FACTA UNIVERSITATIS (NIŠ)
SERIES: ELECTRONICS AND ENERGETICS VOL. 6, NO. 1, (1993), 63-67

A FAULT TOLERANT SCALABLE PARALLEL COMPUTER WITH A RECONFIGURABLE COMMUNICATION STRUCTURE

R. Miederer and Wolfgang Weber

Abstract: The paper presents a new concept of parallel computer architecture which provides additionally to an array structure of equally structured basic modules a second layer of an online observation and controlling network. Each basic module assembles a group of four processing units, their control unit, four dual ported RAM and all communication channels to the neighboor units and to the observation network. The proposed method permits as well a reconfiguration in case of hardware faults as - with the help of a new distributed operating system - a new solution of load balancing which determines free computing capacities under the aspect of processor- memory- and communications channel - loads, avoiding high amounts of communication.

Key words: Computer, fault tolerant computer, parallel computer architecture, reconfigurable structure.

1. Introduction

The existing parallel computer architectures-mainly the massively parallel structures - have been discussed by many authors [1,2,4,6]. These as well as the medium grain parallel computers are restricted in their scalabilty, i.e. capacity and performance cannot be deliberately increased by additional processor items in the array of already existing processors. This restriction results from the bussystems which are usually communication channels between the processors or between a processor and the RAM. Adding new processing units to the array results in a bottle-neck in the communication channels. And even before reaching these restrictions the adding of new components and therefore increase in probability of hardware failures cannot be handled by known methods.

Manuscript received September 14, 1993.

The authors are with Lehrstuhl für Datenverarbeitetung, Ruhr–Universität Bochum, Universitätsstrasse 150. 44780 Bochum, BRD.

This paper proposes a parallel computer which can be upgraded step by step by adding system elements. The system is additionally enabled by its structure to detect faults, to isolate and to analyze defective hardware devices and to generate adapted new configured communication structures. The detection of faults or defects does not involve the online operation of the system.

2. Hardware concept and design: The RESCUE System (reconfigurable and scalable parallel computer)

Fig. 1 explains the structure of a basic module and a connection with the connecting network and control unit layer. The lower layer shows one of the basic processing units constructed by four processing units (PU), each connected by its own communication channels to the upper switching network layer. The switching network itself can be connected by socalled worker netlines to the other modules. A shared-memory (SM) is used by the control unit (CU) to observe the four PU during normal operation. Active control lines between PU and CU serve as error channels for status observation in case of interrupts. The CU can by a configuration line alternate the communication structure of the switching network and start the analysis of eventually defect PU stations.

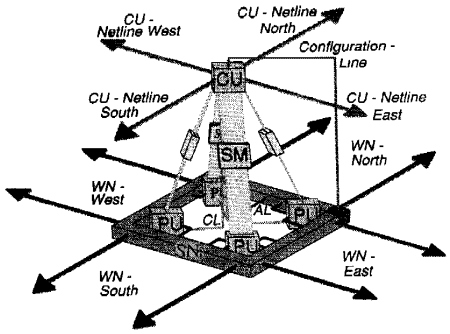

Fig. 1. Basic module of the RESCUE-system.

To underline the reconfigurability Fig. 2 shows a twodimensional top view with the four PU, connected by SM with the CU and the connected switching network of PU (N-E-W-S). This PU network (N-E-W-S net) has already proved by former research [3], [7] as more economical as, e.g., the hypercubes. The observing properties of the CU via SM can so be used to

control a single PU. Interpreting the informations in SM allows the CU to detect a defect PU and to arrange a reconfiguration isolating the defect unit and furtheron to inform the communication partners (Fig. 3).

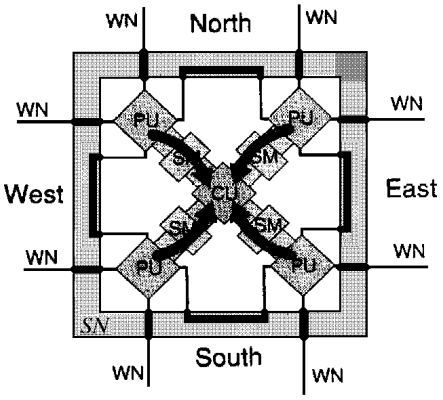

Fig. 2. Connection of the basic Module with CU-level.

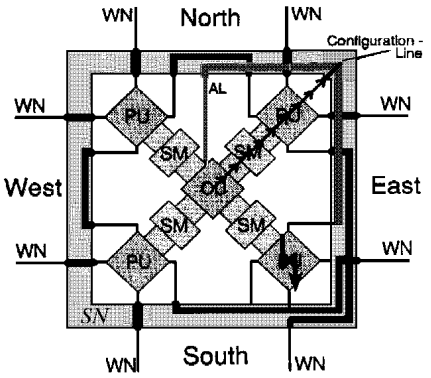

Fig. 3. Isolation of a defect PU and reconfiguration.

