



SPEC ResearchSM Group Newsletter

CONTENTS

- 2 SPEC RG Officers and Working Groups
- 3 Welcome to the SPEC RG Newsletter
- 3 Announcements
- 3 ICPE 2018: Statistics
- 3 ICPE 2019 in Mumbai - Preliminary Announcement
- 4 Reports of the Working Groups
- 7 ContinulTY: New Research Project
- 8 Towards a Serverless Benchmark

ICPE 2019 WILL BE HELD IN MUMBAI, INDIA

Varsha Apte and Antinisca Di Marco, the General Chairs of the next ACM/SPEC International Conference on Performance Engineering (ICPE 2019), invite interesting high-quality submissions. The conference will take place April 6-11, 2019 in Mumbai, India.

Read more on page 3

FIVE SPEC RESEARCH WORKING GROUPS REPORT ON THEIR PROGRESS

The SPEC Research Working Groups IDS, Cloud, Big Data, DevOps Performance and Power report on their progress, articles, benchmarks, and technical reports published in the year 2018. The Working Groups are always open for new members, feel invited to join us!

Read more on pages 4-7

SPEC KAIVALYA DIXIT DISTINGUISHED DISSERTATION AWARD 2017

This year, the selection committee has chosen to select two dissertations: a winner, Aleksandar Milenkoski of the University of Würzburg, and a runner up, Benjamin Heintz of University of Minnesota Twin Cities, based on the high quality of both submissions.

Read more on page 3



CONTACT

Standard Performance Evaluation Corporation (SPEC)
7001 Heritage Village Plaza, Suite 225
Gainesville, VA 20155, USA

SPEC Research Group

Chair: Samuel Kounev (rgchair@spec.org)

Web: <http://research.spec.org>

TOWARDS A SERVERLESS BENCHMARK

Following a durable trend of miniaturization of software services, there is an emergence of an architectural style called serverless for cloud-based software that focuses on executing functions hiding resource management. The goal is to design and release an initial version of a benchmark for serverless computing platforms.

Read more on page 8

SPEC RESEARCH GROUP OFFICERS

Chair:

Samuel Kounev, University of Würzburg, Germany

Vice-Chair:

André van Hoorn, University of Stuttgart, Germany

Secretary:

Anna Queralt, Barcelona Supercomputing Center, Spain

Steering Committee:

J. Nelson Amaral, University of Alberta, Canada

Alexandru Iosup, TU Delft, NL

Samuel Kounev, University of Würzburg, Germany

Klaus-Dieter Lange, HPE, USA

Meikel Poess, Oracle Corporation, USA

Anna Queralt, Barcelona Supercomputing Center, USA

Tilmann Rabl, Bankmark UG, Germany

Petr Tůma, Charles Univ. of Prague, Czech Republic

André van Hoorn, University of Stuttgart, Germany

Boris Zibitsker, BEZNext, USA

Publicity Officer:

Christoph Heger, NovaTec Consulting GmbH, Germany

Release Manager:

Vojtěch Horký, Charles Univers. Prague, Czech Republic

Newsletter Editors:

Nikolas Herbst, **André Bauer**, University of Würzburg, Germany

<http://research.spec.org>

SPEC RESEARCH WORKING GROUPS

Cloud Working Group

Chair:

Alexandru Iosup, TU Delft, NL

Vice-Chair and Release Manager:

Nikolas Herbst, University of Würzburg, Germany

<https://research.spec.org/en/working-groups/rg-cloud.html>

SPEC, the SPEC logo and the names SERT, SPEC SFS, SPECjbb, SPECvirt.sc, Chauffeur WDK, and SPEC PTDaemon are trademarks of the Standard Performance Evaluation Corporation. The SPEC Research Logo and the name SPEC Research are service marks of SPEC. Additional company, product and service names mentioned herein may be the trademarks or service marks of their respective owners. Copyright ©1988-2018 Standard Performance Evaluation Corporation (SPEC). Reprinted with permission. All rights reserved.

