NOVEL TECHNIQUES FOR QUANTIFYING CONFIDENCE DURING MULTI-PROCESSOR VERIFICATION, VALIDATION, DEBUG, AND DIAGNOSIS

Faculty Investigator: Ray Mercer, CE Group, Department of EE, Texas A&M

IBM Technical Contact: Matthew W. Mehalic, PowerPC Engineering Manager
High Performance Processor Development, IBM Microelectronics

Abstract

The focus of this research will be to understand the fundamental problems associated with the process of designing, developing, verifying, validating, debugging, and diagnosing multi-processor systems composed of microprocessors (such as those of the PowerPC product line) and associated memory and peripheral interface components. This understanding should lead to novel techniques to reduce the time from initial concept to full-scale manufacture and delivery of multi-processor computing systems. Particular areas of emphasis will be: (1) CAD algorithms and tools to assist designers and developers in these tasks, (2) automatic generation of test sets and testing sequences (using both formal and traditional methods) to reveal unexpected and undesirable behaviors in the multi-processor system, and (3) potential design rules to simplify the required algorithms and test sequence generation process. The ultimate goal will be to quantify the level of confidence in a particular design activity based upon the set of test cases that have been exercised and the structure of the design being tested.

Introduction

Over time, logic design organizations have been forced by economic necessity to develop powerful tools to enhance the quality and reliability of the digital systems delivered to end customers. For example, simulators at various levels (circuit, logic, register transfer, and system) allow designers to usually predict the response to an input sequence and compare the result with what was intended. Unexpected results are generally easily traced to the associated flaw in the designed structure, and the results are modifications to the structure to eliminate the unexpected (and undesired) behavior. However, the effectiveness of simulation to enhance the fundamental confidence in the correctness of a design is extremely sensitive to several parameters. First, the simulation model must match the actual manufactured device – including variations associated with operational parameters such as ambient temperature and voltage supply variations. After the "obvious" design problems have been detected and remedied, more subtle problems may simply be beyond the limitations of efficient simulation models. In this case, one or more manufactured copies of the system are normally operated for some extended time to demonstrate the viability of the actual final product and – in many cases – reveal additional unexpected behaviors.

In both the case of simulated models and actual system operation, the growth rate for confidence in the correctness of the design is extremely dependent upon the quality of the selected input sequence or sequences. For example, while test cases produced by the designer probably are relatively compact and efficient in the early design verification stages, these sequences usually suffer from a bias on the part of the designer to select sequences compatible with the intended function of the system. In contrast, individuals with a less detailed understanding of (and confidence in) the design may well (based upon their skepticism) produce sequences that are more robust. These inputs investigate the performance of the design under the influence of less probable (but still fully possible) operating circumstances. In the limit, some form of randomization in the input sequence is often used in an attempt to extend this separation between the person who knows how the device should work and the sequence source to sample how in actually works under as many varied situations as possible. No method for quantifying the level of confidence achieved during these activities currently exists. Instead, "engineering judgment" or time to market pressures determine when these processes are concluded.

Proposed work

Overview

In the particular case of multi-processor based computing systems, several aspects of this particular design problem exacerbate the situation described above. For example, the size of multi-processor systems is great, and the timing relationships between the individual processor elements – especially in the case of asynchronous communications – add yet another dimension of complexity to the problem.

The ultimate objective of this work will be to quantitatively relate the level of confidence achieved by exercising a design (simulated or physical) with the test sequence used in the verification / validation process. More aggressively, this quantitative measure should be based upon the aggregation of a "micro-confidence" measure across the structure of the design. With such an approach, the confidence measure first provides a figure of merit for when adequate cases have been investigated. More importantly, if "micro-confidence" is not uniform over the structure under test, then truly effective additional sequences can be targeted at those parts of the structure with the current lowest levels of "micro-confidence."

Previous results

In contrast to the tasks of verification and validation, manufacture testing has for many years included active research into the quantification of test pattern quality. For example, our previous work with Texas Instruments in Dallas, Texas, included development of a metric (MPG-D) to predict the defective part level based upon the circuit under test and the test pattern set applied [DWOR 99 and DWOR 00a]. Utilization of this more accurate defective part level prediction method allowed optimization of test pattern sets to efficiently minimize the final defective part level [GRIM 99, DWOR 00b, and DWOR 01]. Head-on comparisons between our optimized test pattern sets and standard commercial practice have demonstrated the superiority of our method on both academic and commercial circuits. In fact, these methods are now being utilized by Texas Instruments in large-scale manufacture testing operations for selected parts, and a report on this work won the Best Paper Award at the 1999 VLSI Test Symposium.

