Fault-Tolerant Multiprocessor Networks Through an Extended G-Network

Jie Wu and Eduardo B. Fernandez

Department of Computer Engineering
Florida Atlantic University
Boca Raton, FL 33431

Abstract

Fault-tolerance is one of the key criteria in deciding the structures of interconnection networks for parallel/distributed systems. In this paper, we present an extended G-network which is 3-insensitive, meaning that the network can tolerate up to three communication links failures. It is shown that the extended G-network still keeps the characteristics of the G-network: Efficient routing, a small number of links, a small number of file servers, and fault tolerance. The performance of the extended G-network is compared to that of the G-network.

1. Introduction

Recent developments in technology have made it possible to interconnect a large number of computing elements. There are many different networks for parallel/distributed systems discussed in the literature [1]-[3]. which can be general or special purpose. The general purpose interconnections range from basic ring and star to more complicated hypercube, tree, torus, and mesh networks [4]. Special purpose interconnections are usually constructed for some special objectives, such as fault tolerance [5]-[7] or small diameter [8]. In this paper, we propose a specialized interconnection called the *extended G-network* with fault tolerance as the objective.

In general, a system is made fault-tolerant by providing redundant or spare processors, and/or by providing redundant communication links. If a fault in a particular component (link or processing element) is to be tolerated, then the system must have redundant or spare facilities to take over the duties of the faulty component. In case of running out of spare facilities the system should support graceful degradation and therefore prolong system inertia. The design of a fault-tolerant network is usually based on graph theory [6], [15], where a graph G = (V, E) is used where V represents the node set and E represents the edge set. An element from V denotes a processing element, and an element from E denotes a communication link.

The G-network [9] is a novel interconnection design which is a suitable architecture for point-to-point communication. This network can tolerate up to two link faults (it is 2-insensitive) and has better performance than some other interconnection structures. In this paper, an extended G-network is discussed which can tolerate up to three link faults. By comparing the performance of the extended G-network with that of the G-network, it is shown that the extended G-network still keeps the characteristics of efficient routing and small number of links, but it also has a better fault tolerance than the G-network.

In section 2, some definitions are given. The general properties of the extended G-network are discussed in Section 3. A comparison between the performance of the extended G-network with that of the G-network is made in Section 4. Finally in Section 5, conclusions and possible further extensions of the G-network are discussed.

2. The G-Network and An Extended G-network

For certain network interconnections it is desirable to provide a strong core group which can directly communicate with nodes not in the core graph. It is also desirable that the interconnections keep these properties when some components fail. To design network interconnections to meet these objectives, we first need some concepts from [9], [10].

(1) A subset of nodes $D \subset V$ is a *dominating set* for a graph G(V, E) if every node of G is either in D or is adjacent to some node of D. The *domination number* $\partial(G)$ is the minimum size of all the dominating sets.

The practical value of the dominating set can be seen from the modeling of a set of file servers for network-based distributed systems [11]. Let us consider a set of processing elements (workstations), which is connected by some kind of network, with its graph model as discussed in Section 1. The workstations need to share resources, maybe because of economical reasons, or because of the nature of the application. Management of shared resources is an important service that should be provided by a trusted authority to meet reliability and security requirements. One method is to use server machines called file servers to administrate the shared resources and support applications running on the workstations (the various characteristics of file servers and their corresponding implementation issues are of no interest here). We can make correspond an element from the dominating set to a file server, then each file server can communicate directly with all the workstations it serves.

(2) A graph G is *n-insensitive* if the number of nodes needed to dominate G is unchanged when any n edges are removed.

CH2728-4/90/0000/0216$01.00 © 1990 IEEE

This definition characterizes fault tolerant properties of the networks modeled by this type of graphs. It is very restricted since it requires the domination number to remain unchanged in case of failure of n communication links. In general, a set of file servers selected may not constitute a minimum size of dominating set. It is of more important that the set of file servers can continue service in case of failure of n links, and this can be achieved through reconfiguration. Some file servers can migrate to other nodes (workstations). Each file server can still directly communicate with all the workstations it serves. But it is preferable that the number of file servers remain unchanged. We use here a less restricted n-insensitive definition: A graph G is n-insensitive if the number of nodes in the selected dominating set remains unchanged when any n edges are removed. Comparing with the previous definition, the size of the selected dominating set of a graph G in this new definition may be greater than $\partial(G)$.