3. RESCOS (RESCUE Operating System)

To use the new hardware concept of scalability (i.e. unrestricted adding of further array elements) and reconfigurability (online reconstruction of the parallel operating processor array) the development of an assisting operating system became compulsory. RESCOS is a multiuser, multiprocessor and multitask system. It uses a socalled virtual processor as an abstract interface between users and physical processors, which are protected by fault

tolerant mechanisms. Fig. 4 shows that a user (U) intends to start a process or a process environment on the RESCUE computer. The user has to inform RESCOS about the quantity, the names and the demanded communication structure of the ordered processes. RESCOS accepts the order and offers him the demanded structure using the already occupied virtual processors (VP). The realized transformation between the user demanded process environment to the physical processors then has to be done under the lowest operation system level. This procedure is needed to transfer (in case of known processor defects) the demand to other processors. Even user fixed communication structures can be thus bypassed by message passing principals.

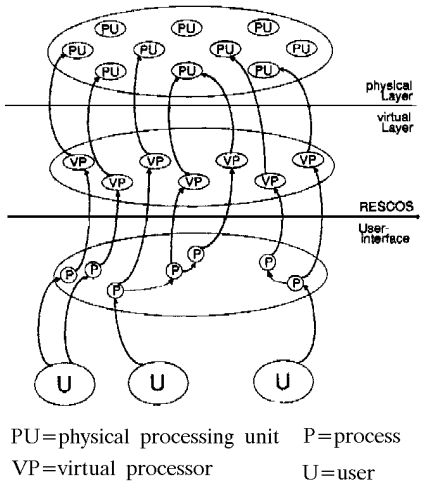

Fig. 4. Virtual processor - layer model.

The operating system is implemented as a distributed operating system. It is realized by a collection of independent processes which can be collected in logical groups.

The first group is formed by the socalled monitor processes containing all processes for the realisation and control of fault tolerant mechanisms and observing routines. Subdivisions are worker side and control side monitor processes. Among the worker side processes, mainly the worker nucleus process (WNP) is of importance because it collects the complete information on each worker from chained lists and transfers that information into the SM stage. A further workerside oriented process is the heartbeat pulser (HBP). It informs the CU via SM that it is still working.

Monitor processes on the side of the CU can furtheron be devided into master control processes (MCP) for load balancing calculations, into error checkers (EC) and into error solution processes (ESP), which are responsible for the error analysis and the reconfiguration procedures.

A second group is formed by the control processes. They control the administration of normal tasks and the use of means as well as message passing, load balancing and administration of the periphery and the RAM. The last of this group are the application processes, mainly the command shell and the user-processes.

4. Conclusions

The paper presents a new computer concept of medium grain parallel processing with the main point of the combination of observations strategies for parallel processors by shared memory arrays and strategical reconfiguration possibilities of communications between the existing array processors in case of arising hardware faults.

The authors have implemented both the hardware (RESCUE) as also the software (RESCOS) up to a demonstration level to be tested by interested research people.

REFERENCES

1. G. Cybenko and D. J. Kuck: *Revolution or evolution.* IEEE Spectrum, Sept. 1992, pp. 39-41
2. A. reuter: *Limitations of parallelism.* Informationstechnik it 34 (1992) 1, it 1/92, R. Oldenbourg Verlag, pp. 62-92.
3. R. Miller and Q. F. Stout: *Mesh computer algorithms for computional geometry.* IEEE Transaction on computers, Vol. 38 No 3, March 1989, pp. 321-340.
4. H.W. Meuer and E. Shchmaier: *Present parallel computer-Concepts and architectures.* Informationstechnik, it 34 (1992) 1, it 1/92, R. Oldenbourg Verlag, pp. 17-27.
5. S. Dutt and J. P. Hayes: *Some practical issues in the design of fault tolerant multiprocessors.* IEEE Transaction on computers, Vol. 41, No 5, May 1992, pp. 588-598.
6. G. Zorpette: *The power of parallelism.* IEEE Spectrum, Sept. 1992, pp. 28-33.
7. G. Zorpete: *Varying diameter and problem size in mesh-connected computers.* Proc. 1985 Int. Conf. Parallel Processing, pp. 667-699.
8. J. Schabernack: *Load balancing algorithms in distributed systems - Survey and taxonomy.* Informationstechnik it 34 (1992) 5, it 5/92, R. Oldenbourg Verlag, pp. 280-295.
9. R. Miederer: *A reconfigurable and scalable parallel computer with a distributed operating system and increased fault tolerance.* Dissertation at the department of electrical engineering with the "Lehrstuhl für Datenverarbeitung", Ruhr-Universität Bochum, 1993.