IDS Benchmarking Working Group

Chair:

Marco Vieira, University of Coimbra, Portugal

Vice-Chair:

Alberto Avritzer, Sonatype, USA

Secretary and Release Manager:

Lukas Iffländer, University of Würzburg, Germany

<https://research.spec.org/working-groups/rg-ids-benchmarking.html>

Big Data Working Group

Chair:

Tilmann Rabl, bankmark, Germany

Vice-Chair:

Todor Ivanov, Goethe University Frankfurt, Germany

Secretary:

Meikel Poess, Oracle Corporation, USA

Release Manager:

John Poelman, IBM, USA

<https://research.spec.org/en/working-groups/rg-big-data.html>

DevOps Performance Working Group

Chair:

André van Hoorn, University of Stuttgart, Germany

Vice-Chair:

Cor-Paul Bezemer, Queen's University, Canada

Robert Heinrich, KIT, Germany

Secretary:

Jürgen Walter, University of Würzburg, Germany

Release Manager:

Vincenzo Ferme, Università della Svizzera Italiana, Switzerland

<https://research.spec.org/en/working-groups/rg-devops-performance.html>

Power Working Group

Chair:

Jóakim von Kistowski, University of Würzburg, Germany

Vice-Chair:

Klaus-Dieter Lange, HPE, USA

Secretary:

Mike Tricker, Microsoft, USA

<https://research.spec.org/working-groups/rg-power.html>

WELCOME TO THE SPEC RESEARCH GROUP NEWSLETTER

We are delighted to present to you the next issue of the SPEC Research Group Newsletter. This regular publication provides information on latest developments, news, and announcements relevant to the benchmarking and quantitative system evaluation communities. Our newsletter is part of our mission to foster the exchange of knowledge and experiences between industry and academia in the field of quantitative system evaluation and analysis. The past year has been very intense and exciting for the SPEC RG. To the major activities and milestones reached, we include:

- 8th ACM/SPEC **ICPE 2017** in L'Aquila, Italy
- 14th IEEE International Conference on Autonomic Computing **ICAC 2016** in Columbus, Ohio
- 3rd International Workshop on Quality-aware Dev-Ops **QUDOS 2017** in L'Aquila, Italy
- 1st Workshop on Hot Topics in Cloud Computing Performance **HotCloudPerf 2018** at ICPE 2018
- SPEC RG face-to-face **meeting 2017**

We have been actively working on preparation, planning and organization of the ICPE 2018. We hope that the vivid exchange of ideas during the upcoming ICPE 2018 will be a great motivation for the next year of scientific and engineering work.

We hope that you will enjoy reading the newsletter. We welcome and encourage your contributions for articles and suggestions for future coverage.

André Bauer, Nikolas Herbst
(Newsletter Editors, University of Würzburg).

SPEC KAIVALYA DIXIT DISTINGUISHED DISSERTATION AWARD 2017 WINNER AND RUNNER UP

The Selection committee has chosen this year to select a winner and a runner up based on the high quality of both submissions. The winning dissertation – Evaluation of Intrusion Detection Systems for Virtualized Environments – was authored by Aleksandar Milenkoski of Julius-Maximilians-Universität (JMU) Würzburg, under the supervision of Prof. Dr. Samuel Kounev. Given the high quality of dissertations nominated for this award, the committee decided to publicly recognize another dissertation titled "Optimizing Timeliness, Accuracy, and Cost in Geo-Distributed Data-Intensive Computing Systems" authored by Benjamin Heintz under the supervision of Prof. Abhishek Chandra at the University of Minnesota Twin Cities as the runner-up.

<https://research.spec.org/news/single-view/article/winner-of-spec-kaivalya-dixit-distinguished-dissertation-award-2017.html>

PLAN FOR NEW WORKING GROUP

The University of Würzburg plans to initiate a new Working Group in the field of Predictive Data Analytics. Methods in scope learn from historical data to predict future developments. In more detail, the interest lays in the assessment and development of predictive methods (either classic or machine learning), combining of existing methods, and designing of new measures/metrics for quantifying the performance of predictions. In the context of machine learning, the group is also interested in feature engineering and feature selection. The scope can be divided into three categories: (i) pure forecasting, (ii) workload forecasting in the context of cloud computing, and (iii) predictive maintenance, failure prediction and root cause analysis in the context of Industry 4.0/Smart Factory. If you are interested, please contact André Bauer (andre.bauer@uni-wuerzburg.de).

ICPE 2018: STATISTICS

With all different tracks, the 9th ACM/SPEC International Conference on Performance Engineering (ICPE 2018), being held in Berlin, Germany from April 9 to 13, 2018, attracted a number of high-quality submissions selected in intensive review processes by different track committees.