Definition 1 [9]: A G-network is constructed using the following steps: First, select r nodes to be file server nodes and label them $a_1, a_2, \ldots, a_r$. Then between each pair (a_i, a_j), $i \neq j$, add two nodes adjacent to both of the special nodes. Label the degree two nodes b_i, $1 \leq i \leq r^2 - r$.

As it is shown in [9], the G-network has $N = r^2$ nodes and $E = 2N - 2r$ links where $r = \partial(G)$. The G-network is 2-insensitive, and its other properties are illustrated in [9], [12]. Figure 1 shows the G-network when r = 3.

Definition 2: An extended G-network can be constructed in the following way:

For a file server set $(a_1, a_2, \ldots, a_r)$, add three nodes between each pair (a_i, a_j), $i \neq j$. Three nodes should all be adjacent to both a_i and a_j, and there are also two edges among these three nodes.

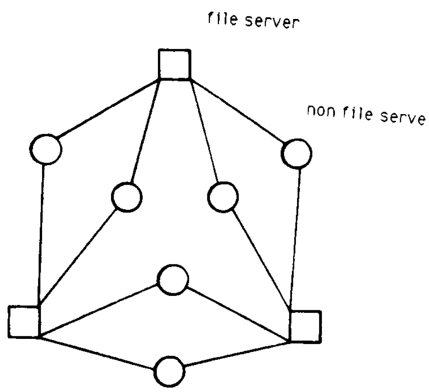


Figure 1 The G-network when r=3

Figure 2 shows the extended G-network when r = 3, and Figure 3 shows the extended G-network when r = 4.

It is easy to see that the extended G-network has $N = 3r(r-1)/2 + r = (3r - r)/2$ nodes and $E = 4r(r-1)$ links.

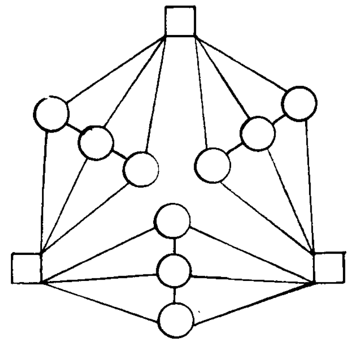


Figure 2 The extended G-network when r=3

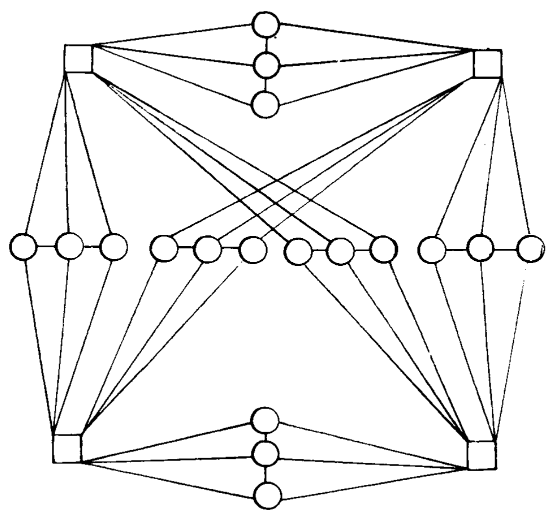


Figure 3 The extended G-network when r=4

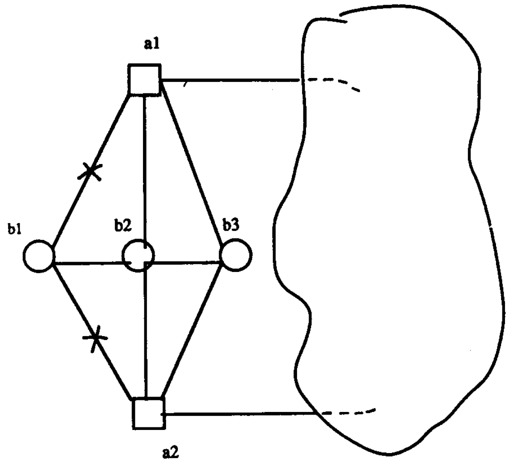
3. Properties of the Extended G-network

The extended G-network still keeps some properties from the G-network. One of them is defined by Theorem 2 from [9]: One file server or non file server node in the extended G-network can fail without disrupting service to the remaining active nodes. We will discuss the properties of the extended G-network which differ from that of the G-network. The following are two theorems about the fault tolerance and fast routing properties of the extended G-network.