Recently, we have also begun to investigate extension of similar metrics to software based testing [WILL 01].

Proposed extensions to previous work

We propose to extend the metrics developed for manufacture testing to verification and validation activities for multi-processor based computing systems. However, the tasks of manufacture testing and verification/validation differ in several ways.

Most significantly, the cardinality of the test pattern space is finite for manufacture testing (where the number of possible patterns is *2** number of circuit inputs* for a combinational circuit). Because multi-processor systems can have asynchronous interfaces, the set of test conditions becomes infinite, and the associated delays are not deterministic due to operating conditions and manufacturing variations. In order to extend our models from the discrete to the continuous domain, we plan to investigate modifications using interval arithmetic. In such a model, the confidence will grow based upon the fraction of possible operating space verified or validated. The interval sizes consistent with continuity in behavior will be a key issue to study.

We also want to investigate the relationship between a failure in the function of a digital circuit and the manufacturing or design flaw that cause that failure. In the case of manufacturing flaws, yield calculations can be used to demonstrate that the probability of multiple flaws on the same die is vanishingly small as the quality of tested die increases to acceptable commercial levels. In contrast, the probability of multiple design errors is extremely large throughout the design verification/validation process, and only at the very end is it reasonable to expect that either one or no additional design flaws exist. In such a situation, we wish to evaluate the validity of the assumption of ergodicity normally utilized in manufacture testing and its extensibility to the multi-processor validation/verification environment.

Finally, we want to investigate the interaction between the excitation of a design error in a multi-processor system and the observation or detection of erroneous logic values, states, and circuit conditions that result. These problems are much simpler in the case of manufacture testing of combinational circuits where primary input values determine excitation conditions and primary output values communicate erroneous results. In contrast, multi-processor systems include a complex (not necessarily unidirectional) graph of input conditions and their resulting outputs. It remains to be determined if this complication requires fundamental enhancements to the defective part level model we currently employ.

References

[GRIM 99] M. R. Grimaila, S. Lee, J. Dworak, K. M. Butler, B. Stewart, H. Balachandran, B. Houchins, V. Mathur, J. Park, L-C. Wang, and M. R. Mercer, "REDO -- Random Excitation and Deterministic Observation -- First Commercial Experiment," *Proc. 1999 IEEE VLSI Test Symposium,* Dana Point, Calif., April 25 - 29, 1999, pp. 268-274. **(Best Paper Award at 1999 VLSI Test Symposium)**

[DWOR 99] J. Dworak, M. R. Grimaila, S. Lee, L-C. Wang, and M. R. Mercer, "Modeling the Probability of Defect Excitation for a Commercial IC with Implications for Stuck-at Fault-Based ATPG Strategies," *Proc. 1999 International Test Conference,* Atlantic City, NJ, September 28 - 30, 1999, pp. 1031-1037.

[DWOR 00a] J. Dworak, M. R. Grimaila, S. Lee, L-C. Wang, and M. R. Mercer, "Enhanced DO-RE-ME Based Defect Level Prediction Using Defect Site Aggregation – MPG-D," *Proceedings of the 2000 International Test Conference,* Atlantic City, NJ, October 3 - 5, 2000, pp. 930-939.

[DWOR 00b] J. Dworak, M. R. Grimaila, B. Cobb, T-C. Wang, Li-C. Wang, and M. R. Mercer "On the Superiority of DO-RE-ME / MPG-D Over Stuck-at-Based Defective Part Level Prediction," *Proceedings of the 2000 Asian Test Symposium,* Taipei, Taiwan, December 4-6, 2000.

[DWOR 01] J. Dworak, J. Wicker, S. Lee, M. R. Grimaila, K. M. Butler, B. Stewart, L-C. Wang, and M. R. Mercer, "Defect-Oriented Testing and Defective-Part-Level Prediction," *IEEE Design and Test of Computers*, January-February, 2001, Vol. 18, No. 1, pp. 31 - 41.

[WILL 01] T. W. Williams, M. R. Mercer, J. P. Mucha, and R. Kapur, “Code Coverage, What Does It Mean in Terms of Quality?” *Proceedings of the 2001 Annual Reliability and Maintainability Symposium,* Philadelphia, PA, January 22-25, 2001.