SPEC Research Group's Cloud Working Group

[RG Cloud Group]

Alexandru Iosup*
Distributed Systems Group
Delft University of Technology
The Netherlands
A.Iosup@tudelft.nl

Samuel Kounev†
Descartes Software
Engineering Group
University of Würtzburg
Germany
Smauel.Kounev@uni-wuertzburg.de

Kai Sachs
SAP SE
Walldorf
Germany
Kai.Sachs@sap.com

Keywords

Cloud WG; RG Cloud Group; SPEC Research Group; benchmarking; performance evaluation

1. INTRODUCTION

Classical performance evaluation and benchmarking has produced a variety of methodologies, techniques, and tools that support the design and development of systems, the procurement of IT capacity, performance engineering and tuning of operational infrastructure and applications, etc. However, new challenges in performance evaluation and benchmarking appear every year. The Research Group of the Standard Performance Evaluation Corporation (SPEC) consists of several dedicated working groups to address these challenges, by topic. In this poster, we introduce the scope, membership, activities, and an exemplary outcome of the working group focusing on challenges associated with cloud computing (RG Cloud Group).

The mission of the RG Cloud Group of SPEC is "taking a broad approach, relevant for both academia and industry, to cloud benchmarking, quantitative evaluation, and experimental analysis [...] This group focuses on novel cloud properties such as elasticity, performance isolation, dependability, and other non-functional system properties, in addition to classical performance-related metrics such as response time, throughput, scalability, and efficiency."

Current participants in the RG Cloud Group include the Delft University of Technology (Delft), the IBM T.J. Watson Research Center (USA), Lund University (Sweden), MITRE (USA), Oracle (USA), Salesforce.com (USA), SAP (Germany), Tata TCS (India), Umea University (Sweden), and the University of Wuertzburg (Germany).

Other working group are part of the SPEC Research Group, and are often collaborating with each other. The Big Data Working Group addresses the challenges of volume, variety, and veracity, and possibly also "'V"'s, by "specifying and classifying big data systems, developing rules and tools for big data benchmarking, and fostering collaboration between benchmarking efforts". The IDS Benchmarking Working Group addresses the crucial arising challenge of intrusion detection of security in datacenters and virtualized environments. The DevOps Performance Working Group addresses the challenges of "combining application performance management (APM) and model-based software performance engineering (SPE) activities for business-critical application systems".

ICPE'16 March 12-18, 2016, Delft, Netherlands

ACM ISBN 978-1-4503-4080-9/16/03.

DOI: http://dx.doi.org/10.1145/2851553.2858675

2. SCOPE OF THE RESEARCH EFFORTS

The scope of the group is "to develop new methodological elements for gaining deeper understanding not only of cloud performance, but also of cloud operation and behavior, through diverse quantitative evaluation tools, including benchmarks, metrics, and workload generators".

Developing concepts and translating them into quantitative evaluation tools covers work on measurement of a diverse set of cloud characteristics and situations, such as elasticity, and performance variability or isolation, respectively; but also on profiling and even workload characterization. Collecting and sharing operational traces from cloud systems is also part of the scope of this group, resulting not in open-source software but in open-access data artifacts, such as the Grid Workloads Archive [4] and the Failure Trace Archive [5].

The group focuses on a broad understanding of the term performance, which includes both classical and *new* performance properties, such as response times and throughput, scalability and elasticity/auto-scaling, resource- and energy-efficiency; and classical and *new* dependability-related properties, such as availability, reliability, but also various risk characteristics and metrics.

The group specializes in collaborative work on evaluation prototypes and on facilitating joint research on topics related to performance, but members also develop full-blown implementations in their extensive research. For example, the Descartes Software Engineering group has developed a variety of tools, including the Descartes Modeling Language, the LIMBO Load Intensity Modeling Tool [8], the BUNGEE Cloud Elasticity Benchmark [3], and the Queueing Petri net Modeling Environment; the Delft University of Technology's group is developing the Graphalytics [1] graph analytics

benchmarking and monitoring tool for Big Data Platform-as-a-Service and Infrastructure-as-a-Service clouds; etc.

3. REPOSITORY OF PEER-REVIEWED TOOLS

Through the SPEC Research Group, the RG Cloud Group is publishing SPEC-endorsed tools addressing recurring issues in quantitative cloud evaluation and analysis. Among them,

BUNGEE is a Java-based framework focusing on cloud elasticity, especially for IaaS cloud platforms and auto-scaling environments. The tool provides load and stress-testing functionality, and automates the analysis of the quality of the elastic behavior of the system under test through several elasticity metrics [2]. Currently, BUNGEE supports CloudStack- and Amazon AWS-based deployments. The group is currently extending BUNGEE to support more cloud environments and metrics (see Section 4).

