

RELIABILITY PREDICTION OF PARSYS HYPERCUBE ARCHITECTURE

INFORMATICA 2/89

Keywords: parallel architecture, routing reliability, tours, hypercube

Rihard Piskar
Iskra Delta, Ljubljana

Abstract

The paper presents the combinatorial reliability model of a parallel computer PARSYS. The architectural design is based on 64 processor system with 64 memory modules, that are connected via network of routing nodes. The network is called 6-cube because 2^6 is the number of processors. At first there is derived the combinatorial reliability model of 6-cube architecture for 64 processors and memories connected in the hypercube. Then the reliability of the rerouting procedure is evaluated as the probability of the successful packet transfer from a processor to a memory. Further research showed that torus architecture is the enhancement of the packet transfer reliability because of redundant routing capabilities.

Introduction

PARSYS is MIMD parallel processor research project of Iskra Delta Computers Company. The architectural design is based upon 64 processor system with 64 memory modules, that are connected via network of routing units. Each memory module has the capacity from 2 Mbyte to 8 Mbyte and is divided to the local memory and global memory. Each routing unit supports fast routing and incorporates functions and logic that avoid or minimize degradations mostly encountered in the multiprocessor environments [1].

The routing network architecture design has passed several mayor development stages, the most important of them being hypercube and torus architecture. Theoretical studies of network topology had raised some doubts regarding high dimensional networks which require more and longer wires than medium dimensional networks. For that reason, binary N-cube was chosen with 2^N nodes in hypercube architecture with $N=6$. This is a special case of the family of m-arry N-cubes with N dimensions and m nodes in each dimension [2]. Latter on the transformation was done on torus architecture preserving hypercube functional characteristics and obtaining constant wire density. Torus is 8-arry 2-cube in comparison to hypercube 2-arry 6-cube with the same number of nodes, $n=64$ [3]. Separately to routing performance analysis the reliability analysis was done bringing some interesting cognitions.

The reliability analysis started first with hypercube architecture. As the evolution of the project proceeded, latter on torus reliability analysis showed even better reliability performances.

Modeling

Hypercube. The hypercube architecture of the parallel processor PARSYS is the essential actor of accounting the global memory to the processors. All communications between processors and the memory are passing at least one routing-node. If a routing-node fails, then the packet passing it, will not reach the destination.

To estimate the probability of that random event, there is necessary to have a probability model of the architecture, to make mathematical expression and to find out the relation between architecture and reliability. The probability that the packet transfers predefined number of routing-nodes, is combinatorial expressed and depends on the architecture. The architecture enables the rerouting in a case of failure in a routing-node. In that case the packet is sent via another path again. The model is expressing that possibility, but it was never realized because of great changements of the routing hardware.

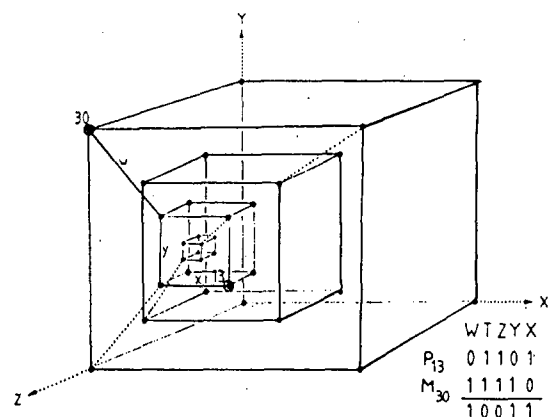


Fig.1.
The principle of hypercube architecture,
the example of 6-cube.

From Fig.1. there are seen two addresses which differ in three bits, so that destination is three edges distant from the origin.

Let be in the system $n = 2^N$ processors and memories connected in N-cube, where N is the number of system levels. Let be the probability, that a processor want to get data from a memory, or to send changed data to the memory, uniformly distributed. Let this operation be called the packet transfer. Suppose, that connecting units fail with a failure rate π due to the Poisson law.

In the system N-cube each processor is connected to n memories of the appropriate level. The number of levels is N, so that the number of processors and memories is $n = 2^N$.

From each node there are N possible connections, in our case six. The addresses of possible connections are different in Hamming distance 1. Each node is containing the processor, the memory and the routing-node. We are interested in the probability, that the processor and the memory are distant in Hamming distance H. Suppose that the address consists of N bits and that k pairs of bits is different. Then $H = N - k$. In fact we are looking for the probability, that the addresses are distant for distance H or, that their addresses have k different pairs of bits.

