Microelectronics Reliability 55 (2015) 679–695

Contents lists available at ScienceDirect

Microelectronics Reliability

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

Improving the reliability of the Benes network for use in large-scale systems

Mohsen Jahanshahi [a,*], Fathollah Bistouni [b]

[a] Young Researchers and Elite Club, Central Tehran Branch, Islamic Azad University, Tehran, Iran
[b] Young Researchers and Elite Club, Qazvin Branch, Islamic Azad University, Qazvin, Iran

ARTICLE INFO

Article history:
Received 10 September 2014
Received in revised form 11 November 2014
Accepted 12 December 2014
Available online 14 February 2015

Keywords:
Parallel computers
Multistage interconnection networks
Rearrangeable non-blocking MINs
Benes network
Large-scale systems
Reliability

ABSTRACT

Systems with high computational capabilities are usually made of a large number of processing elements in order to solve complex problems efficiently. To achieve this purpose, the processing elements must be able to communicate with each other, which can be provided by interconnection networks. Meanwhile, multistage interconnection networks (MINs) are often recommended for use in such systems due to the efficiency and cost-effectiveness. However, there is a fundamental problem in the general structure of these networks called blocking. An important class of MINs is rearrangeable non-blocking MINs that can be used as a cost-effective solution to cope with the blocking problem. Benes network is one of the most popular rearrangeable non-blocking MINs, which has been considered by many researchers over the years. And yet its reliability improvement is an important factor that must be considered in the review of most systems, especially large-scale ones. Based on previous works, there are three main approaches to improve the reliability of MINs: (1) Adding a number of stages to MIN. (2) Using multiple MINs in parallel. (3) Using replicated MINs. In this paper, in order to find the best solution to improve the reliability of the Benes network, all three of these methods are investigated. Reliability analyses show that the use of multiple parallel Benes networks can obtain more advantages than other methods for Benes network. The approach improves the Benes network to be used in large-scale systems in various aspects of reliability namely time-independent terminal reliability, time-independent broadcast reliability, time-dependent terminal reliability, and time-dependent broadcast reliability.

1. Introduction

The demand for systems with high computational power has always existed for many years and has not ever stopped. A number of important problems have been identified whose solutions require a tremendous amount of computational power that should be available in the future [1–3]. Instances of those problems include long-range weather prediction, nuclear reactor design, human genome, condensed matter theory, speech recognition, natural language processing, computer vision, image processing, and automated reasoning, to name a few. In the other hand, performance of processors has been doubled in approximately every three-year [4,5]. However, high computational power cannot be achieved solely by increasing processors' performance. That is, exploitation of other efficient means such as parallel processing techniques is very important to raise this capability. Nowadays, one of the main methods to improve the speed and computational power is the use of multiple processors in parallel computers. In fact, a parallel computer is a collection of processing elements that communicate and cooperate to solve large problems fast [6,7]. Therefore, parallel computers can be achieved by using parallel processing to speed up and gain more power to solve complex problems. Parallel processing is the use of multiple processors that are working simultaneously on the same problem. The hop is if a single processor can generate x floating point operations per second (FLOPS), then 100 of these may be able to produce $100x$ FLOPS, and 1000 processors may produce $1000x$ FLOPS [8].

Given the above discussion, several questions may arise. How do these processing elements communicate and cooperate? How is data transmitted among processors, what sort of interconnection is provided? In other words, a parallel computer requires some kinds of communication subsystem to interconnect processors, memories, disks, and other peripherals. The task of communicating between different nodes is the responsibility of the interconnection networks. An interconnection network is a complex connection of switching elements and links which allow communication

* Corresponding author.
E-mail addresses: mjahanshahi@iauctb.ac.ir (M. Jahanshahi), f_bistouni@qiau.ac.ir (F. Bistouni).

http://dx.doi.org/10.1016/j.microrel.2014.12.008
0026-2714/

between processors and memories. Therefore, the design of an efficient interconnection network is very critical for the construction of efficient parallel computers [4,7].

Interconnection networks can be divided into two main classes: Static and Dynamic. In a static network the connection between input and output nodes is fixed and cannot be changed and reconfigured. The examples of these types of networks are linear array, ring, tree, star, fat tree, mesh, tours, systolic arrays, and hypercube. Dynamic interconnection networks are implemented with switched channels, which are dynamically configured to meet the communication needs of user programs. In this paper, our focus is on a particular class of multistage interconnection networks (MINs) known as Benes network, which is a dynamic network. Now, the question that arises here is that why in this paper, a dynamic network is selected for the study? We have several reasons for this choice: (1) As mentioned previously, dynamic networks are more flexible in providing connections between nodes and can be reconfigured when it is needed. On the other hand, in many modern applications, the flexible and reconfigurable connections are very important. Therefore, dynamic networks are more useful in multi-processor systems compared with static networks such that these networks are known as multi-processor interconnection networks [9]. (2) As it will be discussed below, dynamic networks (i.e. bus, crossbar, and MINs) can provide a variety of options to satisfy both cost and performance requirements. In addition, MINs achieve a reasonable balance between cost and performance compared with bus and crossbar. (3) Dynamic networks have two important solutions to deal with the blocking problem in interconnection networks: use of non-blocking networks and use of rearrangeable non-blocking networks. Also, since use of the rearrangeable non-blocking networks is a cost-effective solution (because they require the minimum number of crosspoints to connect N terminal nodes in a rearrangeably non-blocking manner), this paper is focused on the Benes network, which is a well-known rearrangeable non-blocking network [10,11].