LIMBO is an Eclipse-based tool for creating, managing, and using load-intensity models. LIMBO supports different arrival rates and processes, and can be used for example for generating time series of user requests for benchmarking, or re-scaling existing traces for "what-if" scenarios.

4. THREE ACTIVE TOPICS

The RG Cloud Group members define their own joint activities. Among the research-oriented activities of the group, we detail here three:

Cloud Usage Patterns (CUPs): The goal of this activity is to define a formalism for expressing cloud usage patterns and scenarios. The joint authors, who represent seven organizations, propose [6] a simple yet expressive textual and visual formalism, which can be used by both general users and cloud experts. A key feature of the textual formalism is its conciseness; this goes in contrast to other formalisms that also focus on the executability of the specification. By expressing over ten patterns commonly seen in academic and industrial practice, the authors show that CUP is practical.

Cloud Metrics Survey and Design: This ongoing activity focuses on surveying existing cloud metrics and on the design of key missing metrics that allow the quantitative assessment and characterization of typical cloud usage scenarios. Among the key new metrics, the joint authors focus on various forms of elasticity [2] and risk-quantifying metrics [7].

Benchmarking Auto-Scaling Techniques: This activity, which various members of the RG Cloud Group are just starting, is aiming to conduct a quantitative analysis and comparison of auto-scaling techniques in virtualized environments.

The RG Cloud Group also has various presentation and service activities. It maintains a web site[1], and helps with organizing various workshops and conferences, among which the flagship ACM/SPEC ICPE conference.

[1]https://research.spec.org/working-groups/rg-cloud-working-group.html

5. CONCLUSION

The RG Cloud Group of SPEC is an active inter-organizational research group focusing on all aspects of modern performance evaluation and benchmarking of cloud environments.

The group is actively looking for new members, to jointly develop benchmarking tools such as BUNGEE and LIMBO; to join ongoing activities such as surveying and designing cloud metrics, and benchmarking auto-scaling techniques; but also to propose new and exciting new activities.

Acknowledgments

The authors would like to thank their collaborators in the Cloud WG, and their funding agencies and organizations.

6. REFERENCES

[1] M. Capota, T. Hegeman, A. Iosup, A. Prat-Pérez, O. Erling, and P. A. Boncz. Graphalytics: A big data benchmark for graph-processing platforms. In *Proceedings of the Third International Workshop on Graph Data Management Experiences and Systems, GRADES 2015, Melbourne, VIC, Australia, May 31 - June 4, 2015*, pages 7:1–7:6, 2015.

[2] N. R. Herbst, S. Kounev, and R. H. Reussner. Elasticity in cloud computing: What it is, and what it is not. In *10th International Conference on Autonomic Computing, ICAC'13, San Jose, CA, USA, June 26-28, 2013*, pages 23–27, 2013.

[3] N. R. Herbst, S. Kounev, A. Weber, and H. Groenda. BUNGEE: an elasticity benchmark for self-adaptive iaas cloud environments. In *10th IEEE/ACM International Symposium on Software Engineering for Adaptive and Self-Managing Systems, SEAMS 2015, Florence, Italy, May 18-19, 2015*, pages 46–56, 2015.

[4] A. Iosup, H. Li, M. Jan, S. Anoep, C. Dumitrescu, L. Wolters, and D. H. J. Epema. The grid workloads archive. *Future Generation Comp. Syst.*, 24(7):672–686, 2008.

[5] B. Javadi, D. Kondo, A. Iosup, and D. H. J. Epema. The failure trace archive: Enabling the comparison of failure measurements and models of distributed systems. *J. Parallel Distrib. Comput.*, 73(8):1208–1223, 2013.

[6] A. Milenkoski, A. Iosup, S. Kounev, K. Sachs, P. Rygielski, J. Ding, W. Cirne, and F. Rosenberg. Cloud usage patterns: A formalism for description of cloud usage scenarios. *CoRR*, abs/1410.1159, 2014. Extended article, with new core concepts, under submission.

[7] V. van Beek, J. Donkervliet, T. Hegeman, S. Hugtenburg, and A. Iosup. Self-expressive management of business-critical workloads in virtualized datacenters. *IEEE Computer*, 48(7):46–54, 2015.

[8] J. von Kistowski, N. R. Herbst, and S. Kounev. LIMBO: a tool for modeling variable load intensities. In *ACM/SPEC International Conference on Performance Engineering, ICPE'14, Dublin, Ireland, March 22-26, 2014*, pages 225–226, 2014.