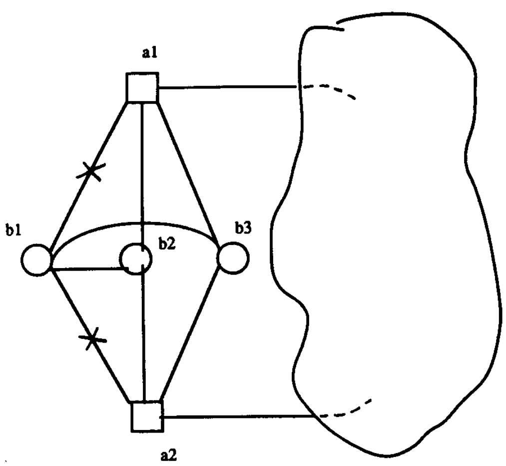
Theorem 1: The extended G-network is 3-insensitive.

Proof of Theorem 1:

Arbitrarily remove three edges from the extended G-network; say, e_1, e_2 and e_3. If no two of these three edges are two edges which connect a non file server node with its two adjacent file servers, then every non file server node has at least one edge to a file server. Hence, the r file server nodes still dominate the graph. If two of these three edges are such that they connect a non file server node with its two adjacent file servers, then there are two cases as shown in Figure 4. Assume in this figure that a_1b_1 and a_2b_1 are two of these edges.



Case 1



Case 2

Figure 4 Two cases of faulty edge distribution

Clearly for case 1: If the third edge is either a_1b_2, a_2b_2, or b_2b_3, then (b_1, b_3) and the set of file server nodes minus (a_1, a_2) dominate G - $\{e_1, e_2, e_3\}$. If the third edge is not a_1b_2, a_2b_2, or b_2b_3, then (b_1, b_2) and the set of file server nodes minus (a_1, a_2) dominate G - $\{e_1, e_2, e_3\}$.

For case 2: If the third edge is either a_1b_2 or a_2b_2, then (b_1, b_3) and the set of file server nodes minus (a_1, a_2) dominate G- $\{e_1, e_2, e_3\}$. If the third edge is not a_1b_2 or a_2b_2, then (b_1, b_2) and the set of file server nodes minus (a_1, a_2) dominate G- $\{e_1, e_2, e_3\}$.

In all the above cases, the size of the dominating set remains unchanged. So the extended G-network is 3-insensitive.

Theorem 2: The maximum number of routing steps (hops) required between any two nodes in the extended G-network is four. Such maximum number will remain four if one edge fails. In case of two or three edge failures, the maximum number is five for non isolated nodes.

Proof of Theorem 2:

For the case of no faults and one fault, the theorem can be proved following the method in [9].

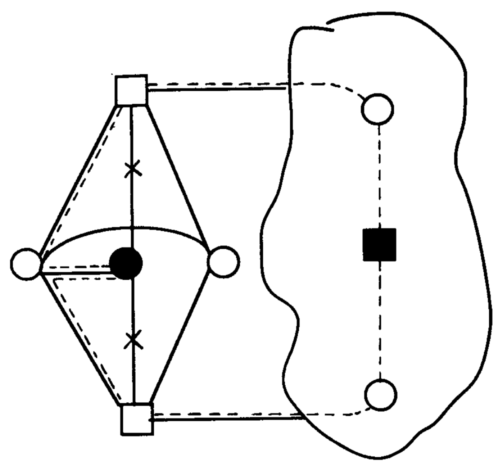
For the case of two faults and three faults, we have the following observations:

(1) The maximum number of routing steps between two file server nodes in case of two faults is two, and in case of three faults is three.