In the research track, 14 out of 59 papers were accepted as full papers. Two full papers received an ACM artifact badge after the subsequent review process in the newly introduced artifact evaluation track. 7 submissions were accepted as short research papers. In the industry/experience track, 4 out of 16 papers were accepted as full papers. 6 submissions were accepted as short papers. In the work-In-progress/vision track, 10 out of 23 papers were accepted. In addition to those scientific papers, the technical program features three keynotes, three tutorials, the presentation of the SPEC Kaivalya Dixit Distinguished Dissertation Award, a poster and demonstration session, and six workshops. Details about the conference and its program are provided at <http://icpe2018.spec.org/>.

André van Horn (University of Stuttgart) and
Katinka Wolter (Freie Universität Berlin)

ICPE 2019 IN MUMBAI - PRELIMINARY ANNOUNCEMENT

The ACM/SPEC International Conference on Performance Engineering (ICPE) provides a forum for the integration of theory and practice in the field of performance engineering. It brings together researchers and industry practitioners to share ideas, discuss challenges, and present results of both work-in-progress and state-of-the-art research on performance engineering of software and systems.

ICPE 2019 will be held in Mumbai (India) from April 6 to 11. For a long time, Mumbai (previously 'Bombay') has been regarded as India's most 'cosmopolitan' city. It is the financial capital of the country and the home of 'Bollywood'. The conference will take place in the Victor Menezes Convention Center within the IIT Bombay Campus. IIT Bombay is worldwide recognized as a leader in the field of engineering education and research. Reputed for the outstanding calibre of students graduating from its undergraduate and postgraduate programs, the Institute attracts the best students from the country for its bachelor's, masters and doctoral programs.

IIT Bombay is a small township in itself. Rich in natural flora and fauna to begin with, the campus' green cover has been maintained and even increased over the years. An island of green in the otherwise concrete jungle that is Mumbai, the campus at Powai nestles among hills and is flanked by the Powai and Vihar lakes. It is special both in terms of its physical beauty and in terms of location, a place where you can be away from the busy world, yet still be a part of it. Education and research are the twin pillars of this institute, the ambience is one in which new ideas, and creativity can flourish.

ACM India is organizing the conference locally. The contact person for ICPE 2019 is Varsha Apte from IIT Bombay, who will be General Co-Chair along with Antinisca Di Marco from University of L'Aquila, Italy.

The PC Co-Chairs will be Greg Franks from Carleton University, Ottawa, Canada and Marin Litoiu from York University, Canada. The industrial track chair will be David Schmidt, HP Enterprise, Houston, USA.

Varsha Apte (IIT Bombay)

REPORT: POWER RESEARCH AND SPECPOWER COMMITTEE

Since the Power Research working group has been started in 2017, it operates tightly coupled with the OSG SPECpower Committee.

In December the Server Efficiency Rating Tool (SERT) suite 2.0.1 was released to address changes to the metric that shipped in 2.0.0, along with an online tool to upgrade v2.0.0 results to be compatible with those from v2.0.1. The final SERT metric was added for the use in government server efficiency programs in the U.S., European Union, Japan, South Korea and China, which are considering the adoption of the SERT suite for several server efficiency programs.

The SPECpower Committee is now able to focus on developing the replacement for the SPECpower_ssj2008 power benchmark. Internal development kits of the SPECpower-

2018 benchmark are already available, including a new GPGPU workload. These kits are intended to validate the new SERT-derived infrastructure and to try out innovative ideas in structuring a benchmark. The SPECpower sub-committee welcomes contributions of suitable code, and resources to help with development and testing.

The SPEC RG Power Working Group has already contributed significantly to the work of the SPECpower Committee. Jóakim von Kistowski of the University of Würzburg provided significant input to creating the metric for the SERT suite 2.0.0, including writing papers that explain and validate the underlying mathematics. The prototype of the GPGPU workload for SPECpower2018 was developed by his colleagues, and Jóakim is working closely with the committee to create the "Measuring and Benchmarking Power Consumption and Energy Efficiency" tutorial to be presented at ICPE 2018 in Berlin.

Klaus-Dieter Lange (HPE)

<https://research.spec.org/working-groups/rg-power.html>

REPORT: IDS WORKING GROUP

The research group on intrusion detection systems (RG IDS) is preparing to broaden its focus in the field of benchmarking security of different types of systems. This will include the interaction with other security systems, the placement inside complex security architectures and the possibilities and risks associated with new trends like network function virtualization (NFV) and software-defined networking (SDN).

