Simulation Concepts to Model Real-Time and Dependability Properties of Symmetric Multiprocessor Systems

August 10, 1998

Peter Magnusson, Bengt Werner
Computer Architecture Simulation Group
Computer and Network Architectures Laboratory
Swedish Institute of Computer Science
{psm,werner}@sics.se

Summary

Symmetric multiprocessor systems (SMPs) will be a cost-effective technology in providing the processing capacity that is required by future transaction-oriented telecom applications. Recent developments in instruction-level simulation techniques, as exemplified by the SimICS technology, have made it possible to verify the function and explore performance consequences of design alternatives of future systems at the application, the system software, and the hardware level. Unfortunately, state-of-the-art simulation technologies lack concepts to also verify quality-of-service and dependability requirements of multiprocessor systems. The aim of this project is to fill this gap.

In two Ph. D. projects, we intend to develop concepts that extend state-of-the-art system simulation techniques with capabilities to (1) analyze timing to meet QoS requirements and to (2) analyze the fault coverage of fault-tolerance mechanisms. The distinguishing feature of our approach is to take advantage of the complete observability and determinism of simulation technologies. By including the developed concepts in simulation models, it will be possible to verify quality-of-service and dependability aspects of design alternatives of future systems e.g. telecom and transaction processing.

To factor in requirements from transaction-oriented applications, we will apply our developed concepts to a proprietary database engine run on a cluster of multiprocessor workstations at Ericsson UAB. The project involves two PAMP-nodes: SICS and Ericsson UAB in addition to Virtutech, a recently started company. Two senior scientists and two Ph.D. students at SICS and at least one senior scientist and one MSc student at Ericsson will actively work on the project. The project will exchange ideas and experimental resources with the PAMP-nodes at Chalmers as well as at Karlskrona/Ronneby.

Key words: Symmetric multiprocessors, transaction processing, hardware and software system design, high-performance, quality of service, real-time systems, system level simulation.

Problem Statement

Emerging information technology applications such as information servers, multimedia servers and telecommunications systems, are characterized by high performance requirements. SMP technology is promising in cost-effectively meeting this requirement. However, because of the rapid generation shift of commodity components (for example, a new generation of a high-performance microprocessor is currently shipped once every 18 months), multiprocessor application designers' tasks are challenging; they have to make cost-effective design tradeoffs across multiple layers of interacting hardware and software subsystems in a fraction of this time. Lacking adequate tools for performance and function modeling, design decisions are currently mostly based on intuition which often results in conservative design decisions which severely hampers the competitiveness of future products.

Simulation methodologies are an effective approach in the analysis of functional and performance behavior of a complete computer system. The major advantage is that hardware/software interactions can be studied in a deterministic manner without the availability of the hardware platform. The functional and timing specification of the hardware platform is implemented in a software simulation model which makes the interaction of the complete software system, encompassing application and system software, with the hardware platform visible. Therefore, such a simulation system provides a powerful environment to verify function as well as tracking performance bottlenecks.

Instruction-level simulation techniques were for a long time too slow to model the interaction between complex applications run on commodity operating systems on multiprocessor systems. Recently, however, two simulation systems, SimICS (SICS)[1] and SimOS[2] (Stanford), with similar design goals but using quite different design principles have emerged that are capable of applying the complete system simulation approach to large commercial applications. Key properties of both approaches are to provide a fast instruction-level simulation kernel and provide a functional interface to the OS to make it possible to run completely unmodified system and application binaries on the simulated hardware platform. Unlike the SimOS system, however, the SimICS system can afford to model hardware-level events, such as cache misses, at a reasonable slow-down of less than 100. Moreover, SimICS became in 1997 the first generally available simulator capable of booting and running unmodified commercial strength operating system environments (Solaris 2.6 and Linux 2.0.30).[3]

While it is possible to use the SimICS system to identify functional and performance-oriented design errors (bugs) thanks to its advanced profiling system, key concepts are missing to make accurate and efficient timing and dependability analysis. Such capabilities are fundamental in advancing state-of-the-art of *complete system simulation methodologies* to be applicable to industrial applications on multiprocessors with explicit quality-of-service and dependability requirements.

