Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 69 (2015) 66 – 75

Procedia Computer Science

7th International Conference on Advances in Information Technology

Fault-tolerant routing in dual-cubes based on routing probabilities

Junsuk Park, Yuki Hirai, Keiichi Kaneko*

Graduate School of Engineering, Tokyo University of Agriculture and Technology, Koganei-shi, Tokyo 184-8588, Japan

Abstract

Recently, there has been a growing demand for large-scale computing in various fields. A binary hypercube has been used as an interconnection network topology in parallel computers for large-scale computing. However, the hypercube has a limitation that the degree, that is, the number of links for each node augments rapidly according to the increase in the number of total nodes. This limitation has been complemented in a dual-cube. In this paper, we propose a fault-tolerant routing algorithm in a dual-cube based on routing probabilities. The routing probability is developed as limited global information for routing ability of a node in a hypercube for an arbitrary node located at a specific distance. Unlike other previous researches that use limited global information in each cluster on fault tolerant routing in a dual-cube, our routing algorithm applies limited global information for an overall dual-cube. Each node selects a neighbor node to forward the message to the destination node by considering routing probabilities of its neighbor nodes. The results of a computer experiment show better performance of our algorithm.

Peer-review under responsibility of the Organizing Committee of IAIT2015

Keywords: Dual-cube; Fault tolerance; Interconnection networks; Routing probability

1. Introduction

A binary hypercube[1] was one of the most popular topology for interconnection networks and it was adopted in a wide variety of parallel systems such as the Connection Machine CM02, Intel iPSC, SGI Origin 2000 and the nCube[2,3] due to its properties of symmetric and recursive structures and a low diameter[4,5]. A hypercube network of a dimension n, Q_n, includes 2^n nodes. Each node has a unique n-bit binary address and has n edges. However, the hypercube has a limitation that the degree, that is, the number of links for each node augments rapidly according to the increase in the number of total nodes. Since the current IC technology has a limit of links per IC, the number of nodes in a parallel processing system based on a hypercube topology is restricted.

* Corresponding author. Tel.: +81-42-388-7489; fax: +81-42-385-9747.
E-mail address: k1kaneko@cc.tuat.ac.jp

Peer-review under responsibility of the Organizing Committee of IAIT2015
doi:10.1016/j.procs.2015.10.007

Li et al. have introduced a dual-cube as a new topology for interconnection networks. The dual-cube has reduced the degree in comparison with the hypercube while maintaining most of the topological properties of the hypercube[3]. Hence, the dual-cube can interconnect much more nodes in a parallel processing system compared with the hypercube of the same node degree. An $(r + 1)$-connected dual-cube, D_r, consists of 2^{r+1} hypercubes where each hypercube is called a cluster. There are two kinds of links: the links in a cluster are called *cube links* and the links between two different clusters are called *cross links*. Each node in a D_r has r cube links and only one cross link.

In a parallel processing system, multiple nodes are used together simultaneously, and they communicate with each other based on message passing. Therefore, how to route messages is a very important issue on performance of parallel processing systems. Furthermore, if the number of nodes will increase, the probability of occurrence of faulty nodes will also increase. Accordingly, for the better performance, finding a shorter path of routing between a source node and a destination node is an important issue in a parallel processing system with faulty nodes.

In general, an efficient fault-tolerant routing algorithm must satisfy the following two conditions. First, for a source node and a destination node, it must find a fault-free path. Second, each node should store small amount of information for time and space complexities[4]. Therefore, in this paper, we will use an approach based on a limited-global-information model[2,6,7,8].

There are three approaches: local-information-based, global-information-based, and limited-global-information-based to collect information regarding faulty nodes in a system. In a local-information-based approach, each node collects fault information of neighbor nodes only, and forwards a message to a destination node. Hence, only small amount of space complexity is used in this approach. However, the approach may induce many backtracks and long path lengths. On the other hand, in a global-information-based approach, each node collects fault information of all of the nodes in the system. Hence, it can forward a message along a shortest fault-free path to the destination. However, this approach would be impractical since it must collect global information and the process requires huge time and space complexities. A limited-global-information-based approach is an eclectic approach between a local-information-based one and a global-information-based one[2]. In the limited-global-information-based approach, each node collects abstracted faulty information of all of the nodes in the system[7], and forwards a message to the destination node along a quasi optimal path based on the information.

It is important to design the abstracted faulty information of all of the nodes in the system to work with the routing algorithm in the limited-global-information-based approach. If so, the routing algorithm based on the limited global information can generally achieve an optimal or a quasi optimal solution by requiring a simple process to collect and maintain information regarding faulty nodes of the neighbor nodes[9].

The rest of this paper is structured as follows. In Section 2, we survey the related works. Section 3 describes the dual-cube architecture and introduces requisite notations. Next, in Section 4, we introduce the routing probabilities in a hypercube with faulty links and show how to calculate them. Moreover, we propose our fault-tolerant routing algorithm in Section 5, and carry on a computer experiment in Section 6. Finally, Section 7 concludes this paper and presents some future works.

2. Related works

This section explains several researches related to the dual-cube after it has been introduced.

Li and Peng have introduced efficient routing and broadcasting algorithms in a non-faulty dual-cube. They have shown that the broadcasting in the dual-cube can be completed in optimal time under the restricted one-port communication model[10].

Li et al. have extended the broadcasting algorithm to all-to-all broadcast for collective communication[11]. In addition, they have shown that a fault-free Hamiltonian cycle can be constructed in a D_r with $(r - 1)$ faulty links[12,13].

Shih et al. were also involved in a study on Hamiltonian cycles in the dual-cube. They have proved that a D_r ($r \geq 2$) contains $(n + 1)$ mutually independent Hamiltonian cycles[5].

Jiang and Wu have proposed a fault-tolerant routing algorithm based on limited global information[2]. They have applied the safety level approach, which is based on limited global information for a hypercube, to each cluster of the dual-cube. Their routing algorithm makes use of the safety levels when the message is forwarded inside clusters.