EXTENDED ABSTRACT

To guarantee the vision of Quality of Service (QoS) different goals in terms of SLAs have to be dynamically met between the Cloud provider and the customer (Breskovic et al., 2013). This SLA enactment should involve little human-based interaction in order to guarantee the scalability and efficient resource utilization of the system. To achieve this we start from Autonomic Computing, examine the autonomic control loop and adapt it to govern Cloud Computing infrastructures. We propose an approach to manage Cloud infrastructures by means of Autonomic Computing. The basic structure of the autonomic systems is represented by a control loop that monitors (M) Cloud parameters, analyses (A) them, plans (P) actions and executes (E) them; the full cycle is known as MAPE. MAPE-K loop stores knowledge (K) required for decision-making in a knowledge base (KB) that is accessed by the individual phases. This talk addresses the research question of finding a suitable KM system (i.e., a technique of how stored information should be used) and determining how it interacts with the other phases for dynamically and efficiently allocating resources.

We first hierarchically structure all possible adaptation actions into so-called escalation levels (Chimaobi et al., 2011). We then focus on one of these levels by analysing monitored data from virtual machines and making decisions on their resource configuration with the help of knowledge management (KM). The monitored data stems both from synthetically generated workload categorized in different workload volatility classes and from a real-world scenario: scientific workflow applications in bioinformatics. As KM techniques, we investigate two methods, Case-Based Reasoning (CBR) and a rule-based approach. We design and implement both of them and evaluate them with the help of a simulation engine. Simulation reveals the feasibility of the CBR approach and major improvements by the rule-based approach considering SLA violations, resource utilization, the number of necessary reconfigurations and time performance for both, synthetically generated and real-world data. Both approaches are evaluated with the real world data from the life science domain by using the traces of RNA sequencing. RNA sequencing represent a new challenging approach for the evaluation and profiling of the gene structures.

With the rapid development of high-throughput technologies in recent years, huge amounts of data are being generated and stored in databases in the field of life science, which necessitates significant advances in computing capacity and performance. RNA sequencing has the potential to transform how gene structure and gene expression profiling are studied. Scientific workflow applications are crucial in enabling scientists to determine important information from those huge amounts of stored data. Existing workflow applications are process- or data-, rather than resource oriented. Thus, they lack efficient computational resource management capabilities, such as those provided by Cloud computing environments. Insufficient computational resources disrupt the execution of workflow applications, wasting time and money. To address this issue, advanced resource monitoring and management strategies are required to determine the resource consumption behaviors of workflow applications for a dynamical allocation and deallocation of resources (Maurer et al., 2013).

Thus, we utilize the knowledge management strategies for Clouds (rule and CBR-based) to manage computational resources for workflow applications in order to guarantee their performance goals and their successful completion (Maurer et al., 2013).
REFERENCES


BRIEF BIOGRAPHY

Dr. Ivona Brandić is Assistant Professor at the Distributed Systems Group, Information Systems Institute, Vienna University of Technology (TU Wien). Prior to that, she was Assistant Professor at the Department of Scientific Computing, Vienna University. She received her PhD degree from Vienna University of Technology in 2007. From 2003 to 2007 she participated in the special research project AURORA - Advanced Models, Applications and Software Systems for High Performance Computing and the European Union's GEMSS - Grid-Enabled Medical Simulation Services project. She is involved in the European Union's SCube project and she is leading the Austrian national FoSII - Foundations of Self-governing ICT Infrastructures project funded by the Vienna Science and Technology Fund (WWTF). She is management committee member of the European Commission's COST Action on Energy Efficient Large Scale Distributed Systems. From June to August 2008 she was visiting researcher at the University of Melbourne, Australia.

In 2011 she received the Distinguished Young Scientist Award from the Vienna University of Technology for her HALEY project on Holistic Energy Efficient Hybrid Clouds. Her interests comprise Service Level Agreement and Quality of Service management in large scale distributed systems, autonomic computing, workflow management for scientific applications, and energy efficient large scale distributed systems (Cloud, Grid, Cluster, etc.). She published more than 50 scientific journal, magazine and conference publications and co-authored a text book on federated and self-manageable Cloud infrastructures. I. Brandić co-authored European Union's Cloud Computing report paving future research directions of the EU. In 2010 she chaired the International Conference on Utility and Cloud Computing held in Chennai, India. She has been serving more than 50 program committees (among others EuroPar, COMPSAC, CloudCom) and was invited reviewer of more than 10 international journals. In 2011 she edited two special issues for Future Generation Computer Systems (Elsevier) and Scientific Programming Journal (IOS Press). I. Brandić has been invited expert evaluator of the European Commission, French National Research Organization (ANR), National Science and Engineering Research Council Canada (NSERC) and Netherlands Organization for Scientific Research (NWO).