Fault Tolerance of BSN and HCN

Turki F. Al-Somani and Mostafa Abd-El-Barr

Computer Engineering Department, King Fahd University of Petroleum & Minerals, Dhahran 31261, Saudi Arabia, \{tsomani, mostafa\}@ccse.kfupm.edu.sa

Abstract. This paper presents two recent hierarchical interconnection networks, the Block Shift Network (BSN) and the Hierarchical Cubic Network (HCN). Then a comparison between them has been done in terms of reliability and fault tolerance issues such as the fault diameter and the 2-terminal reliability. The comparison between them is depending on different factors such as the topological properties and different configurations and parameterizations of BSN. The result shows that HCN is favorable in some situations and BSN in some others.

Keywords: BSN, HCN, HIN, Fault Tolerance, Reliability

1. Introduction

In non-shared memory multiprocessors systems, interconnecting processors will be either directly, which is costly for large number of processors i.e. $O(N^2)$, or indirectly by routing via intermediate processors. Three important factors should be considered at the design stage for the indirect connection: minimizing the message delay, keeping the cost low and maximizing the reliability. This is because indirect connection has four effects on the system performance:

1) An extra delay will be added in data transmission because of intermediate stages.
2) Each processor must have the ability to perform this routing intelligently.
3) There could be a larger queuing delay in each processor’s output queue, because there is more traffic load in the system due to message relaying.
4) The system reliability becomes a problem.

The hypercube network [2] is very popular for parallel computing systems. Combining several networks has been used to construct new networks. Recent examples are Hyperbanyan [3], Hyper-deBruijn [4] and Banyan-hypercube [5]. The basic idea is to synthesize a network from simple building blocks in an incremental fashion. The hypercube networks have been also used as the interconnection network in a number of commercial and experimental parallel computers. A variety of interconnection networks based on the hypercube have also been proposed [12], [13], [14], [15], [16], [17], [18], [19]. For large-scale systems, the number of links for the hypercube may become prohibitively large.

Ghose and Desai presented a new interconnection network called the Hierarchical Cubic Network (HCN) [10], [20], [21]. The HCN uses almost half as many links as a comparable hypercube and yet has a smaller diameter than a comparable hypercube and emulates desirable properties of a hypercube very efficiently. Yi Pan also proposed another new interconnection network called the
Block-Shift Network (BSN) [1], [6], [7], [9]. The BSN has been designed to eliminate the drawbacks of the hypercube network while retaining its advantages. Because many application algorithms use the links on lower dimensions extensively, connections on lower dimensions are provided in BSN.

This work compares the reliability and fault tolerance aspects of BSN and HCN such as the fault diameter and the terminal reliability. Comparisons between them using different parameters are included showing the performance of BSN and HCN which is better than the hypercube as in the literature [9][11] in many aspects and it is very promising for massively parallel computers [1][7]. The comparison also includes the topological cost of BSN and HCN which is expressed by using the nodes degree and network diameter with different network sizes. Section 2 will introduce the BSN connection types, construction and topological properties. Section 3 introduces the HCN construction, degree and diameter. The topological cost, the fault diameter and the reliability of BSN and HCN are addressed in section 4. Section 5 will discuss the results of both networks. Finally, the conclusion and future work will be presented in section 6.

2. Block-Shift Network

This section presents different BSN connection types, edge groups, degree and diameter and the construction of BSN.

2.1 BSN Connection Types

The connections between two processors are called the connections on dimensions two if the two processor element (PE) differ only in bits from \( i \) to \( j \) and are connected by the links on these dimensions. For example, the link between PE 000 and 010 is the dimension-1 connection while the link between PE 0010 and 0100 is the dimension-2 connection. There are several variations for the connections on dimensions \( i \) to \( j \); this paper considers the following connection methods:

1) Concurrent-Connection Method: If two processors whose addresses differ only in positions \( i \) to \( j \) are directly connected. That is, in a concurrent-connection method, bits \( i \) to \( j \) in one address can be changed in 1 step to reach another address. When \( i=0 \) and \( j=n-1 \) (covering all dimensions), then the connection scheme corresponds to the fully connected topology.