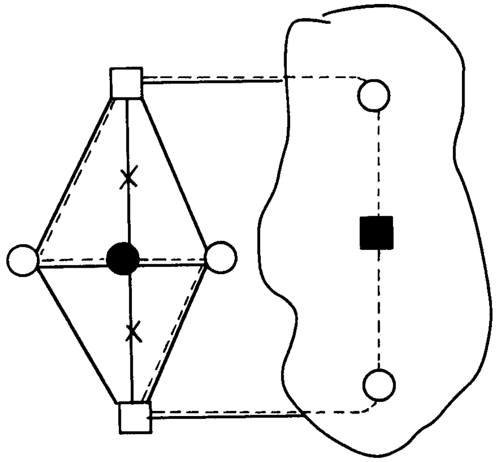
(2) The maximum number of routing steps between one file server node and one non file server node in case of two or three faults is four. Figure 5 shows such worst cases. Note that, since there exist at least two disjoint paths, we do not indicate the location of the third edge fault in case of three faults.

(3) The maximum number of routing steps between two non file server nodes in case of two or three faults is five. Figure 6 shows such worst cases.

From the above observations, we can prove that the maximum number of routing steps is four for the no fault case, and is five for two faults or three faults.



case 1



case 2

Figure 5 Two cases of fault distribution in a routing between a file server and a non file server

4. Cost and Performance Evaluation

When proposing a new interconnection, we should study its cost and performance as compared to some existing interconnections. Although there are no consensus on which properties or features should be selected to compare different interconnections [13], we use here four commonly used properties to compare the extended G-network to the G-network.

The first criterion is their fault tolerance properties which are measured by n-insensitiveness. It has been shown that the extended G-network is 3-insensitive and has a better fault-tolerance (for communication links failures) than the G-network.

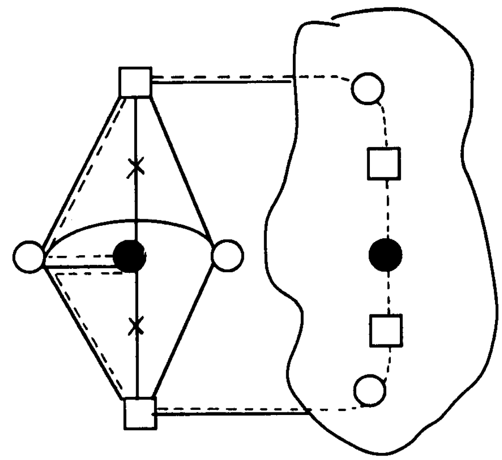
The extended G-network also has a more efficient routing than the G-network. In case of two or three edge faults, Theorem 2 shows that the maximum number of routing steps is five for non isolated nodes, while it is six for non isolated nodes in the G-network [9].

The next two properties are cost factors to evaluate the networks. The first cost factor is small number of links, which can be measured by the ratio of the number of links to the number of nodes. This ratio for network-i can be denoted as R1(network). Clearly ,

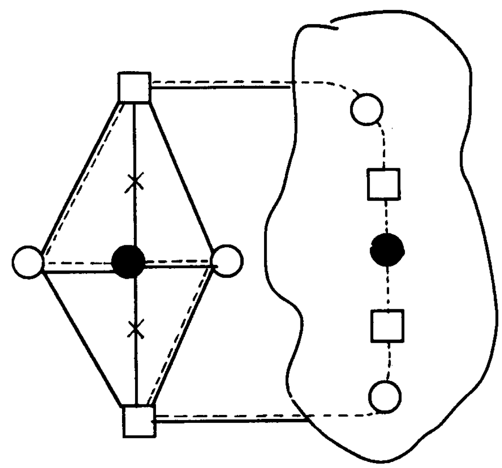
$$R1 (\text{G-network}) = 2 (r-1) / r$$
$$R1(\text{extended G-network}) = 8 (r-1) / (3r - 1)$$

When r becomes infinite,

$$\lim_{r \to \infty} R1(\text{G-network}) = 2$$
$$\lim_{r \to \infty} R2 (\text{extended G-network}) = 8/3$$



case 1



case 2

Figure 6 Two cases of fault distribution in a routing between two file server nodes

The extended G-network needs a few more links, but still it is much better than the barrel shifter [14].