This probability is:

$$p(k) = \frac{\binom{N}{k}}{\sum_{k=0}^N \binom{N}{k}} \quad (1)$$

The combinatorial probability of the successful packet transfer from the processor to the memory in other words the reliability of packet transfer is:

$$R(t) = \frac{1}{N} \sum_{k=0}^{N-1} [1 - p(k)q(k)] \quad (2)$$

where $p(k)$ is the probability of k different pairs of bits in the processor and the memory addresses, $q(k)$ is the unreliability of k routing nodes:

$$q(k) = 1 - \text{Exp}[-k\pi t] \quad \text{for hypercube without,} \quad (3)$$

$$q(k) = [1 - \text{Exp}(-k\pi t)]^2 \quad (4)$$

for hypercube with rerouting, where π is the failure rate of single routing-node.

Equ.(3) is straight forward solution of the probability, that k routing nodes does not function correctly for time t. Equ.(4) involves some redundancy of packet transfer routes. Let us suppose the possibility of the rerouting in a case of a fault in the packet transfer because of a physical failure in a unit, where the packet is transferred over. The packet is unable to reach the destination because of the failure in a routing-node. The rerouting is realized by the selection of another path.

Torus. The evolution of the parallel system proceeded to another architecture with better characteristics (less and shorter wires per board). This is 2-dimensional 8-arry torus with the same number of processors 64. The communication between processors and memories located at each node start from a node in four possible directions; north,south,west and east. In a case of a failure in one of nodes the alternative path is established. In the communication protocol two directions are provided, each time the one with more available routs to avoid the saturation.

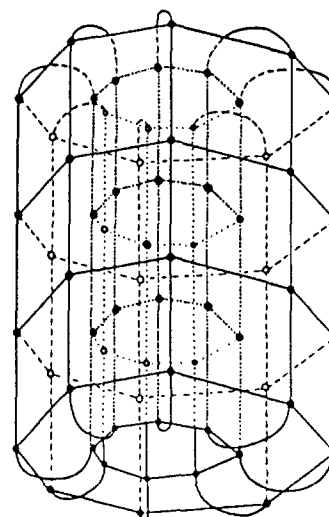


Fig.2.

Torus archyecture of 64 routing-nodes connecting 64 processors and memories.

There is a relation between N-dimensional m-aries of torus and hypercube. In 6-dimensional 2-arry hypercube there is 6 different possibilities of combining the number of nodes in one packet transaction and 64 in total. In 2-dimensional 8-arry torus there exist 8 different possibilities and 256 in total. The sum of probabilities of an established route with its reliability leads to greater reliability of torus than hypercube architecture. The reliability of randomly chosen single communication route for a packet transfer equals to Equ.(2) using (3) and of redundant route which is normal in torus architecture, equals to Equ.(2) using (4).

Reliability evaluation results

From the evolution point of view the analysis of different dimensions of hypercube and torus architecture is obvious. There must be the evidence of reliability versus dimension N and the number of processors $n=2^N$. Upon the combinatorial model the comparison of the reliability showed the following results that are summarized on the table 1.

Table 1.

The unreliabilities of packet transfer for hypercube and torus architecture for constant product of node failure rate and time $\pi t = 0.1$

	TORUS	HYPERCUBE
Dimension	2-dim, 4-arry	4-dim, 2-arry
No. of processors	16	16
Max. no. of nodes	4	4
Single unreliability	0.031	0.031
Redundant unreliability	0.0010	0.0010
Dimension	2-dim, 8-arry	6-dim, 2-arry
No. of processors	64	64
Max. no. of nodes	8	6
Single unreliability	0.041	0.043
Redundant unreliability	0.0017	0.0018
Dimension	2-dim, 16-arry	8-dim, 2-arry
No. of processors	265	265
Max. no. of nodes	16	8
Single unreliability	0.034	0.041
Redundant unreliability	0.0011	0.0017

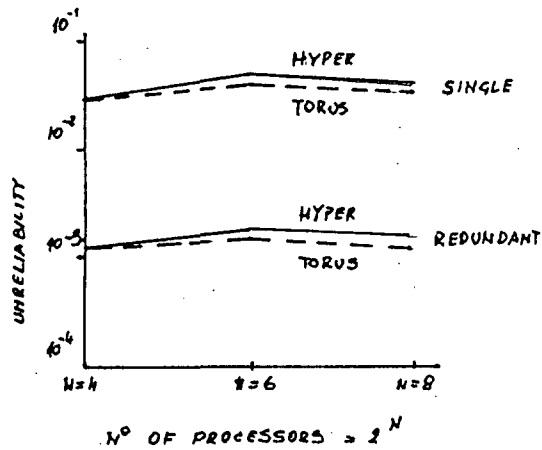


Fig.3.