New to the group is Lukas Iffländer, a Ph.D. student from the University of Würzburg. The current working title for his Ph.D. thesis is "Attack-aware Security Network Function Management". Thus, he focuses on the ordering of network security functions inside the chain between the incoming traffic and the protected service. The order is to be changed depending on the composition of the incoming traffic. It is optimized to assure the best performance with the available resources. The paper's [1] abstract can be found below:

Services provided online are subject to various types of attacks. Security appliances can be chained to protect a system against multiple types of network attacks. The sequence of appliances has a significant impact on the efficiency of the whole chain. While the operation of security appliance chains is currently based on a static order, traffic-aware reordering of security appliances may significantly improve efficiency and accuracy. In this paper, we present the vision a self-aware system to automatically reorder security appliances according to incoming traffic. To achieve this, we propose to apply a model-based learning, reasoning, and acting (LRA-M) loop. To this end, we

describe a corresponding system architecture and explain its building blocks. In the upcoming weeks, RG IDS will work on updating its focus and structure. The results will be presented in the next issue of this newsletter.

Lukas Iffländer (University of Würzburg)

<https://research.spec.org/working-groups/rg-ids-benchmarking.html>

[1] L. Iffländer, J. Walter, S. Eismann, and S. Kounev. The Vision of Self-aware Reordering of Security Network Function Chains (Vision Paper). ACM/SPEC ICPE 2018, Berlin, Germany, April 9–13, 2018.

REPORT: CLOUD WORKING GROUP

In 2017, the SPEC RG Cloud Group has pursued through a broad palette of activities its long-term mission of furthering cloud benchmarking, quantitative evaluation, and experimental analysis, in directions relevant for both academia and industry. Through our mission, we focus 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. 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’. Among the keywords most discussed in the group, over 2017, are serverless computing and function-as-a-service, cloud metrics, cloud experiment methodology, elasticity and autoscaling. In 2017, through monthly online meetings facilitated by WebEx and SPEC as well a numerous activity meetings, we have advanced work on five main topics and a number of other sub-topics:

1) Cloud Metrics Survey and Design [1]: This established 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 we have introduced in 2016, the joint authors focus on various forms of elasticity and risk-quantifying metrics. An extended version of the technical report [1] is under review at the ACM Transactions on Modeling and Performance Evaluation of Computing Systems (TOMPECS) journal.

2) Cloud Experiment Methodology: This activity can be seen as a spin-off of the previously mentioned topic and started in spring 2017. The idea is to review how performance experiments are conducted and reported in leading research venues over the past years. The group defines a number of principles that should be fulfilled for reliable and reproducible experiment results. Meanwhile, this new activity resulted in a submission to a top-tier conference.

3) Autoscaling: We achieved two publications: (i) Our work on benchmarking autoscalers for complex workflows has been accepted to appear in ACM TOMPECS. (ii) A cost-aware mechanism for enhancing autoscalers, called FOX was accepted as full paper at the ICPE 2018 [3].

4) Serverless Computing: This activity started in May 2017 motivated by the emerging trend of executing arbitrary functions without much server- and resource-management burden in cloud environments. A detailed description of this activity can be found on page 8 in a separate section “Towards a Serverless Benchmark”.

5) Reach-out and dissemination: Our group has been active in disseminating its knowledge and artifacts to industry and academia. The working group initiates the 1st Workshop on Hot Topics in Cloud Computing Performance (HotCloudPerf 2018) [4] featuring 6 accepted papers and 4 talks complemented by a joint panel discussion with WOSP-C.

To conclude, 2017 was a full and successful year for the RG Cloud Group. We are looking forward to an even more successful 2018. For this, we are actively seeking new participants and activities.

Nikolas Herbst (University of Würzburg)

<http://research.spec.org/working-groups/rg-cloud.html>

[1] N. Herbst, R. Krebs, G. Oikonomou, G. Kousiouris, A. Evangelinou, A. Iosup, and S. Kounev. Ready for Rain? A View from SPEC Research on the Future of Cloud Metrics. Technical Report SPEC-RG-2016-01, SPEC Research Group — Cloud Working Group, Standard Performance Evaluation Corporation (SPEC), 2016.

