S. Iyengar, N. M. Singhi, and S. Sanyal

Fault-Tolerance in Hypercube-based Parallel Architecture

Proceedings of the First International Workshop on Parallel Processing (IWPP '94), Bangalore, 26-31 Dec., 1994, pp. 430-434.

Proceedings of the First International Workshop on Parallel Processing (IWPP '94) held in Bangalore, 26 - 31 Dec., 1994, pp 430 - 434.

FAULT-TOLERANCE IN HYPERCUBE-BASED PARALLEL ARCHITECTURES

IYENGAR SHRINIWAS
Department of Computer Technology,
Victoria Jubilee, Technical Institute,
V.J.T.I. Road, Matunga, Bombay-19, India
E-mail:shriniwas@vjti.ernet.in

N.M. SINGHI
School of Mathematics,
Tata Institute of Fundamental Research,
Homi Bhabha Road, Colaba, Bombay 05, India
E-mail:singhi@tifrvax.ernet.in

S. SANYAL
Computer Systems and Communications Group,
Tata Institute of Fundamental Research,
Homi Bhabha Road, Colaba, Bombay-05, India
E-mail:sanyal@tifrvax.ernet.in

Abstract

Fault-tolerance in Hypercube-based parallel computers is discussed. Methods for estimating and identifying fault-tolerant partitions in Hypercubes are also discussed. These problems are converted into equivalent representations in extremal sets theory and it is illustrated as to how such a conversion lends itself to intuitive solutions to the problems more easily. The identification of such fault-tolerant partitions has implications in the implementation of gracefully-degradable parallel architectures as well in the implementation of cube-sharing/allocation algorithms in faulty hypercubes.

1 Introduction

Given the modern-day over-dependence on computers and the consequent demand for processing-speed and fail-safe operation, parallel architectures incorporate fault-tolerance more as a rule than as an exception.
Hypercubes are one of the most popular structures adapted for the construction of parallel architectures because of their versatility. Several researchers have studied the properties and fault-tolerant behaviour of hypercube-based parallel computers in [6],[7],[8],[9] and [10]. The excellent topological properties of the Hypercube [11] can be exploited for providing fault-tolerance. Hypercubes also offer asymptotically maximal probability for reliability (1 connected component) amongst all graphs with the same number of vertices and edges [4] (c.f.[5]).
The rest of this paper is organized as follows: in Section I , we discuss basic aspects of tolerant partitions , in Section II, we outline the problems that are studied by us, in Section III we review existing results and present our methods and results. Details and proofs will appear in [1]. We conclude the paper with our observations and a list of useful references

2 Section I

2.1 *m-Partitions of an n-cube*

A Hypercube H^n of n dimensions has 2^n nodes with each node being associated with an n-dimensional vector $v \in \{0,1\}^n$. v(i) denotes the ith co-ordinate of v , $0 \leq i \leq (n-1)$. Let $I \equiv \{0,1, \ldots, n-1\}$ .We can make a one-one correspondence between n-vectors and subsets B of I using the mapping Ψ defined as follows:
For $v \in H^n$, $\Psi(v) = \{i \in I \mid v(i) = 1\}$

For example: the finite-set associated with the 4-vector (1010) is {1,3}.

Thus we can identify the n-cube with P(I) i.e the power-set of I. Any node of the n-cube can then be associated with a subset of I.

Let $M \subseteq I$ such that $|M| = m$ and let $A \subseteq I$, $M \cap A = \phi$. We can easily see that under the above identification, an m-subcube of the n-cube H^n corresponds to the family of sets $T(M,A) = \{B \cup A \mid B \subseteq M\}$. The m-partition of an n-cube is defined as the set of all m-subcubes of the n-cube that agree in exactly (n-m) dimensions (or co-ordinates). From above, it clearly corresponds to $T(M) = \{T(M,A) \mid A \subseteq I, M \cap A = \phi\}$. Thus, for a given m, there are ${}^nC_m$ possible m-partitions of the n-cube and each of these will contain 2^{n-m} m-subcubes, with every m-subcube having 2^m nodes. Also, in every m-partition, there will be exactly one m-subcube that includes the node with the label corresponding to the all-zeroes n-vector $0 \equiv \Phi$; we shall call this m-subcube M_0. Note that $M_0 \equiv M$.

2.2 Simplicial Complex

In the context of finite-sets, a simplicial complex S on I is a family of subsets of I that satisfies the following property:

Property 2.1: If K is an element of S, then every subset of K is also an element of S.
Clearly, n-cube H^n corresponds to the trivial simplicial complex P(I). Similarly the subcube M_0 corresponds to a simplicial complex P(M).

2.3 Other notations and terms

We define floor(a) = a, if a is an integer, and
= a-1 if a is not an integer.

