J. Parallel Distrib. Comput. 74 (2014) 1858–1871

Contents lists available at ScienceDirect

J. Parallel Distrib. Comput.

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

Dual partitioning multicasting for high-performance on-chip networks

Jianhua Li [a,*], Liang Shi [a], Chun Jason Xue [b], Yinlong Xu [a]

[a] College of Computer Science and Technology, University of Science and Technology of China, Hefei, Anhui, PR China

[b] Department of Computer Science, City University of Hong Kong, Kowloon, Hong Kong

HIGHLIGHTS

- Multicast traffic threatens the scalability of on-chip unicasting mechanisms.
- We propose Dual Partitioning Multicasting to balance the network link usage.
- DPM simultaneously yields high performance for unicast traffic.
- DPM effectively improves average packet latency and network power dissipation.
- DPM yields better scalability compared with previous work.

ARTICLE INFO

Article history:
Received 17 December 2011
Received in revised form
23 July 2013
Accepted 26 July 2013
Available online 6 August 2013

Keywords:
On-chip network
Multicast routing
Rectilinear Steiner tree
Latency-aware
Load-balance

ABSTRACT

As the number of cores integrated onto a single chip increases, power dissipation and network latency become ever-increasingly stringent. On-chip network provides an efficient and scalable interconnection paradigm for chip multiprocessors (CMPs), wherein one-to-many (multicast) communication is universal for such platforms. Without efficient multicasting support, traditional unicasting on-chip networks will be low efficiency in tackling such multicast communication. In this paper, we propose Dual Partitioning Multicasting (DPM) to reduce packet latency and balance network resource utilization. Specifically, DPM scheme adaptively makes routing decisions based on the network load-balance level as well as the link sharing patterns characterized by the distribution of the multicasting destinations. Extensive experimental results for synthetic traffic as well as real applications show that compared with the recently proposed RPM scheme, DPM significantly reduces the average packet latency and mitigates the network power consumption. More importantly, DPM is highly scalable for future on-chip networks with heavy traffic load and varieties of traffic patterns.

1. Introduction

The continuous decrease in transistor size has led to the persistent increase in the number of cores that can be integrated into CMP systems [4,16,17,36]. On-chip network [5] provides an efficient as well as scalable communication paradigm for CMP systems. Recent work [18] found that, without efficient multicasting support, many applications for CMPs, such as cache coherence protocols [25,31,32] and operand networks [4,29], will suffer from significant performance degradation. For on-chip networks without multicasting support, N unicast packets are injected into the network in order to transmit each multicast packet with N multicasting destinations. Such routing mechanism is traffic-intensive which can cause network congestion as well as unacceptable packet latency. As presented in [18], in a 4×4 mesh network with the state-of-the-art packet-switched unicast router, the saturation point drops rapidly when 1% of the injected packets are multicast packets. As a result, efficient multicasting support is imperative for such systems. This paper proposes a novel multicast routing scheme which can effectively route the on-chip multicast traffic with low latency and low power dissipation.

How to effectively choose the paths for transmitting the multicast packets is the major challenge for efficient multicasting. Judicious routing can reduce the power dissipation and on-chip network traffic. Several research works have been done on the multicasting problem for on-chip networks. Lu et al. [22] proposed a path-based connection-oriented multicasting for wormhole-switched on-chip networks. The advantage of path-based multicasting is its simple implementation. However, as the constructed path may be long, it will force the packets to be transmitted along

* Corresponding author.
E-mail addresses: jianhual@mail.ustc.edu.cn (J. Li), shil0704@mail.ustc.edu.cn (L. Shi), jasonxue@cityu.edu.hk (C.J. Xue), ylxu@ustc.edu.cn (Y. Xu).

0743-7315/$ – see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.jpdc.2013.07.002

the long path which increases the packet latency even under low or medium network load. Abad et al. [1] proposed on-chip hardware support for multicasting which is based on a special router called multicast rotary router (MRR). A fully adaptive tree is used to transmit the multicast traffic. A considerably complex mechanism is proposed to avoid deadlock in [1]. Virtual Circuit Tree Multicasting [18] is a routing table based multicasting mechanism and its advantages and disadvantages have been analyzed in recent work [28,35]. bLBDR [28] is a logic-based multicasting scheme primarily proposed for irregular topology. In addition, the logical-based bLBDR scheme removes the area-consuming routing table. RPM [35] is also not based on a routing table. The route calculation in RPM is based on the global distribution of the multicasting destinations in a mesh-based on-chip network. The transmission scheme in RPM could lead to ineffective routing decisions as indicated in [21,23].

All the previous works are either based on routing table support or dedicated to multicast communication while neglecting unicast communication. In this paper, we propose a novel multicasting scheme, Dual Partitioning Multicasting, to route multicast packets with low latency and low power consumption without routing table support. More importantly, DPM also achieves high performance for unicast communication. Under the DPM scheme, multicast packets are partitioned into two categories through exploiting the link sharing patterns based on the knowledge of global distribution of multicasting destinations. For unicast packets, we propose an adaptive and *multicast-sensitive* packet type assignment mechanism which further enhances the scalability of DPM through balancing the network traffic. For each category of packets, we propose corresponding multicast routing algorithms for transmission. The proposed DPM scheme can obtain the near-optimal worst-case throughput for unicast traffic which makes DPM adaptable under on-chip traffic with varieties of unicast traffic patterns and different percentage of unicast traffic. Extensive experimental results show that DPM can significantly reduce the average packet latency and mitigate the power consumption compared with RPM [35].

