Hybrid Analytical-Simulation Model for Performance Prediction of Distributed Systems

Eng. Cosmina Chişe

Thesis Summary

The thesis presents a new hybrid approach for performance modeling and analysis, to be applied in case of distributed systems described by UML diagrams with MARTE (Modeling and Analysis for Real-Time and Embedded systems) annotations. The proposed method is implemented, additionally to an enhanced simulator, in a tool called PHYMSS (Performance Hybrid Model Solver and Simulator).

The introductory chapter places the thesis subject in the context of current field research. The benefits of applying Software Performance Engineering (SPE) practices to distributed systems are briefly reviewed. Distributed systems are widely adopted as a response to the increasing demand of computational resources and also to facilitate integration of remote components. Such systems are at risk of being useless, since design flaws often lead to critical performance issues at deployment, when usually they are too costly to fix. SPE is concerned with modeling and prediction of system non-functional properties early in the software development process, leading to informed design decisions. A complementary research direction covered by the thesis addresses the need of analysis automation using tools and standardization of the modeling languages for input models provided to such tools. The chapter also establishes clear thesis objectives both in theory and in practice (implemented application), without omitting proper validation.

Chapter 2, *„Software Performance Engineering Approaches”,* presents the main research directions in SPE and the classification of hybrid modeling approaches. The entire performance analysis process is covered, from system modeling languages, performance models solved analytically or simulated, intermediate performance models, particular models for component-based systems, and automatic interpretation of performance results providing better design alternatives. Considering the classification of hybrid methodologies, the hybrid model proposed in the thesis belongs to class III.

Chapter 3, *„Standardization of Distributed Systems Modeling Languages"*, briefly reviews the existing approaches of model transformation from generic models to performance models, performance evaluation automation and modeling language standardization in a chronological order. Before year 2000, there were various system specification formalisms (no standard notation existed) and only few tools to implement model transformation techniques. Afterwards, along with the development of software tools implementing both transformations and solution algorithms, the need of a unified standard notation emerged. This need was first addressed by the definition of the UML SPT (Schedulability, Performance and Time) profile in 2004, which was rapidly adopted by analysis tools. A more comprehensive profile, MARTE, is under development since 2007, as an extension for the new UML 2.0 diagrams. Since there are few tools that rely on UML 2.0 and the MARTE profile, the current thesis addresses this issue by presenting the implementation of a robust software application, having a friendly user interface, supporting UML MARTE models as inputs and covering the entire performance prediction process for distributed software systems, application called PHYMSS.

The first main contribution of the thesis is described in chapter 4, „*Hybrid Analytical/Simulation Model and Solver*": a new hybrid analytical-simulation approach regarding performance model solving. It is based on a flexible scenario oriented meta-model from which hybrid models can be derived and then solved using an iterative algorithm defined by the thesis author.

The hybrid performance model is decomposed into layers in order to be able to address submodels. Depending on the information specified in the input diagrams, in case a hierarchical decomposition is not possible, a sequential breakdown is implemented. Performance model layering allows separation of a simulation submodel from a certain layer downwards, layer that will be referred to as „simulation level". The simulation model was extended by the thesis author to allow nested calls (from sequence diagrams), additionally to sequential calls (as defined in activity diagrams). The analytical model is adaptive: Layered Queueing Network for a hierarchical layering if it is possible, or Queueing Network for a sequential breakdown.

Concerning the proposed hybrid solver, each iteration consists of two steps: submodel simulation and analytical solving by Mean Value Analysis extensions. Simulation results for the submodel are used in the higher level submodels by the analytical solver. Both pure analysis techniques are enhanced by the author: multi-threaded version of the originally coroutine-based simulator; analytical solver implemented using two approximations, Bard-Schweitzer and Chandy-Neuse, the latter being an improved estimation. Also, multiple passive resources and mixed requests (open and closed) are modeled and the formulae for solving open systems have been adapted by the thesis author to obtain better accuracy.

Chapter 5, „*Transformation of UML MARTE Models to the Hybrid Model*", presents the transformation rules from UML diagrams with MARTE or SPT performance annotations into the proposed hybrid performance model. The chosen set of supported diagrams consists of the most widely used UML diagrams, in order to facilitate performance modeling for the software developer. Active and passive resources are modeled using deployment diagrams, while system behavior is described in two ways: use case diagrams detailed with activity diagrams, or sequence diagrams.

Method evaluation is performed with a tool called PHYMSS, described in chapter 6. This tool is developed by the thesis author in C# and allows hybrid analysis, pure simulation and pure analytical calculus. The pure simulator uses an improved multithreaded version of a coroutine-based simulation model, model which is extended to support nested calls, additionally to sequential actions. PHYMSS allows performance prediction automation, the only user intervention being the analysis of the results: in case values of performance parameters are satisfying, the system implementation may start, otherwise the system is redesigned according to the detected issues and the automated analysis step using PHYMSS is repeated on the new model. The tool requires two input files: UML MARTE model in XMI format and a configuration file in JavaScript format (.JS). The user's guide is also presented, proving how friendly the GUI is, by allowing analysis configuration and also intermediate results monitoring during the simulation.

Several case studies are described in chapter 7, in order to prove the efficiency of the new hybrid model and solver, compared to pure analytical and simulation approaches. The validation strategy implies a specific order in considering the implemented approaches. First, the simulation is validated in order to use simulation results as reference values when validating the other methods. Then, the analytical solver is considered for validation. The proposed hybrid technique is compared to the pure methods, focusing on accuracy and analysis duration. An original case study is built by the author (an air-traffic control system model) to cover the potential complexity of activity diagrams and also to illustrate how the performance prediction process aids in detecting design flaws and improving the model. The hybrid method proved to converge much faster than simulation and to yield more accurate results than analytical calculus.

Chapter 8, „*Contributions and Future Work*", emphasizes the contributions of the thesis and future research perspectives. The most important contributions to the SPE field, both in theory and in practice, are the following:

- Comprehensive study and systematization of publications in the SPE field and related areas;
- New hybrid performance model and solver defined by the thesis author;
- Implementation of an application, PHYMSS, supporting input UML diagrams that comply with the most recent standard for performance annotations, MARTE, and allows several analysis methods to be run, including the proposed hybrid method.
- Validation strategy, starting with pure approaches in order to consider their results as reference values when validating the hybrid method, the measurement strategy for an online real system, and building an original case study.

Future research directions, whose results will be continuously disseminated in conferences, are presented below:

- Extended range of input diagrams, and also extending PHYMSS GUI;
- Heuristically detect the appropriate value for the simulation level in order to assist the user when choosing this level, prior to running the analysis;
- Extensions to the hybrid meta-model to include QoS (Quality of Service) parameters and to the solver (worst-case analysis).