2) Sequential-Connection Method: If two processors whose addresses differ only in positions \( i \) to \( j \) can reach each other by changing bits \( i \) to \( j \) of their addresses one by one. Thus, for 2 processors with addresses differing in all bits from \( i \) to \( j \), one needs unit-routes to send a message from one to the other. If only 1 bit is different in bits \( i \) to \( j \), then 1 step is needed. When \( i=0 \) and \( j=n-1 \), then the connection scheme corresponds to a hypercube topology.

3) Partial-Connection Method: Between the previous two extremes (method 1 &2), other methods can be defined. For example, define a connection method which can change the section (bits \( i \) to \( j \)) 2 bits at a time, 3 bits at a time, and so on. Assume that the section (bits \( i \) to \( j \)) has \( b \) bits, and can be divided into \( b/a \) subsections. A partial-connection can change a whole subsection into any pattern in 1 step by modifying 1 bit, 2 bits, … , or \( a \) bits in the subsection.
2.2 BSN Edges Groups

BSN consists of 3 groups of edges:

Group #1: It connects nodes to their counterparts with addresses shifted cyclically positions left in 1 step; *i.e.*, it connects the processor at address, \(a_{n-1}a_{n-2}...a_1a_0\) to the processor at, \(a_{n-b-1}a_{n-b-2}...a_1a_0a_{n-1}...a_{n-b}\). These connections are called L-Shift links and the data transfers over these links are L-Shift operations. For example, if \(b=3\) then PE 0010011 must be connected to PE 0011001.

Group #2: It connects nodes to their counterparts with addresses shifted cyclically positions right in 1 step; *i.e.*, it connects the processor at address, \(a_{n-1}a_{n-2}...a_1a_0\) to the processor at, \(a_{b-1}a_{b-2}...a_1a_0a_{n-1}...a_{n-b}\). These connections are called R-Shift links and the data transfers over these links are R-Shift operations. For example, if \(b=3\) then PE 1010011 must be connected to PE 0111010.

Group #3: It contains the connections over the rightmost \(b\) dimensions. The links in group #3 are called R-change links; and the data transfers over these links are R-change operations. For example, if \(b=3\) then PE 0010000 could be connected to PE 0010001, PE 0010010, PE 0010011, PE 0010100, PE 0010101, PE 0010110 or PE 0010111.

2.3 BSN Construction

A block in BSN is defined as the nodes which are connected by the links in group #3 (R-change links). BSN has \(N=2^n\) nodes and in each step only \(a\) bits can be changed within the section of the rightmost \(b\) bits, and is labeled BSN(\(a, b\)). Changing these two parameters defines the network connection type as follows:

1. If \(a=1\) then the network (*i.e.*, BSN(1, \(b\))) has a sequential connection method and each block is a hypercube with \(2^b\) nodes as shown in figure 1.
2. If \(a=b\) then the network (*i.e.*, BSN(\(b, b\))) has a concurrent connection method over the last dimensions and contains \(2^{n-b}\) blocks each with \(2^b\) nodes. Within each block is a complete graph (thin lines), and blocks are connected by either R-Shift or L-Shift links (thick lines which are double link), as shown in figure 2.
3. If \(1 < a < b\) then the network has a partial connection method.

The loops in Figure 1 and 2 are R-Shift or L-Shift links which happen to connect the processors themselves. The BSN is a hierarchical structure with nodes connected tightly within blocks and with blocks connected loosely. This property matches the communication requirements of most parallel application algorithms. The design of the BSN is motivated by the fact that although the hypercube is useful for many parallel algorithms, it lacks flexibility and costs too much when it is large. On the other hand, BSN is flexible because its parameters can be changed to meet the performance & cost requirements. The BSN is scalable in the sense that changing the size of the network does not require changing the hardware within the nodes. Many existing networks are special cases of the BSN (within a block). For example,

1. BSN(1, 1) is the shuffle-exchange network.
2. BSN(\(l, n\)) is the \(n\)-dimensional hypercube.
3. BSN(n, n) is the complete network. Thus, the BSN study is useful for comparing the performance of these networks.

![Figure 1: A BSN(1, 2) with N=16](image1)

![Figure 2: A BSN(2, 2) with N=16](image2)

2.4 BSN Degree and Diameter

In BSN(a, b) of N=2^n nodes, the network degree (the maximum number of ports per node) is (2^n - 1)*(b/a) + 2. Note that the network degree in BSN depends only on the parameters a and b regardless to the total number of nodes. This cost measure compares the BSN favorably with the hypercube because BSN has a constant degree once a and b are fixed, while the degree of the hypercube increases with its number of nodes.

The diameter (the maximum shortest path, with distinct hops, between any two nodes) of BSN(a, b) of N=2^n nodes, with nodes and degree (2^n - 1)*(b/a) + 2 is (1+(b/a))*[n/b]. Note that as b increases, the diameter of a BSN decreases, and its number of links increases. For more details on the performance, routing algorithms, parallel algorithms on BSN, see [1] and [6].
3. Hierarchical Cubic Network

This section presents HCN degree and diameter and the construction of HCN.

3.1 HCN Construction

An HCN\((n, n)\) is a hierarchical network consisting of \(2^n\) clusters, each of which is an \(n\)-dimensional hypercube. Each node of the HCN\((n, n)\) is addressed by a pair of numbers \((I, J)\), where \(I\) is an \(n\)-bit cluster number and \(J\) is an \(n\)-bit address of the node within a cluster. Each node in the HCN\((n, n)\) has \((n + 1)\) links connected to it. The links within a hypercube cluster are referred to as local links and the links between two clusters are referred to as external links. Clusters are interconnected by using external links to construct the HCN\((n, n)\) using the following rule:

- If \(I' \neq J\), a node \((I, J)\) is connected to the node \((J, I)\) using its external link, which is called nondiameter link.
- A node \((I, I)\) is connected to the node \((I', I')\), where \(I'\) is the bitwise complement of \(I\), using its external link, which is called diameter link.

An HCN\((2, 2)\) and HCN\((3, 3)\) network are shown in Fig. 3 and Fig 4 respectively.

![Figure 3: A HCN(2, 2) network](image)

3.2 HCN Degree and Diameter

In HCN\((n, n)\), the network degree (the maximum number of ports per node) is \(n + 1\) since \(n\) is within the same cube and one for the external link. The HCN\((n, n)\) diameter (the maximum shortest path, with distinct hops, between any two nodes) were computed in [22] and shown to be equal to \(n + \left\lfloor \frac{n + 1}{3} \right\rfloor + 1\), which is a correction to the one provided by [11].
4. Fault Tolerance & Reliability

This section presents the fault tolerance and reliability of BSN and HCN. This includes the topological cost, fault diameter and reliability.

4.1 Topological Cost

In general, the cost of a network can be expressed by using the nodes degree and diameter of the network. The networks with small degrees have large diameters, and on contrary, networks with small diameters have large degrees. It appears that there is a tradeoff between the degrees and the diameters of networks. So that, there have been a commonly used measure for the cost of a network which computes the product of diameter and degree. Table 1 shows the cost of BSN & HCN and Table 2 & 3 shows the cost when \( n=16 \) for both HCN and BSN with different parameter.

Table 1: Network Cost of BSN and HCN

<table>
<thead>
<tr>
<th>BSN((a,b))</th>
<th>Nodes</th>
<th>Degree</th>
<th>Diameter</th>
<th>Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>(2^n)</td>
<td>((2^a-1)*(b/a)+2)</td>
<td>((1+(b/a))*(n/b))</td>
<td>(\left\lfloor n/b \right\rfloor * 2^a/b + 2^2b^2 + ab - b^2 + 2a^2)</td>
<td></td>
</tr>
<tr>
<td>HCN((n,n))</td>
<td>(2^{2n})</td>
<td>(n+1)</td>
<td>(n + \left\lfloor \frac{n+1}{3} \right\rfloor + 1)</td>
<td>((n+1)*(n + \left\lfloor \frac{n+1}{3} \right\rfloor + 1))</td>
</tr>
</tbody>
</table>

Table 2: Topological Cost BSN\((b, b)\) and HCN\((n, n)\) with \( N = 65536 \)

<table>
<thead>
<tr>
<th>HCN((8, 8))</th>
<th>BSN((2,2))</th>
<th>BSN((4,4))</th>
<th>BSN((8,8))</th>
</tr>
</thead>
<tbody>
<tr>
<td>108</td>
<td>80</td>
<td>88</td>
<td>524</td>
</tr>
</tbody>
</table>
Table 3: Topological Cost BSN(1, b) and HCN(n, n) with \( N = 65536 \)

