. Mapping algorithms are proposed first including the heuristic mapping algorithm [14], genetic algorithm based approaches [15], the algorithms with QoS [16] and Mesh oriented multi-target algorithms [17]. In algorithms, the mapping process is determined before the operation of the system. It is an advantage that the mapping can be optimized and has a good performance. However, such mapping cannot be adjusted according to the specific situations at run-time. As the changes occur, the mappings have to be re-computed. This is a very frustrating and time-consuming process.

Online mapping is also an important research area. Run-time mapping is similar to the mapping in a traditional operating system. The tasks will be mapped dynamically according to the run-time environments. [18] focused on the resource allocation and thread immigration at run-time based on the network characteristics of NOC. [19] proposed task predication and allocation based on the common task sequence constructed through the user habits as the reference impact. [20] assumed the cores had different power consumption level. And it proposed that the power consumption could be reduced by allocating tasks to the cores with lower power-consuming as much as possible. Scenario based mapping was provided in [21]. The scenarios were taken as a state machine and the scenario transitions were the task mapping. Common tasks were separated from the tasks in an operating system in [22]. They were mapped to different cores independently to reduce the interferences of such tasks. DVS/DVFS was also used in task mapping for power saving [23–25]. In these approaches, the frequency or voltage of the cores might be adjusted according to the run-time analysis. The frequency or voltage scaled down could reduce the power consumption of the whole system.

Tasks running on NOC have the communication requirements to achieve the target. The mapping of tasks on the network determines the traffic density. Related works show that the mapping algorithm was important for the performance of NOC. Optimized mapping algorithm can also reduce the power consumption. In this paper, a dynamic online mapping algorithm is proposed. It relies on the analysis of the traffic in the on-chip network at run-time. This algorithm can reduce the power consumption of NOC and have good impact on the performance.

3. System model

NOC has new features and different designs. [26] provided a general description of NOC architectures, applications and the re-

lated algorithms. In this section, the system model is introduced including the on-chip network topology, the routing policy, task model and the energy model. They are the basis of the mapping algorithm.

3.1. On-chip network topology

NOC is a novel design paradigm of system on chip (SOC), which has many advantages over bus-based communication. The on-chip network connects the on-chip devices and provides better performance. Various network topologies have been proposed for NOC architecture such as Ring, Mesh and Torus [27]. Mesh is the typical mainstream choice of NOC design. The mesh-based NOC architecture is shown in Fig. 1.

Mesh has similar structure to matrix. The wires are used to connect the tiles as shown in Fig. 1 (a). Tiles are the nodes of the on-chip network. Each tile has the following components including routers, input/output interfaces, processor cores and on-chip memory(cache/SPM) as shown in Fig. 1. Routers and input/output interfaces are responsible for the data forwarding and the communication between the cores. The tile may have non-conventional process core. There are probably processing elements (PE) for special purposes in tiles. Cache/SPM is local memory for the cores or PEs.

3.2. Routing policy

Routing policy provides the mechanism to determine the paths from source nodes to destinations. A routing algorithm is designed to improve the performance of on-chip communication and solve the problems such as deadlock and congestions. Routing algorithms can be classified into two types including deterministic routing and adaptive routing [8]. If a deterministic routing algorithm is used, the traversal paths are determined by all packets transmitted from the source node to the destination node in the network. Adaptive routing chooses the routing direction according to the run-time environments dynamically. Such adaptive routing has strict requirements on router design which increases the complexity.

X-Y routing is deterministic routing algorithm for mesh-based NOC [28]. In X-Y routing, the mesh structure is marked as X direction and Y direction. The routing is first along the X direction to forward the packets and then the Y direction as shown in Fig. 2. For a given destination, X-Y routing can reach it without deadlock. It has high simplicity. In this paper, X-Y routing is chosen as the on-chip routing algorithm for the least impact by routing algorithm itself.

Switching mode is important for data transferring on NOC. Wormhole switching is packet based switching mode and it is one of the primary switching mechanisms currently [29]. The packets are split into several small segments called flits. These flits are transferred through the network. If there are enough buffers for one flit, this flit can be buffered and forwarded by the router. Such method has reduced the network latency and saved the buffer. If the flits are blocked, the following flits can be forwarded by the other routers, which enhance the throughput of the network. In this paper, wormhole switching is adopted as the switching mode. The disadvantage of wormhole routing is the latency may increase significantly when the traffic is heavy. Such situations are rare according to our setup and have little impact on the achievements.

3.3. Task model

Application Control Graph (ACG) is used to represent an application, which is similar to the offline analysis in [16]. A typical ACG is shown in Fig. 3. An application can be presented as $G(V, R)$, in which V is the vertices set and R is the directed edge set between

Download English Version:

<https://daneshyari.com/en/article/4956304>

Download Persian Version:

<https://daneshyari.com/article/4956304>

[Daneshyari.com](https://daneshyari.com)