$\text{ball}(n,r) = \sum_0^r {}^nC_i$ i.e. the number of all n-vectors of Hamming weight $\leq r$.
$\text{Ball}(n,r) = \{A \mid A \subseteq I, |A| \leq r\}$ i.e. the collection of all sets of cardinality $\leq r$.
$\text{rad}(n,t) = \max\{r \mid \text{ball}(n,r) \leq t\}$.
$\text{rem}(n,t) = t - \text{ball}(n,r)$.

2.4 Conditions for tolerant partitions

Let F denote a finite set whose elements are subsets of I corresponding to faulty nodes of the n-cube-based parallel architecture. We assume that faulty nodes can neither perform calculations nor can they route information, that all faults are known beforehand and that these faults remain static during the course of their analysis.

An m-partition is said to tolerate the fault-set F (or is F-good) if the distribution of the faults in F is such that the induced sub-graph on healthy nodes in every m-subcube of the m-partition has a connected-component with atleast $2^{m-1}+1$ healthy nodes. Such components will be called Large Healthy Components of the cube H^n. It can be easily seen that under such conditions the large healthy connected-components of the m-subcubes in that partition collectively form a single large connected-component of atleast $2^{n-1}+1$ nodes (see *Remark 2.1* below). An m-partition of the n-cube is said to be F-bad if the distribution of faults in F is such that there exists in the m-partition atleast one m-subcube that has no connected-component of size more than 2^{m-1} healthy nodes.
The following remark has also been observed in [2].

Remark 2.1: Let T(M) be an F-good m-partition and let $M \subseteq M'$, $|M'| = |M|+1$, then the (m+1)-partition M' is also F-good.
Proof: Consider any two m-subcubes of the F-good m-partition which are neighbours. Clearly, we can find a pair of nodes one from a large healthy component in each of these m-subcubes which have an edge connecting them. From this the remark follows.

One of the basic tools in [1] and [2] is a well-known result of Kleitman [3]:
Kleitman's Condition: If the number of faults in an m-subcube is atmost ${}^mC_{\text{floor}(m/2)}$, then the subcube has a large healthy connected-component.

3 Section II

3.1 Problems under consideration

Consider an n-dimensional Hypercube-based parallel computer with a known fault-set F consisting of f arbitrarily placed faulty nodes (i.e. $|F| = f$). F can be determined by using methods suggested in [10]. Following the method outlined in [2], we wish to partition the faulty hypercube into small sub-cubes such that each of these subcubes include a small number of tolerable faults from the original set of faults, F. Then, the majority of the fault-free nodes in such subcubes form a connected-component that spans the hypercube and can be used to implement various hypercube algorithms with a constant factor slowdown (see [2]).

The main problems we consider are:

1. Find m such that the m-partition of the n-cube will tolerate the given set of F faulty nodes. Because of *Remark 2.1* we should try and find the smallest such m. Smaller m-partitions also imply more efficient implementations of regular and single-port algorithms on faulty n-cubes [2].

2. For the smallest m described in 1 above, find an F-good m-partition.

The problems 1 and 2, in this generality are quite hard. A simpler, related problem is the following:

3. For a given m, find a fault-set F_{min} of the smallest size such that every m-partition is F-bad.

It is expected that for any fault-set F, $|F| < |F_{min}|$, the structure of F_{min} will help in identifying an F-good m-partition. Problem 3 has mainly been studied using Kleitman's condition giving rise to the case $c={}^{m}C_{floor(m/2)}$ of the following closely-related problem:

4. For given m and c, find a simplicial complex S on I with the smallest possible size |S| such that for every m-subset M, $|S \cap P(M)| \geq c$

4 Section III

4.1 *Review of earlier work*

In [2], several interesting methods to study problem 3 were developed. We re-state some of the results obtained in [2] and use the same as reference for deducing and for comparing our results.

Theorem 4.1: (c.f. *Theorem 3.9* in [2])
For all $n \geq m \geq 0$, given any set F of fewer than $\phi(n,m)$ faulty nodes in an n-cube, there exists an m-partition of the n-cube which tolerates F.
Here,

$$\phi(n,m) = \text{bal}(n,r) + \frac{{}^{n}C_{m} \cdot \text{rem}(n,x)}{{}^{n-r-1}C_{m-r-1}}$$

where,

$x = {}^{m}C_{floor(m/2)}$ and $r = rad(m,x)$.

4.2 Transformation of problem to extremal sets domain

In [2], a method is given to transform the original fault-set F into an equivalent set that has certain desirable properties.

Transformation T(i):
Let $F \subseteq P(I)$ and $i \in I$.
Define T(i) [F] to be the family $F_i \subseteq P(I)$ where F_i is defined as follows:

$$F_i = \{ A \mid A-\{i\} \in F, A \in F \} \cup \{ A-\{i\} \mid A \in F, A-\{i\} \notin F \}$$

Let T[F] denote the family obtained by applying transformation T(i) to F in a sequential manner from 0 through (n-1) for the n-dimensional case.