An example of an application with such requirements is the one provided by Ericsson UAB that is used as a case study in this project. Efficient databases are an essential technology in controlling a telecom system. Ericsson UAB is currently implementing a parallel database on a shared memory multiprocessor (SMP). The SMP is composed of a cluster of multiprocessor workstations connected by SCI. Predicting and understanding the performance of such a system involves getting access to detailed statistics from the system, e.g. cache hit ratios and resource contention. This knowledge is crucial to aim performance optimization on the true bottlenecks of future system releases.

From the point of view of important design issues, the essential characteristics of the aggregate system are its overall performance, the variability of that performance, and the reliability of the system. In particular, state-of-the-art simulation methodologies must be extended with concepts to enable (1) accurate worst and best-case timing analysis of code sequences and (2) fault-coverage analysis of fault-tolerance mechanisms at any system level. Such concepts are currently lacking in the open literature in the system simulation area and provide the scientific basis for the two sub-projects in this research proposal.

Proposed Approach

Together with Ericsson UAB and Virtutech, SICS will develop simulation concepts that make it possible to extend SimICS to verify functional, real-time, and dependability requirements of this (and similar) applications. A key goal is to make it possible to include the developed concepts in the simulation environment with as little impact on simulation speed as possible.

Our overall research methodological approach to evaluate developed concepts is by developing an experimental simulation platform of the complete system with its application. By comparing the simulation results with measurements on the real system, we can make sure that all system aspects are incorporated in the model. A complete system simulation platform provides the first major milestone of the project and will engage personnel at SICS as well as personnel at Ericsson UAB (for details, please refer to Project Plan).

Having established the experimental simulation platform, two Ph. D. projects with quite different goals, although with the same general goal of advancing complete system simulation methodologies, will be carried out by the two Ph.D. students. The basic approaches along with some initial ideas are outlined below.

---

[1] P. Magnusson *et al*. SimICS/sun4m: A Virtual Workstation. In *Proceedings of the 1998 USENIX Annual Technical Conference (USENIX'98).* June 15-19

[2] Stephen Herrod. *Using Complete Machine Simulation to Understand Computer System Behavior*, Ph.D. Thesis, Stanford University, February 1998. The thesis is available at http://simos.stanford.edu

[3]Virtutech, a recent SICS spin-off company, has licensed commercial rights to the SimICS environment and will participate in this project.

Timing Analysis

The first Ph. D. project aims at developing concepts for analyzing timing requirements in applications with quality-of-service requirements. In such applications, the execution time for a particular code sequence should be within certain bounds. This calls for concepts to establish best and worst-case execution times of performance-critical code sequences, which is the major focus of this project. The timing analysis should take all the features of high-performance multiprocessor hardware platforms into account such as pipelining and multi-level memory hierarchies.

In the past, we have developed a timing analysis methodology that aims at very accurate timing analysis of code sequences with minimal impact on simulation speed.[4] The basic methodology is the following. First, with the profiling system of the SimICS system, the performance-critical code sequences are first identified.. Second, the fast instruction-level simulation kernel then produces instruction traces of these code sequences and feed them into a detailed timing model of the hardware platform. The timing model is then used to accurately determine the execution time of the actual execution. We plan to extend this basic methodology to determine execution time bounds of code sequences.

Unlike other approaches in the real-time literature, which mainly focus on the WCET problem of entire programs in hard real-time environments, we are focusing on *performance-critical code sequences* to meet quality-of-service goals in e.g. telecom and transaction processing applications.

Dependability analysis

The second Ph. D. project aims at developing concepts to verify a number of dependability requirements of typical telecom applications. In order for such applications to meet specified reliability goals, a fault model reflects the type of faults and their distributions. To tolerate faults, the designer can choose among a number of approaches ranging from fault-tolerance mechanisms at the application, system software, and hardware level. The particular approach chosen reflects a tradeoff between design complexity, design time as well as performance. Consequently, it is important that the designer have access to an environment in which such tradeoffs can be studied in a systematic fashion. A complete system simulation approach provides such opportunities but has not, to the best of our knowledge, been used before to analyze fault-tolerance mechanisms. This is the goal of the second sub-project.

We plan to develop concepts to incorporate fault-injection mechanisms at all system levels and support for fault-coverage analysis. The particular challenge lies in integrating these concepts with as little impact as possible on the simulation speed. We will compare the observability properties of this approach with other fault-injection approaches in the literature and carry out a number of case studies based on relevant mechanisms dictated by the application case.

Expected Results and Impact

The expected academic results of the project is:

- ***A complete system simulation model of an industrial transaction-oriented application.*** Very little experience with using complete system simulation on commercial applications is reported in the literature.
- ***Concepts and tool support for integrating quality-of-service timing analysis in simulation models of multiprocessor systems.*** Such concepts are currently missing in the open literature.
- ***Concepts and tool support for integrating fault-coverage analysis in simulation models of multiprocessor systems.*** Such concepts are currently missing in the open literature.
- ***Knowledge of performance characteristics of a commercial application***
- ***Design methodologies to systematically make design tradeoffs to meet functional, performance, quality-of-service, and dependability requirements of multiprocessor systems using simulation techniques.***

Overall, the project will advance state-of-the-art in complete system simulation methodologies in a number of fundamental ways.

---

[4] Werner, B. and P. S. Magnusson. 1997. A Hybrid Simulation Approach Enabling Performance Characterization of Large Software Systems, *Proceedings of MASCOTS'97*

Ericsson UAB's outcome of the project is, in addition to the above, important knowledge about the characteristics of their application.

- Performance analysis of a database system.
- A tool to develop and optimize applications on.
- Increased knowledge in performance measurement and optimization

Project Plan

The approach taken to develop methods to exploit the performance potential of SMP technology is to use applications provided by the industrial partners as study objects or cases. Functional and performance requirements of the applications will serve as sources for identifying relevant research issues to focus on. Based on the functional and performance requirements of the applications, the second phase of the project is devoted to development of concepts (hardware and software design methods) aiming at shortening the design cycle for designers. The third phase, called the evaluation phase, aims at applying the methods to the applications provided by the industrial partners to understand the strengths/weaknesses of the developed methods. Finally, the fourth phase aims at refining and generalizing the concepts developed so as to make them applicable to a wider scope of applications/systems. The schedule for these activities is as follows:

- Application analysis phase (ends June 1999)
- Conceptual development phase (ends June 2000)
- Evaluation phase (ends June 2001)
- Refinement and generalization phase (ends June 2002)

Application Analysis

The first phase of the project, *the application analysis phase,* aims at developing a simulation model of the target system used as a case study. This phase will be finalized after approximately one year and delivers a report on the analysis of function performance, dependability, and quality-of-service required by the system.

In the first phase of the project, the following tasks have been identified as SICS responsibility and will mainly be carried out by two Ph. D. students. The work will be based on the SimICS simulation platform. Virtutech will provide SICS with necessary licenses and support for the SimICS/sun4u simulator, which will be implemented as a new generation of the existing SimICS/sun4m toolkit during the early phase of the project (during 1998).

- Creating a timing simulator for UltraSparc.
- Generalizing the timing simulator to support a multiprocessor.
- Generalizing the timing model to calculate best-case and worst-case timing.
- Building a memory hierarchy for the Ericsson prototype configuration including an SCI interface.
- Verifying the function and timing of the model with the prototype
- Simulator support for Solaris lightweight threads

These tasks all aim at supplying a measurement environment. The project leverages on earlier experiences of performance measurements in cooperation with Ellemtel and processor simulator development of SparcV8 and Motorola 88000 architectures. Ericsson UAB is responsible for a number of sub-tasks within the project. The following have been identified:

- Database design
- UltraSparc Plex compiler
- Plex interpreter
- Performance measurement and optimization
- Adjusting the application software to run on the simulator

- Creating software/traces that generate load for the application

The simulator has a well-defined application that should run on it. This defines the minimal functionality that will be supported. The simulator supports emulation of parts of the Solaris operating system. While the user mode instruction set will be fully supported, including some extensions, not every possible system call will be supported. The project may decide to proceed with a system-level model instead, in which case Solaris 2.7 will boot on the simulator, on top of which in turn the database application will execute.

Verification of the instruction set functionality will be automated as much as possible. In addition to this verification of the timing model can be done by comparing executions on the simulator with executions on a prototype system. The application can also run on the prototype system where some data can extracted e.g. execution time and hardware counter measurements.

SimICS supports loadable objects. Two such objects that will be developed are a memory hierarchy and a timing simulator. The memory hierarchy contains a functional model of the memory system including caches. The timing simulator is fed with a trace of functions being performed and calculates the time for these.

The Ericsson application prototype is a proprietary database engine based on AXE 10 technology. To maintain software compatibility an emulator for APZ, Ericsson's proprietary hardware architecture, runs on an open workstation, currently Sparc. The shared memory is used to implement efficient message passing between APZ emulators. To speed up execution some software (which is written in Plex) can be compiled directly to native Sparc code instead of being interpreted.

The system is typically 5-15 nodes but should allow for a large number (100-1000) of nodes to work concurrently. The prototype node is a 2 processor UltraSparc Server. The network is SCI, which gives the system a distributed shared memory. Currently SCI is connected to the I/O bus but in the future, it may be connected directly to the memory bus (for increased performance). The database may be accessed via an ATM interface or via SCI.

The application is a complex mixture of operating system, statically compiled source code, and binary translation within an emulation layer, all executing on a parallel architecture. This is a mixture of technologies that we expect to become more common as industrial projects pursue hybrid solutions to achieve partial backward compatibility; high performance, quality of service, and reliability; and shortened time to market. Studying the characteristics of such a system is a challenge.

Conceptual Phase

In the conceptual phase, we will pursue a variety of techniques of leveraging new simulation tools to support design tasks of this class of application. In particular, very recent tools, such as SimICS/sun4m, allow designers to study an entire, realistic workload within a controlled environment.

This phase aims at developing concepts to incorporate support mechanisms to enable analysis of quality-of-service and fault-coverage in the simulation platform. The two sub-projects (as outlined in Proposed Approach) are carried out by the two Ph. D. students focusing on QoS and fault-tolerance support concepts, resp. This phase is anticipated to last for one year and the two Ph.D. students are expected to write reports that together with the work in the application analysis phase corresponds to the requirements for a Lic Degree.

Evaluation Phase

The evaluation phase will ascertain the validity of the simulation support concepts (in terms of accuracy of measurements in comparison to other data, including hardware profilers), as well as evaluate the efficacy of the design techniques and metrics applied and/or developed in the conceptual phase.

Refinement/Generalization Phase

This final phase will attempt to generalize from the successes, or failures, of manual, semi-automated, or automated methods developed in the project. The results should hopefully be applicable to a broad class of applications on shared memory multiprocessor targets in general.

Preliminary Budget

The project will involve two senior scientists from SICS (Peter Magnusson and Bengt Werner), two graduate students from SICS, and staff from Virtutech and Ericsson UAB. Moreover, Per Stenström at Chalmers will act as a scientific advisor in the project. SICS senior scientists will participate at a level of up to 20% each, and the two SICS graduate students between 80% and 100%. Virtutech will assist to the extent needed to support licenses for SICS of the SimICS/sun4u platform, which will be developed during 1998. From Ericsson at least one senior scientist and at least one M. Sc. student will be using and developing the simulator.

SICS has standard rates which include all overhead, and depends on scientist category.

The total costs for the SICS scientists are shown in the following table:

SICS Staff	Annual Cost (kkr)
Peter Magnusson	285 (at 20%)
Bengt Werner	214 (at 20%)
Ph.D. Student 1	912 (at 100%)
Ph.D. Student 2	912 (at 100%)
Total	2323

We are asking the funding agency to cover an annual cost of one MSEK, starting on July 1st 1998 and through the end of 2002. Thus the funding asked for is 0.5 MSEK for 1998 and one MSEK for each of the remaining four years. SICS expects to cover costs exceeding the rate of one MSEK per year by internal research funds.

The two Ph.D. students will be recruited specifically for this project.