The dynamic networks include bus, crossbar, and multistage interconnection networks (MINs), which are often used in the multiprocessor systems [11,12]. But which of these topologies can be more suitable for use in large-scale systems? In the choice of an interconnection network for use in large-scale systems, two basic parameters should be considered: *cost* and *performance*. The crossbar networks are the most efficient but also the most expensive ones. On the other hand, shared bus networks have lower costs but also lower performance. MINs provide a compromise between the above networks because they provide efficient performance using a reasonable number of switches [12,13]. Therefore, MINs are often used in the context of single-instruction multiple-data (SIMD) and multiple-instruction multiple-data (MIMD) parallel machines and are also increasingly adopted for implementing the switching fabric of high-capacity communication processors, including asynchronous transfer mode (ATM) switches, gigabit Ethernet switches, and terabit routers [14,15].

MINs can be divided into three main classes: (1) *blocking MINs*, (2) *non-blocking MINs*, and (3) *rearrangeable non-blocking MINs* [11,16–18]. A MIN is said to be blocking if any free output may be unavailable to any free input because existing connections prevent a path from being established across the network. The Banyan-type MINs such as shuffle-exchange network [19], Baseline [20], and Generalized Cube [21] are blocking to traffic with a random destination distribution. A MIN that is always capable of connecting a free input to a free output, but which may require existing connections to be rearranged in order to do so, is called rearrangeable non-blocking MIN. The $(2n - 1)$ stage $N \times N$ shuffle-exchange network [22] ($n = \log_2 N$) and Benes network [23,24] are examples of rearrangeable non-blocking MINs. In addition, a MIN which is always capable of connecting a free input to a free output, regardless of the connection already established across the MIN, is said to be non-blocking MIN. The Clos network [25] is non-blocking MIN. The Clos network gets more performance compared with the other MINs [12] that this feature makes it attractive. However, these networks have two main problems that prevent them from being the practical solution. Generally, non-blocking MINs are more expensive than rearrangeable non-blocking ones and more complex to control [14,18,26,27]. It should be noted here that to estimate the cost of a MIN, one common method is to calculate the crosspoint cost. The crosspoint cost is given by the number of crosspoints within a switching element and by the number of switching elements within the network [18,27]. Therefore, since a non-blocking network requires more switching elements than a rearrangeable non-blocking network, it is not economical, especially for use in large-scale systems. Consequently, at present, the rearrangeable non-blocking MINs are commonly used in large-scale systems. Benes network is a rearrangeable non-blocking MIN that is known to many researchers in this field of research. On the other hand, one of the most important aspects of performance that should be considered in evaluation of most systems is reliability. An interconnection network should be able to deliver information reliably. Interconnection networks can be designed for continuous operation in the presence of a limited number of faults. In other words, in an interconnection network, reliability is a measure of how often the network correctly performs the task of delivering messages. In most situations, there is a need to deliver messages 100% of time without loss. For this reason, an unreliable interconnection network will have a direct negative impact on the overall performance of a parallel computer system. A parallel computer requires that its interconnection network operate without packet loss for ten thousands of hours [4,10]. Generally, reliability is the ability of a system to perform and maintain its functions in routine circumstances, as well as hostile or unexpected circumstances. Therefore, many researchers believe that it is the most important requirement for an effective MIN [12,13,19,28–33]. Hence, in this paper we are focused on improving the reliability of the Benes network.

While this is the case, improving the reliability of a MIN, until now, has been presented in several approaches; (1) *Adding a number of stages to a MIN or MINs*. (2) *Using multiple MINs in parallel*. (3) *Using replicated MINs*. Each of these approaches is described in details in the following section. Subsequently, according to the arguments already made, our motivation in this paper is to find the best solution to improve reliability of Benes network. If we meet the vital requirement of reliability, then it can be claimed that a good progress for improving MINs' performance in large-scale systems takes place.

The rest of the paper is organized as follows: A useful background will be presented in Section 2. Different reliability improvement approaches on the Benes network will be discussed in Section 3. In Section 4, reliability of these approaches will be analyzed, and finally some conclusions will be drawn in Section 5.

2. Background

In Section 2.1, the general structure of MINs will be discussed. Then, the blocking problem and structure of Benes network will be described in Sections 2.2 and 2.3, respectively. Discussions about related works will be done in Section 2.4. Finally, a useful argument on the contribution of the paper will be made in Section 2.5.

2.1. General structure of MINs

MINs are used for the interconnection of a set of N input terminal to N output terminals using sets of switches arranged in stages.