We note that: $|T(i)[F]| = |F| = |T[F]|$.

Remark 4.1 For every F , T[F] is a simplicial complex and if F is a simplicial complex, T[F] = F.
As an example, consider the fault-set F to be:
{(000101),(111101),(100111),(100101),(010101) }
which corresponds to the family
F = { {0,2}, {0,2,3,4,5}, {0,1,2,5}, {0,2,5}, {0,2,4} }.

It can be checked that:

$$T[F] \equiv \{\phi, \{4,5\}, \{1\}, \{5\}, \{4\}\}$$

Thus *transformation T* transforms F into:
T[F] ≡ {(000000), (110000), (000010), (100000), (010000) }

The following is essentially *Lemma 3.7* of [2] (see also [1]) :

Theorem 4.2: Let F be a fault-set of H^n and let $M \subseteq I$ such that |M|=m. Then, a subcube T(M,A) has atleast

$^mC_{floor(m/2)}$ nodes from F iff simplicial complex T[F] is such that $|T[F] \cap P(M)| \ge {}^mC_{floor(m/2)}$.

Using *Kleitman's Condition* and applying *Theorem 4.2* to every m-subset, one can see a close relationship between problems 3 and 4 stated in **Section II** above.

4.2 Equivalent and Stronger Results

We now state results proved in [1]. Details will appear in [1].

Let S be a family of subsets of I. We will say that S satisfies property p(c,k,m) if:
$|S \cap P(M)| \ge c$ for all m-subsets M of I and for every B $\in$ S, $|B| \ge k$
Note : p(c,0,m) is the property considered in problem 4.

Let t(c,k,m) be the minimum size of S such that $S \subseteq P(I)$, and S satisfies p(c,k,m).
Now, using *Theorem 4.2,* one can easily see that *Theorem 4.1* essentially gives the following inequality:

$$t(x,0,m) \ \ge \ \phi(n,m) \ \ge \ b(n,r)$$
where, $x = {}^mC_{floor(m/2)}$ and r = rad(m,x).

Using the same argument as in [2] to prove *Theorem 4.1,* one can see more generally:

Theorem4.3: $t(c,0,m) \ \ge \ \phi(n,c,m) \ \ge \ b(n,r)$ with r=rad(m,c), where

$$\phi(n,c,m) = \ ball(n,r) + \frac{{}^nC_m \cdot rem(n,c)}{{}^{n-r-1}C_{m-r-1}}$$

However, we have a much stronger result (see [1] for details), which gives a structure on the smallest fault-set F such that every m-partition is F-bad:

Theorem 4.4: There exists $S \subseteq P(I)$ such that S satisfies p(c,0,m) ,|S| = t(c,0,m) and $S \supseteq Ball(n,r)$, r = rad(m,c).
Corollary 4.1: t(c,0,m) = b(n,r) + t(c-b(n,r), r+1, m)

Remark: The above corollary suggests that F-good m-partitions may be found iteratively.

Now, let $S \subseteq P(I)$ be such that it satisfies p(c-b(n,r), r+1, m), where r = rad(m,c) .It can be easily observed (see [1] for details) that :

$$ {}^{n-r-1}C_{m-r-1} \cdot |S| \ge \sum_{M \subseteq I, |M|=m} |S \cap P(M)| \ge {}^nC_m \cdot (c - b(n,r))$$

Thus,

$$|S| \ \ge \ \frac{{}^nC_m \cdot (c - b(n,r))}{{}^{n-r-1}C_{m-r-1}}$$

Further, equality holds iff S satisfies the property:

Property 4.2: $B \in S \Rightarrow |B| = r+1$, and
$|S \cap P(M)| = c - b(n,r)$ for all m-subsets M of I

4.2 The use of t-designs
Property 4.2 is closely related to a mathematical structure called t-designs, studied extensively and used in Coding Theory, interconnection networks etc.. (for details see [11]). More precisely,

Lemma 4.1: S satisfies *Property 4.2* iff there exists an (n-m)-(n, n-r-1, c-b(n,r)) design D and S is given as:

S = {I-B | B is a block of D}

Thus we have proved,

Theorem 4.5: If $S \subseteq P(I)$ satisfies p(c-b(n,r), r+1, m) , r=rad(m,c) , then

$$|S| \ \ge \ \frac{{}^nC_m \cdot (c - b(n,r))}{{}^{n-r-1}C_{m-r-1}}$$

and equality holds iff there exists an (n-m)-(n, n-r-1, c-b(n,r)) design D and S is given as:
S = {I-B | B is a block of D}
Using *Corollary 4.1*, we get the following: