Jan Jürjens, Kurt Schneider (Hrsg.): Software Engineering 2017,
Lecture Notes in Informatics (LNI), Gesellschaft für Informatik, Bonn 2017 95

Model-Based Self-Aware Performance and Resource Management Using the Descartes Modeling Language

Samuel Kounev[1], Nikolaus Huber[1], Fabian Brosig[1], Simon Spinner[1], and Manuel Bähr[2,3]

Abstract: We present the results of our recent work published in [Hu17] and summarized in [Ko16]. We introduce a holistic model-based approach for self-aware performance and resource management of modern IT systems and infrastructures. Based on a novel online performance prediction process, we implement a model-based control loop for proactive system adaptation. We evaluate our approach in the context of two representative case studies showing that with the proposed methods, significant resource efficiency gains can be achieved while maintaining performance requirements. These results represent the first end-to-end validation of our approach, demonstrating its potential for self-aware performance and resource management of modern IT systems and infrastructures.

Keywords: Modeling, performance prediction, resource management, self-aware computing

Modern IT systems have increasingly distributed and dynamic architectures providing flexibility to adapt to changes in the environment and thus enabling higher resource efficiency. However, these benefits come at the cost of higher system complexity and dynamics. Thus, engineering systems that manage their end-to-end application performance and resource efficiency in an autonomic manner is a challenge. To tackle this challenge, techniques for *online* performance prediction are needed [Ko10]. Such techniques should make it possible to continuously predict at run-time: a) changes in the application workloads, b) the effect of such changes on the system performance and resource efficiency, and c) the impact of possible adaptation actions at run-time.

We argue that sophisticated modeling and prediction techniques are needed, specifically designed for performance and resource management in online scenarios. Such techniques should allow to capture both static and dynamic system aspects, including all relevant influences of the system's resource landscape, its architecture, as well as its adaptation space, adaptation strategies and processes, in a generic, human-understandable and reusable way. Moreover, we need model-based system adaptation mechanisms that apply such modeling and prediction techniques end-to-end, to drive autonomic decision making at run-time.

In our previous work [Hu14, BHK14, KBH14], we proposed the Descartes Modeling Language (DML), an architecture-level modeling language for online performance and resource management specified by meta-models based on OMG's Meta-object Facility

[1] Department of Computer Science, University of Würzburg, Am Hubland, 97074 Würzburg. E-mail: {samuel.kounev,nikolaus.huber,fabian.brosig,simon.spinner}@uni-wuerzburg.de
[2] Blue Yonder GmbH & Co. KG., Karlsruhe, Germany. E-mail: manuel.baehr@blue-yonder.com
[3] This work was funded by the Deutsche Forschungsgemeinschaft (DFG) under grant No. KO 3445/11-1 and grant No. KO 3445/15-1. Parts of this work have been done during Fabian Brosig's summer internship at Salesforce.com. We acknowledge the many fruitful discussions with Jürgen Walter and Simon Eismann from the Chair of Software Engineering, University of Würzburg.

(MOF) [OM15]. This work was focused on finding suitable abstractions for describing the performance-relevant aspects of applications, the influence of their resource environment, and the processes for adapting the system according to high-level goals specified in service-level agreements (SLAs). In the context of DML, we have developed and released a number of open-source tools, each of them focussed on a given task within our overall approach[7].

In this talk, we provide an overview of the results of a recent paper, published in [Hu17], which uses DML as a foundation and makes the following contributions on top of it: i) an end-to-end approach for self-aware performance and resource management based on a holistic model-based adaptation control loop exploiting the reflective capabilities of DML, ii) a novel online performance prediction process that dynamically tailors the model solving taking into account the requested performance metrics, as well as goals in terms of accuracy and overhead, and iii) the first real-world evaluation and validation of our approach in the context two industrial case studies demonstrating the benefits of model-based adaptation control loop and the online performance prediction process for performance and resource management of modern IT systems and infrastructures.

A more detailed high-level overview of our approach can be found in [Ko16].

References

[BHK14] Brosig, Fabian; Huber, Nikolaus; Kounev, Samuel: Architecture-Level Software Performance Abstractions for Online Performance Prediction. Elsevier Science of Computer Programming Journal (SciCo), Vol. 90, Part B:71–92, September 2014.

[Hu14] Huber, Nikolaus; van Hoorn, André; Koziolek, Anne; Brosig, Fabian; Kounev, Samuel: Modeling Run-Time Adaptation at the System Architecture Level in Dynamic Service-Oriented Environments. Service Oriented Computing and Applications Journal (SOCA), 8(1):73–89, March 2014.

[Hu17] Huber, Nikolaus; Brosig, Fabian; Spinner, Simon; Kounev, Samuel; Bähr, Manuel: Model-Based Self-Aware Performance and Resource Management Using the Descartes Modeling Language. IEEE Transactions on Software Engineering (TSE), PP(99), 2017. To appear.

[KBH14] Kounev, S.; Brosig, F.; Huber, N.: The Descartes Modeling Language. Technical report, Department of Computer Science, University of Wuerzburg, October 2014. http://nbn-resolving.org/urn:nbn:de:bvb:20-opus-104887.

[Ko10] Kounev, Samuel; Brosig, Fabian; Huber, Nikolaus; Reussner, Ralf: Towards self-aware performance and resource management in modern service-oriented systems. In: Proceedings of the 7th IEEE International Conference on Services Computing (SCC 2010), July 5-10, Miami, Florida, USA. IEEE Computer Society, July 2010.

[Ko16] Kounev, Samuel; Huber, Nikolaus; Brosig, Fabian; Zhu, Xiaoyun: A Model-Based Approach to Designing Self-Aware IT Systems and Infrastructures. IEEE Computer, 49(7):53–61, July 2016.

[OM15] OMG: , Meta Object Facility (MOF) Version 2.5, 2015.

[7] `http://descartes.tools`