The main contributions of this paper are as follows:

- We propose to divide multicast packets into two categories through exploiting the link sharing patterns according to the distribution of the multicasting destinations. Two methods, called RST and B-SLC respectively, are proposed to categorize the multicast packets.
- We propose dedicated routing algorithms to route the two categories of packets efficiently. Moreover, the dedicated routing algorithms can obtain near-optimal worst-case performance for unicast traffic.
- We propose the E-SLC approach which can get more accurate multicast packet categorization with very low time complexity. Additionally, we propose *multicast-sensitive* unicast assignment to make DPM highly scalable for non-uniform traffic patterns.

The remainder of this paper is organized as follows. The background of network partitioning and the motivation of this work are illustrated in Section 2. The proposed DPM scheme is introduced in Section 3. The implementation of DPM is presented in Section 4. Simulation and results analysis are discussed in Section 5. Finally, Section 6 concludes this paper.

2. Background and motivation

Recent work [35] proposed a multicasting scheme called recursive partitioning multicasting (RPM) for mesh-based on-chip networks. In RPM, route calculation is based on the global distribution of the multicasting source and all the multicasting destinations. The whole mesh network is logically partitioned into 3, 5 or 8 parts according to the location of the multicasting source. For on-chip network, if a multicasting source locates in a center tile, the whole mesh is partitioned into 8 parts[1] which is depicted in Fig. 1.

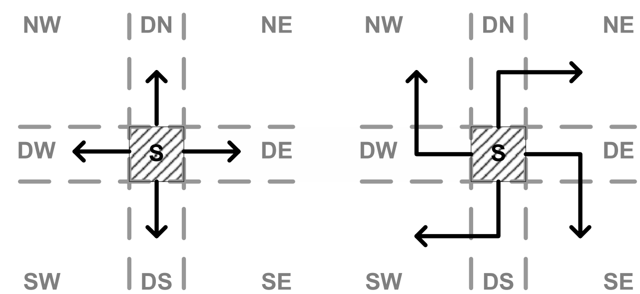

Fig. 1. Partitioning of on-chip networks.

Based on the above network partitioning, RPM proposed several multicasting rules to transmit network packets. Specifically, RPM has higher priority to first route the packet along south or north links. As depicted in Fig. 2(a), if multicasting destinations locate in both the NE and NW part, RPM will first transmit the packet to the north output port of the source tile. In the same way, if multicasting destinations locate in both the SE and SW part, RPM will first transmit the packet to the south output port of the source tile as shown in Fig. 2(b). Such sharing routing will reduce network traffic as well as link usage. However, west-sharing routing and east-sharing routing are restricted in RPM which are demonstrated in Fig. 2(c) and (d). The restrictions in Fig. 2(c) and (d) could cause unbalanced link usage which can exacerbate network performance.

Fig. 3(a) shows the RPM's routing solution for a sample multicast packet whose source is tile 14 and destinations are tiles 5 and 15. The optimal routing solution for this packet is shown in Fig. 3(b). RPM uses 10 network links which is 1.67χ of the link usage in the optimal routing as demonstrated in Fig. 3. However, the optimal routing is restricted by the routing rules in RPM which leads to more network link usage. Unnecessary link usage and router traversal consume more power and generate more network traffic. Moreover, if any one of the transmissions (packet to south destinations and packet to north destinations) is blocked at the multicasting source, the subsequent packets generated by the same source will be blocked. Therefore, a *good multicasting scheme should balance the network link usage*.

Fig. 1(b) presents the routing rules for unicast packet under RPM [35]. RPM utilizes shortest path routing, wherein the unicast traffic to NE and SW parts is routed based on deterministic YX routing and the unicast traffic to NW and SE parts is transmitted by deterministic XY routing. Seo et al. [30] showed that such static routing suffers from poor worst-case and average-case throughput. As a result, a multicasting scheme, which is also *unicast-aware*, is beneficial for on-chip network performance. In other words, a *robust multicasting scheme should simultaneously yield high performance for unicast traffic.*

Motivated by the above analysis, in this work, we propose a novel multicasting scheme called DPM. Under the DPM routing scheme, multicast packets are divided into two categories. One category exhibits more vertical links sharing like the cases in Fig. 2(a) and (b). The other category owns more horizontal links sharing similar to the scenarios in Fig. 2(c) and (d). How to categorize the multicast packets is presented in Section 3.1. We then propose four multicast routing algorithms, which are illustrated in Section 3.2,

[1] A multicasting source located in a corner tile will partition the network into 3 parts while a boundary tile source will partition the network into 5 parts. The capital letters in Fig. 1 (NW, DN, etc.) represent the eight orientations.