ADAPTIVE FAULT TOLERANCE IN MULTI-COMPUTER SYSTEMS USING DATA COMPACTION AND TWO-LEVEL VOTING

Andreas Stopp
Central Institute for Cybernetics and Information Processes,
Institute for Computer Architecture,
Kurstr.33, 0-1086 Berlin

ABSTRACT

Multi-computer systems offer good conditions for utilization of various fault tolerance principles. In order to increase reliability with little expense, this paper proposes an Adaptive Two-Level Voting scheme for multi-computer systems. This scheme enables a process-oriented dynamically determined fault tolerance by fault masking. The basic principle of the first voting level is a decentralized 2-out-of-N Signature Voting of exchanged process signatures generated by data compaction of N identical process copies on N computer nodes. It can tolerate N-2 faulty computer nodes. On the second voting level the results of the first voting level are compared by V-Version Programming using diverse software. For Two-Level Voting (V,N), the number of involved program versions V and the number of involved program copies N can be chosen. This paper describes four variants of Adaptive Two-Level Voting, a sequential, a parallel, a comparison variant and a very efficient dynamic variant. An estimation of efficiency of the suggested variants is also given.

1. INTRODUCTION

Commercial multi-computer systems with a large number of identical computer nodes exist already on the market (e.g. nCUBE, iPSC, Transputer Systems). Such multi-microcomputer systems have a good price/performance ratio and offer good conditions for utilization of software-implemented fault tolerance, e.g. reconfiguration, error recovery /6/, /2/, error compensation and fault masking. Different principles of fault masking techniques are known from literature /3/, /11/, /8/. An interesting complex system project is delta-4 /10/. In /1/ a good survey on multiple computation by replication in the three redundancy domains, namely time, space and information and the extension to diverse software is given. Multi-version programming /1/, /5/, is a very interesting but expensive procedure. We have to ensure that this procedure is mainly used against software design faults and we should treat the other faults by cheaper fault masking procedures. A survey on the pathology of faults is given in /7/. In this paper efficient procedures of Two-Level Voting (V,N) are proposed that utilize data compaction on the base of signature principles for replicated programs on different computer nodes. Replicating programs is simple in multi-computer systems and idle computer nodes are normally available.
The used 2-out-of-N Signature Voting is different to the adaptive voting in SIFT /14/ and the threshold voting in the FMPA-concept /9/. The goal is to improve the reliability with little expense by utilizing the advantages of hardware-supported process data compaction for fault tolerance and by eliminating the disadvantages by combining with multi-version programming. Adaptivity is a goal in many fault tolerance approaches. In this context adaptivity means that, after each process time slice the number of included program copies, computer nodes and program versions can be determined depending on the requirements of the actual process.

In Section 2 necessary assumptions are explained for the multi-computer system, the data compaction of process data and output data blocks and the hardware support by a Signature Circuit. Section 3 contains the proposals for 2-out-of-N Signature Voting, the first and lower voting level using replicated programs on N computer nodes and decentralized signature voters against hardware faults. In Section 4 the complete Two-Level Voting is described. Four different variants SEQUENTIAL, PARALLEL, COMPARISON and DYNAMIC of Two-Level Voting are proposed and discussed because they fit to different applications. The cost of the suggested variants are estimated.

2. DATA COMPACTION AND COMPUTER NODE DESIGN

A homogenous multi-computer system of M computer nodes is assumed (see Figure 1) connected by a communication network with redundant communication links (e.g. array or hypercube structures).

![Diagram of Multi-Computer System with Signature Circuit (SC) in each Computer Node]

Figure 1: Multi-Computer System with Signature Circuit (SC) in each Computer Node

All computer nodes must be suitable for executing program copies concurrently (not time-synchronous), for communicating by message exchange and for compacting selected data streams to signatures. In the multi-computer system each computer node should possess its own Signature Circuit for data compaction. The Signature Circuit is implemented in Gate-Array-Circuits (U5201.401 and XILINX 3020) using a 22-bit-polynomial with error detection probability \( P = 0.9999998 \) for 16-bit parallel data compaction. The Circuits are applied in computer nodes with Intel's 80286. A chip for 32-bit data compaction is planned. The Signature Circuit and its various applications for test, diagnosis and fault tolerance were described in the former papers /12/ and /13/. The Signature Circuit is