<table>
<thead>
<tr>
<th>HCN(8,8)</th>
<th>BSN(1,2)</th>
<th>BSN(1,4)</th>
<th>BSN(1,8)</th>
</tr>
</thead>
<tbody>
<tr>
<td>108</td>
<td>160</td>
<td>248</td>
<td>436</td>
</tr>
</tbody>
</table>

For the same number of nodes \( N = 65536 \), it is clear from the results that the cost of BSN varies according to the values of \( a \) and \( b \). While it remains constant for HCN all the time since it only depends on \( n \). However, it is clear that for small numbers of \( a=b \) for BSN, the cost is smaller than HCN (see Table 2). But for \( a=1 \), the cost is always larger than HCN. The last column in both Table 2 & 3 gives a huge fast increase in the network cost of BSN compared to HCN which makes HCN favorable in such situations.

4.2 Fault Diameter

Fault diameter is an important aspect for fault tolerance in a multiprocessor system. A fault diameter of a graph is defined as the diameter of a new graph generated after the faulty nodes & links are removed from the old graph. An \( f \)-fault diameter of a graph is defined to be the maximum of distances over all possible graphs that can occur with at most \( f \) faults [12]. The BSN fault diameter [9] is as follows:

- The \((b−1)\)-fault diameter of a BSN(1, b) is \( 2[n/b] + 2(b + 1)b[n/b] \).
- The \((2^b − 2)\)-fault diameter of a BSN(b, b) is \( 6[n/b] \).

The HCN fault diameter is shown in [23] to be \( n + \left\lfloor \frac{n+1}{3} \right\rfloor + 5 \).

4.3 Reliability

Reliability analysis could be intractable if many paths between any 2 nodes can have one or more links in common. Therefore, deriving a lower bound on 2-terminal reliability (also referred to as path reliability) by considering a subset of all available paths between two nodes in the two networks offers an important insight into the value of 2-terminal reliability of a network. As long as the lower bound is quite tight, it can be used to estimate the 2-terminal reliability of a network. The 2-terminal reliability between \( s \) and \( t \) is defined as the probability of finding a path entirely composed of operational links between them.

The 2-terminal reliability of BSN(b, b) is shown in [9] to be:

\[
TR_{BSN(b,b)} = 1 - (1 - p^{4m})^{2^b - 1}, \ m \equiv \left\lfloor n/b \right\rfloor
\]

Similarly, the 2-terminal reliability of BSN(1, b) is:

\[
TR_{BSN(1,b)} = 1 - (1 - p^{2m(b+1)})^b, \ m \equiv \left\lfloor n/b \right\rfloor
\]

The 2-terminal reliability of HCN(n, n) could be calculated as:

\[
TR_{HCN(n,n)} = 1 - (1 - p^n)^{n+1}
\]
where we have totally $n+1$ node disjoint paths as shown in [23] for the HCN($n$, $n$).

5. Results

Table 4 shows the values for the 2-terminal reliability of the BSN and HCN of moderate size of the same number of nodes ($N = 65536$) for various $a = b$ and Table 5 shows the BSN and HCN for various $a = 1$. The link operational rate of changes from 0.90 to 0.98. Figure 5 & 6 demonstrate the results of Table 4 & 5 respectively.

Table 4: 2-terminal reliability of BSN($b$, $b$) and HCN($8$, 8)

<table>
<thead>
<tr>
<th>$P$</th>
<th>HCN($8$, 8)</th>
<th>BSN(2,2)</th>
<th>BSN(4,4)</th>
<th>BSN(8,8)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.90</td>
<td>0.994</td>
<td>0.099</td>
<td>0.953</td>
<td>1.000</td>
</tr>
<tr>
<td>0.92</td>
<td>0.998</td>
<td>0.194</td>
<td>0.989</td>
<td>1.000</td>
</tr>
<tr>
<td>0.94</td>
<td>0.999</td>
<td>0.359</td>
<td>0.999</td>
<td>1.000</td>
</tr>
<tr>
<td>0.96</td>
<td>1.000</td>
<td>0.612</td>
<td>0.999</td>
<td>1.000</td>
</tr>
<tr>
<td>0.98</td>
<td>1.000</td>
<td>0.892</td>
<td>1.000</td>
<td>1.000</td>
</tr>
</tbody>
</table>

