Constraint Based System-Level Diagnosis of Multiprocessors

J. Altmann[1], T. Bartha[2], A. Pataricza[2], A. Petri[2] and P. Urbán[2]

[1] University of Erlangen, Dept. of Computer Science III, Martensstr. 3, 91058 Erlangen, Germany
[2] Technical University of Budapest, Dept. of Measurement and Instrument Eng., Műegyetem rkp. 9, H-1521 Budapest, Hungary

Abstract. The paper presents a novel modelling technique for system-level fault diagnosis in massive parallel multiprocessors, based on a re-formulation of the problem of syndrome decoding to a constraint satisfaction problem (CSP). The CSP based approach is able to handle detailed and inhomogeneous functional fault models on a similar level as the Russel-Kime model [18]. Multiple-valued logic is used to describe system components having multiple fault modes. The granularity of the models can be adjusted to the diagnostic resolution of the target without altering the methodology. Two algorithms for the Parsytec GCel massively parallel system are used as illustrations in the paper: the centralized method uses a detailed system model, and provides a fine-granular diagnostic image for off-line evaluation. The distributed method makes fast decisions for reconfiguration control, using a simplified model.

1 Introduction

The large number of components built into *massively parallel multiprocessor systems* increases the probability of component faults. Since reliable operation over a long time period is also necessary for complex computations, the system must be able to mask the effect of occurring errors by *fault tolerance*. The underlying diagnostic principle is generally *system-level diagnosis*, followed by reconfiguration and recovery in case of a detected fault.

Different models and algorithms were developed for system-level diagnosis, typically originating from the first graph theory based "system-level models" (PMC for symmetric and BGM for asymmetric test invalidation) published in the late-sixties. Their mathematical apparatus is simple and well-elaborated; practical implementations also proved their usefulness. However, the implicit limitations – for instance the over-simplification of the test invalidation mechanism in order to assure a proper mathematical treatment – decrease the level of reality in these models. Moreover, the rapid development of electronic technology and computer architectures radically modified the basic assumptions used in diagnostic models [16]:

- fault rates are in general low and the dominating part of faults is *transient*;
- the complexity of additional components of the system (interface and communication circuits) is comparable with that of the processing elements;

- the number of the components in the systems drastically increased.

The majority of insufficiencies result from the hardest simplification of the PMC-type models: the assumption of a homogenous system and test structure (identical components with the same test invalidation over the entire system). This reduces their applicability due to the increasing practical importance of inhomogeneous systems.

1.1 Required Features of a New Diagnosis Method

The new requirements involved by the latest results in multiprocessor system design characterize the expected features of a general purpose self-diagnosis method:

- it should be applicable to inhomogeneous systems as well as to homogenous ones (components with different test invalidation models are to be considered);
- neither the actual system topology nor the test invalidation model should limit the diagnostic resolution (current methods use rigid, inadaptive algorithms seriously restricting the target system features);
- the algorithm should extract all of the useful information from the elementary diagnostic results (e.g. for estimating the level of diagnosis at run-time);
- it should be able to work in massively parallel computers with several hundreds or even thousands of system components, thus the algorithm should have an excellent efficiency even for a very high number of units under test;
- many applications demand "on-the-fly" diagnosis for a maximal performance, able to identify the fault states of certain units even from partial syndrome information (i.e. before receiving all of the test results).

These requirements necessitate a new approach. A generalized test invalidation model and syndrome decoding algorithm for inhomogeneous systems is published in [1]. However, the efficiency of the algorithm becomes to a crucial factor in case of large-scale systems due to the employed mathematical apparatus—operations on matrices of the dimension of the number of processor in the system.

Syndrome decoding is the most important step in diagnosis, determining the actual fault states of system components from the syndrome. This systematic search is in general NP-complete. The main intention of "artificial intelligence" (AI) methods is to find efficient solutions for difficult to solve problems. A group of them, the CSP (Constraint Satisfaction Problem) solving methods seem especially useful for system-level diagnosis [2].

2 System-level Diagnosis and CSP

2.1 Definition of the CSP

Constraint satisfaction problems (CSP) deal with the estimation of a single or all consistent solutions in large-scale relation systems. More formally, a CSP is a (X, D, C) tuple, where $X = \{X_1, X_2, ..., X_n\}$ is a *set of variables*, defined over the set $D = \{D_1, D_2, ..., D_n\}$ *domains*; and $C = \{C_1, C_2, ..., C_k\}$ is a *set of constraints*. Con-