Contents lists available at ScienceDirect



Simulation Modelling Practice and Theory

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



Adaptive wormhole routing in tori with faults: A mathematical approach

F. Safaei^{a,b,*}, A. Khonsari^{b,c}, M. Fathy^d, M. Ould-Khaoua^e

^a Department of Electrical and Computer Engineering, Shahid Beheshti University, Tehran, Iran

^b IPM, School of Computer Science, Tehran, Iran

^c Department of Electrical and Computer Engineering, University of Tehran, Tehran, Iran

^d Department of Computer Engineering, Iran University of Science and Technology, Tehran, Iran

^e Department of Electrical and Computer Engineering, Sultan Qaboos University, Oman

ARTICLE INFO

Article history: Available online 14 June 2009

Keywords: Interconnection networks Fault-tolerance Torus Routing algorithms Adaptive routing Message latency Virtual channels Queuing theory Performance evaluation

ABSTRACT

Fault-tolerance in a communication network is defined as the ability of the network to effectively utilize its redundancy in the presence of faulty components (i.e., nodes or links). New technologies of integration now enable the design of computing systems with hundreds and even thousands of independent processing elements which can cooperate on the solution of the same problem for a corresponding improvement in the execution time. However, as the number of processing units increases, concerns for reliability and continued operation of the system in the presence of failures must be addressed. Adaptive routing algorithms have been frequently suggested as a means of improving communication performance in large-scale massively parallel computers, Multiprocessors System-on-Chip (MP-SoCs), and peer-to-peer communication networks. Before such schemes can be successfully incorporated in networks, it is necessary to have a clear understanding of the factors which affect their performance potential. This paper proposes a novel analytical model to investigate the performance of five prominent adaptive routings in wormhole-switched 2-D tori fortified with an effective scheme suggested by Chalasani and Boppana [S. Chalasani, R.V. Boppana, Adaptive wormhole routing in tori with faults, IEE Proc. Comput. Digit. Tech. 42(6) (1995) 386–394], as an instance of a fault-tolerant method widely used in the literature to achieve high adaptivity and support inter-processor communications in parallel computers. Analytical approximations of the model are confirmed by comparing them with those obtained through simulation experiments.

© 2009 Published by Elsevier B.V.

1. Introduction

Computer networks comprise of a large number of technologies ranging from millimeters for on-chip networks such as processor-cache communication to the world spanning Internet. Interconnection networks traditionally belong to the smaller part of the range; from chip-to-chip communication to the system area networks, and in particular as the communication medium for multiprocessors. Interconnection networks offer communication with high reliability, high throughput, and low latency, all being vital factors for closely cooperating units. Interconnection networks are represented through technologies such as AutoNet [2], ServerNet [3], Myrinet [4], InfiniBand [5], RapidIO [6], PCI-Express AS [7], and HyperTransport [8]. Ten Gigabit Ethernet [9] with link-level backpressure is also emerging; however, this technology implements a "soft" backpressure which does not guarantee the absence of packet loss.

^{*} Corresponding author. Address: Department of Electrical and Computer Engineering, Shahid Beheshti University, Tehran, Iran.

E-mail addresses: f_safaei@sbu.ac.ir, safaei@ipm.ir (F. Safaei), ak@ipm.ir (A. Khonsari), mahfathy@iust.ac.ir (M. Fathy), mohamed@dcs.gla.ac.uk (M. Ould-Khaoua).

¹⁵⁶⁹⁻¹⁹⁰X/\$ - see front matter @ 2009 Published by Elsevier B.V. doi:10.1016/j.simpat.2009.06.005

An interconnection network is defined by its *topology, flow control*, and *routing*. The topology is the pattern of network node interconnection via physical communication channels. The torus topology has become a popular interconnection architecture for constructing massively parallel computers. Many parallel systems adopt low dimensional torus networks due to their low communication latency and high bandwidth [9,10].