Related Research

In general terms, research into metrics and design methodology to achieve quality of service, performance, and/or reliability, as well as short time-to-market, is a very broad research area, clearly beyond even a brief summary in this document. More specifically, the field of leveraging from system-level simulation to pursue these issues is a very new field, with very little work reported so far. The only other group that we are aware of doing similar work is the SimOS group at Stanford.[5] Their work has produced early results in characterizing application behavior, in particular performance. To our knowledge, research has not yet been conducted on the particular issue of dependability and quality of service design issues that can be addressed with system level simulation.

Relation to ARTES and PAMP Profiles

We believe this project is an excellent fit within both the ARTES and PAMP profiles. In particular, in relation to the ARTES twofold vision, we believe the methodology we intend to explore has great potential in reducing lead time in designing and modifying real time systems, by providing a radically more powerful environment for the designers. Regardless of the level of success of this approach in significantly reducing lead time concerns, the project will result in a transfer of knowledge to industry of the latest research developments in system level simulation techniques as a source of tools to support system design in general, and real time system design in particular.

Industrial Relevance

As we described earlier, the class of applications represented by the case study that we will focus on is one that we believe will be increasingly common. A difficult characteristic of this application is that it is exceedingly difficult to analyze from an aggregate perspective. It is perceived as being of great value to significantly improve the available tools and/or methods for characterizing, designing, and modifying such commercial applications. This will yield a more powerful design methodology that will increase the competitiveness of future systems.

---

[5] *Using Complete Machine Simulation to Understand Computer System Behavior*, Stephen Herrod, Ph.D. Thesis, Stanford University, February 1998. The thesis is available at http://simos.stanford.edu

Context

The CAS Group

The CNA-lab at SICS (Bengt Werner) has a strong background in modeling and analyzing computer systems, especially shared memory architectures. They will develop modular simulation technology based on its simulator platform SimICS to enable accurate performance analysis of a parallel database system developed at Ericsson UAB.

Industrial and Academic Cooperation

Ericsson UAB is studying implementations of network databases for future telecom products where SMP-technology plays an important role. Their interest in this project is to get acquainted with performance prediction methodologies to pinpoint bottlenecks across the hardware/software boundary in SMP platforms. (Contact person: Mikael Ronström). From a PAMP perspective, the project will export simulation methodologies to be used in the projects at Chalmers and Karskrona/Ronneby.

Applicant CVs

PETER S. MAGNUSSON (http://www.sics.se/~psm) is a senior researcher and head of the Computer Architecture Simulation group in the Computer and Network Architectures Laboratory. He received his MSc degree in computer science in 1992 from the Royal Institute of Technology and MBA in 1993 from the Stockholm School of Economics. His research is focused on efficient techniques for implementing instruction set simulators and modeling performance critical resources in modern computer systems such as virtual memory, data cache, and instruction cache. He is a member of IEEE and ACM, a columnist for Sweden's premier computer magazine Datateknik, and an information technology advisor to Öhmans Fondkomission.

BENGT WERNER is a senior researcher at the Computer and Network Architectures Laboratory. He received his MSc degree in 1989 from Lund University. His research is focused on modeling and simulation of computers and networks with special interest in timing accurate modeling of such systems. He has previously been working with performance evaluation at Ericsson's R&D.

Relevant publications:

Magnusson, P. S. 1993. A Design for Efficient Simulation of a Multiprocessor, in Proceedings of MASCOTS, pages 69-78, January .

Magnusson, P. S. and B. Werner. 1995. Efficient Memory Simulation in SimICS, in 28th Annual Simulation Symposium, April .

Magnusson, P. S. and D. Samuelsson. 1994. A Compact Intermediate Format for SimICS, Technical Report T94:17, Swedish Institute of Computer Science, September.

Magnusson, P.S. 1997. Efficient instruction cache simulation and execution profiling with a threaded-code interpreter. In Proceedings of the Winter Simulation Conference (WSC'97). December 7-9.

Magnusson, P. S., F. Dahlgren, H. Grahn, M. Karlsson, F. Larsson, F. Lundholm, A. Moestedt, J. Nilsson, P. Stenström, B. Werner. 1998. SimICS/sun4m: A Virtual Workstation. In Proceedings of the 1998 USENIX Annual Technical Conference (USENIX'98). June 15-19.

Montelius, J. and P. S. Magnusson. 1997. Using SimICS to evaluate the Penny system. In Proceedings of ILPS'97.