Table 5: 2-terminal reliability of BSN(1, $b$) and HCN($8$, 8)

<table>
<thead>
<tr>
<th>$P$</th>
<th>HCN($8$, 8)</th>
<th>BSN(1,2)</th>
<th>BSN(1,4)</th>
<th>BSN(1,8)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.90</td>
<td>0.994</td>
<td>0.012</td>
<td>0.057</td>
<td>0.166</td>
</tr>
<tr>
<td>0.92</td>
<td>0.998</td>
<td>0.036</td>
<td>0.134</td>
<td>0.334</td>
</tr>
<tr>
<td>0.94</td>
<td>0.999</td>
<td>0.099</td>
<td>0.296</td>
<td>0.598</td>
</tr>
<tr>
<td>0.96</td>
<td>1.000</td>
<td>0.262</td>
<td>0.580</td>
<td>0.876</td>
</tr>
<tr>
<td>0.98</td>
<td>1.000</td>
<td>0.614</td>
<td>0.905</td>
<td>0.994</td>
</tr>
</tbody>
</table>

Figure 5: Reliability of HCN(8, 8) and BSN($b$, $b$)
Figure 6: Reliability of HCN(8, 8) and BSN(1, b)

The total number of links of BSN(b, b) is equal to $2^b \times (2^b + 1)$ [9] while the total number of links of HCN(n, n) is found in [11] to be $2^{2n-1} \times (n+1)$. Table 6 shows the product of the reliability multiplied by the number of links for HCN(8, 8) and BSN(b, b) and Figure 7 illustrates these results.

Table 6: Reliability times the number of network links

<table>
<thead>
<tr>
<th>$p$</th>
<th>HCN(8,8)</th>
<th>BSN(2,2)</th>
<th>BSN(4,4)</th>
<th>BSN(8,8)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.9</td>
<td>265420.8</td>
<td>294912</td>
<td>884736</td>
<td>15158477</td>
</tr>
<tr>
<td>0.92</td>
<td>271319.04</td>
<td>301465.6</td>
<td>904396.8</td>
<td>15495332</td>
</tr>
<tr>
<td>0.94</td>
<td>277217.28</td>
<td>308019.2</td>
<td>924057.6</td>
<td>15832187</td>
</tr>
<tr>
<td>0.96</td>
<td>283115.52</td>
<td>314572.8</td>
<td>943718.4</td>
<td>16169042</td>
</tr>
<tr>
<td>0.98</td>
<td>289013.76</td>
<td>321126.4</td>
<td>963379.2</td>
<td>16505897</td>
</tr>
</tbody>
</table>

Figure 7: Reliability X No. of Links

For the same number of nodes ($N = 65536$), the results of Table 4 show that the reliability of the BSN(b, b) depends on $b$ while the reliability of HCN depends only on $n$. So, it is obvious that the reliability of the BSN(b, b) goes better than hypercube for $b$ changing from $b=2$ to $b=8$. This is at the expense that it uses more links than the HCN as shown in Table 6. In Table 5, BSN(1, b) has the least reliability because it uses the least number of links. The reliability and number of link product shown in Figure 7 shows that HCN has the least results among different BSN(b, b).
6. Conclusion & Future Work

BSN as a recent proposed topology seems to be interesting because of its flexibility and scalability. HCN is also a recent proposed network that uses half number of links of the hypercube per node with the same performance as a comparable hypercube. Another interesting advantage of the HCN is the ability to physically partition a hypercube into sub-cubes which is useful in many systems where the system and interconnection is shared among multiple users.

In this paper, reliability and fault tolerance issues of BSN and HCN have been demonstrated in a smooth manner. The construction details of BSN and HCN and their topological properties helped a lot realizing their properties. The results show comparable and competing results between BSN and HCN in terms of reliability and number of links per node.

Future work will include an exhaustive study of other hierarchical networks. The results of such a study could draw a clear picture about the advantages and disadvantages of using them. Also, an emphasize study will be done on generalized BSN and incomplete HCN.

References