Flow control deals with the allocation of channel and buffer resources to packets as they proceed through the network. Since processors in a parallel computer network need to communicate with the others, efficient communication is essential to enhance the performance of the system. The wormhole switching (also widely known as wormhole routing [9–11]) has been dominant for its low latency communication, and it has been adopted by most of the contemporary massively parallel machines. In wormhole switching, a message is divided into a sequence of fixed-size units of data, called *flits*. If a communication channel transmits the first flit of a message, it must transmit all the remaining flits of the same message before transmitting flits of another message. Wormhole switching only requires small buffers in the routers through which messages are routed. Also, it makes message latency largely insensitive to the message distance in the network. The main drawback of wormhole switching is that blocked messages remain in the network, therefore wasting the channel bandwidth and blocking other messages. In order to reduce the impact of message blocking, physical channels may be split into *virtual channels* by providing a separate buffer for each virtual channel and by multiplexing the physical channel bandwidth. The use of virtual channels can increase throughput considerably by dynamically sharing the physical bandwidth among several messages [9,10,12,13].

Routing and fault-tolerance in interconnection networks are issues belonging to the network layer of the OSI model [14]. The network layer deals with the end-to-end problem of moving packets from the source node to the final destination. The network layer has knowledge of the network topology and knows how to route the packets through the network. The routing algorithm is generally classified as being either *deterministic* or *adaptive* [9–11]. Deterministic routing is used in a variety of parallel computers because it is exceedingly simple and provides low latency and high bandwidth. However, deterministic routing has a number of significant disadvantages: poor performance under non-uniform traffic loads and poor fault-tolerance. On the other hand, adaptive routing provides alternative paths to route messages, thus avoiding congested regions in the network and increasing throughput.

Fault-tolerance is important to allow functionality of the interconnection network in the presence of faults. Preferably, faults and changes should be tolerated in run-time without disrupting network operations. If one routing node or one link fails, the rest of the network should not be forced into a halt. Preferably, they should be allowed continuous use of the interconnection network, suffering only from the degradation of performance presented by the fault. If a routing node or a link is inserted into the network, the network should incorporate the extra resource without requiring a stop in the communication between the communicating devices. Such hot-insertion of components simplifies incremental growth, which can also be considered as scalability issue. In a large sized network, it is therefore essential to design a fault-tolerant routing algorithm that can route messages in the presence of faults. Fault-tolerant routing for large-scale parallel computers has been the subject of extensive research in recent studies [1,14–22].

Most network performance evaluation studies have been conducted by means of software simulation [1,14–22]. Studying the relative performance merits of routing algorithms using simulation techniques is, however, limited by the excessive computation time required to run large simulations. Analytical modeling, in contrast, offers a cost-effective and versatile tool to carry out such a study, typically requiring a far lower computational load. This paper proposes a novel analytical model to assess the performance behavior of a number of prominent adaptive routing algorithms in wormhole-switched 2-D tori fortified with an efficient scheme suggested by Chalasani and Boppana [1], as an instance of routing methodology widely used in the literature to achieve high adaptivity and fault-tolerance capability in communication networks.

The remainder of the paper is organized as follows. Section 2 describes the context of this work, torus topology, and prominent adaptive wormhole routings, which are used throughout the paper. Our model assumptions as well as the performance modeling are presented in Section 3. Section 4 compares the message latency predicted by analytical model with those obtained through simulation experiments and finally, in Section 5, we summarize the results presented in the paper.

2. Preliminaries

This section explores the basis on which our analytical model is founded including the structure of a node in the underlying network, and routing schemes used in this study. The definitions in this section adhere to standard notation and definitions in wormhole-switched networks.

2.1. Torus network topology

Formally, a topology can be represented as a graph $\mathbb{T} = (\mathcal{N}, \mathscr{E})$ where the vertices \mathcal{N} are the routing nodes and the edges $\mathscr{E} \subseteq \mathcal{N} \times \mathcal{N}$ are the channels.

Definition 1. A link (channel) is a point-to-point communication medium, connecting two nodes, with buffers at each end. A link is understood to be full duplex.

Download English Version:

https://daneshyari.com/en/article/492091

Download Persian Version:

https://daneshyari.com/article/492091

Daneshyari.com