Unreliability of hypercube and torus architecture for constant product of failure rate and time $\lambda t = 0.1$.

The differences between hypercube and torus architecture are not essential, if we look only single routing path. At $N=6$ there is not more than 5% difference in the unreliability of hypercube and torus. More significant is the difference between single hypercube and redundant torus. Hypercube is 36 times more unreliable than torus. It is true that the redundant hypercube is only 6% less reliable than torus, but to achieve the redundant routing algorithm in hypercube it is demanded to make huge changes in hardware. At the contrary, torus architecture contains the alternative algorithm to decide which direction may be taken. The rerouting is accomplished, when the route fails because of a failure in a routing node. Pragmatical meaning of better torus reliability in failures per thousand hours is that instead of $4 \cdot 10^{-6}$ failures there occur only 10^{-7} failures per randomly chosen packet transfer operation.

System availability

Another aspect of system performance is the availability as the function of failure and repair rate of routing-nodes. The failures of nodes are consequences of mutually independent random events. The natural result of this fact is, the architecture has no influence to the availability.

The technique used to evaluate system availability is the method of multidimensional Markov system, described in Ref.[4]. Briefly explained, the method consists of the allocation of system failure states in a multidimensional space in such a way, that the relation between the number of states S and the dimension d is $S=2^d$. Vice versa relation is $d=\log_2 S$.

From that relation with $S=64$ and $d=6$ the system named d -cube represents a failure state in each cube vertex. The numerical availability evaluation of the system was done by VMS program tool written in Fortran language. The inputs are failure and repair rates of nodes and dimension parameters. The outputs are system availability and MTBF. With 44.5 failures per million hours and 2 repairs per hour the availability is 0.9989 and MTBF is 468 hours.

Communication cost. The system PARSYS consists of 64 processor/memory boards folded into torus configuration. Communication cost is the function of the ratio between processor speed and the speed of routing-node as well as the function of dimension. It was found [3], that torus communication cost is quickly increasing over dimension $N=6$.

Hardware design. The processor board is VME/386 processor with communication node and dynamic 2M to 8M Byt RAM. The packet transfer logic is the most important part of the node. Each packet routine logic consists of mechanism for independent selection of a packet in the direction of more paths to desired node.

Software design. Software task partitioning is based on explicit parallel constructs for two level parallelism; macro or user specified and micro with iterative structures embedded into macro parallelism. A very fast micro task has been developed based on the fact that a processor can interrupt any processor. The synchronization is a simple get-and-link swap mechanism [3].

Conclusion

The reliability evaluation showed some advantages of hypercube architecture. One of them is relatively low average number of routing-nodes in randomly chosen packet transfer operation. There is also a low level of the unreliability of packet transfer. The reliability of the hypercube architecture can be even increased with possible rerouting, but with essential change of deterministic routing algorithm. Better solution due to general better characteristics of torus architecture is to implement redundant capabilities of torus architecture. The combinatorial model can be used in further reliability analysis of parallel systems to find out the optimum number of processors to trade-off the communication cost versus the reliability.

References

- [1]S. Prešern, "A Selected Survey of Parallel Computer Systems", Journal of Computing and Informatics, Vol.11, No.4, pp.40-53, YU ISSN 0350-5596, 1987.
- [2]P. Brajak, "Designing Reconfigurable Intelligent Memory Module Performance Enhancement to Large Scale General Purpose Parallel Processor", Journal of Computing and Informatics, Vol.11, No.1, pp.19-53, YU ISSN 0350-5596, 1987., 1987.
- [3]P. Brajak, S. Prešern, L. Vogel, A.P. Zeleznikar, P. Hitič, "Scheduling and Synchronization Concepts on the PARSYS Parallel Processor", Proc. on ISMM International Conference, Florida, 1988.
- [4]R. Piskar, J. Virant, "Hardware-Software Availability of a System", Proc. of 6th int. conf. on reliab. and maint. Strasbourg 1988.