The last property considered is small number of file servers. This can be measured by the ratio of the number of file servers to the number of nodes. We denote this ratio as R2 (network). It is easy to get the following results:

$$R2 (\text{G-network}) = 1/ r$$
$$R2 (\text{extended G-network}) = 2 / (3r - 1)$$

So the extended G-network has a better property of small number of file servers than the G-network, even though r file servers in the G-network form a minimum dominating set for the graph, while r file servers in the extended G-network do not form a minimum dominating set.

5. Conclusions

We have proposed an extended G-network. In addition to keeping the four characteristics of the G-network, the proposed network has a better fault tolerance property and a faster routing in case of failures of three edges fault than the G-network. Other extensions are under investigation based on different objectives. One extension [12], called multi-layered G-network, is obtained by interconnecting copies of the G-network in parallel, and is suitable for large networks for parallel computation. Since this structure is capable of expansion in such a way that it causes a small disruption of the existing set up, the multi-layered G-network has better expansion capability than normal G-network. Another possible way of extension is the following: Given r file server nodes, then add new nodes among every three file servers instead of between every two file servers. More study is need to show the feasibility of such an extension. The problem again is cost and effectiveness of the interconnection.

Ackonwledgments

This work is supported by a grant from the State of Florida High Tech and Industry Council.

References

[1] Bhuyan, L. N., "Interconnection networks for parallel and distributed processing," *Computer*, p. 9, June 1987.

[2] Pradhan, D. K., *Fault Tolerant Computing: Theory and Techniques*. Vol. i and Vol. ii, Prentice-Hall, Englewood Cliffs, 1986.

[3] Yanney, R. M. and J. P. Hayes, "Distributed recovery in fault tolerant multiprocessor networks," *IEEE Transaction on Computers*, 35, 10, p. 871, Oct. 1986.

[4] Reed, D. A. and R. M. Fujimoto, *Multicomputer Networks: Message-Based Parallel Processing*, The MIT Press, 1987.

[5] Akers, S. B. and B. Krishnamurthy, "On group graphs and their fault-tolerance," *IEEE Transaction on Computers*, 36, 7, p. 885, July 1987.

[6] Hayes, J. P., "A graph model for fault tolerant computer systems," *IEEE Transaction on Computers*, 25, 9, p.876, Sept. 1976.

[7] Pradhan, D. K., "Fault tolerant multiprocessor link and bus network," *IEEE Transaction on Computers*, 34, 1, p. 35, Jan. 1985.

[8] Chudnovsky, D. V., *et al.*, "Regular graphs with small diameter as models for interconnection networks," *Proc. of Third Intl. Conf. on Supercomputing*, Vol. iii, p. 232, 1988.

[9] Rice, T. H., R. K. Guha, R. C. Brigham, and R. D. Dutton, "The G-network and its inherent fault tolerant properties," Technical Report, Dept. of Computer Science. University of Central Florida, 1988.

[10] Dutton, R. D. and Brigham, R. C., "An extremal problem for edge domination insensitive graphs," *Discrete Applied Mathematics*, 20, p. 113, 1988.

[11] Svobodova, L., "File servers for network-based distributed systems," *Computer Survey*, 16, 4, p. 454, Dec. 1986.

[12] Guha, R. K. and T. H. Rice, "Some remarks on K-r-Insensitive graphs in network system design," to be published, 1989.

[13] Agrawal, D. P. and V. K. Janakiram, "Evaluating the performance of multicomputer configurations," *Computer*, p. 23, May 1986.

[14] Hwang, K. and Briggs, F. *Computer Architecture and Parallel Processing*. McGraw-Hill. 1984.

[15] Agrawal, D. P., "Graph theoretical analysis and design of multistage interconnection networks," *IEEE Transaction on Computers*, 32, *7, p. 63,* July 1983.