[2] A. Ilyushkin, A. Ali-Eldin, N. Herbst, A. Bauer, A. V. Papadopoulos, D. Epema, and A. Iosup. An Experimental Performance Evaluation of Autoscalers for Complex Workflows. ACM Transactions on Modeling and Performance Evaluation of Computing Systems (ToMPECS), 2018, ACM. To Appear.

[3] V. Lesch, A. Bauer, N. Herbst, and S. Kounev. FOX: Cost-Awareness for Autonomic Resource Management in Public Clouds. ACM/SPEC ICPE 2018, Berlin, Germany, April 9-13, 2018.

[4] HotCloudPerf 2018 web site: <https://hotcloudperf.spec.org/>

REPORT: BIG DATA WORKING GROUP

The term “big data” was coined only recently. However, the development of systems that are able to transform massive amounts of data into useful information for decision making date back to the 1960s when model-oriented decision support systems (DSS) were first introduced. In the early days, DSS were implemented as isolated, specialized applications that were used by few companies that could afford a custom DSS implementation. Nowadays, DSS are implemented using general purpose database management systems (DBMS) running on commodity hardware that are connected to multiple data

sources. The ability to quickly and accurately integrate data from these data sources became paramount. Data integration (aka. ETL) systems were developed to facilitate this process. Still, these DSS are only connected to a small number of data source. With network technology evolving and becoming widely available, the use of DSS is extending to the Internet of Things (IoT), where data is sent from millions of devices in many cases at a very high frequency. Increasingly, the variation in data format and content is challenging the traditional relational database model. To mitigate the overhead in frequent model changes, more flexible forms of data processing and analysis have been developed, spear-headed by the map reduce paradigm.

Having benchmarks to measure the performance of these systems is very important, because without them data analysis solution providers can make unverifiable marketing claims about system performance. Since the late 80's industry standard benchmark consortia such as the Standard Performance Evaluation Corporation (SPEC) and the Transaction Processing Performance Council (TPC) have developed and maintained industry standard benchmarks, which enable all solution providers to compete on a level playing field and to allow for fair performance comparisons amongst solution providers. TPC developed TPC-D, TPC-H, TPC-R and TPC-DS to measure the performance of standalone DSS, TPC-DI for data integration and TPCx-IoT for the Internet of Things. SPEC has a broader scope, featuring benchmarks for many different forms of computation, however, with no data analytics related benchmark so far. However, seeing the pressing needs and strong trends in data analytics, the SPEC Research Working Group on Big Data was formed to fill the gaps that the TPC, which is focused more on relational databases, is not covering.

While big data processing has been around for a couple of years now, recently, many efforts have been focused especially on two topics that benefit from the advancements in big data processing: stream processing and deep learning. In stream processing, data is continuously evaluated rather than waiting for a larger chunk of data before processing. This enables much lower latency. Deep learning is very useful for approaching problems with complex solution spaces, which are hard to impossible to hardcode into a system. However, for both topics there exist no industry standard benchmark from either TPC or SPEC. This is why the SPEC Research Group Working Group on Big Data has started efforts to collect the state of the art in stream processing and deep learning benchmarking in research and industry and is aiming to fill these gaps with new benchmark proposals.

The group holds weekly meetings, with internal calls and open, public calls alternating. For the open calls, presen-

tations on big data benchmarking, big data systems, performance tuning, and related research are invited. These calls are well attended and received. Presented topics in 2016 were BigBench, SparkBench, OpenML, and Aloja. Presentations can be found on the working group's website.

Tilmann Rabl (bankmark) and Meikel Poess (Oracle)
<https://research.spec.org/en/working-groups/rg-big-data.html>

REPORT: DEVOPS PERFORMANCE WORKING GROUP

The DevOps Performance Working Group focuses on how to address performance concerns in DevOps. DevOps aims to increase the velocity and frequency of bringing software changes into stable production. These goals are achieved by a tighter integration of development (Dev) and operations (Ops), as well as a high degree of automation (e.g., via continuous delivery/deployment).

The DevOps Performance Working Group fosters and facilitates research in combining model-based and measurement-based APM (application performance management) and SPE (software performance engineering) activities for DevOps, e.g., by experience sharing, agreement on definitions, specification of metrics, and dissemination of novel methods, techniques, and tools for quantitative evaluation. We aim to identify cross-community collaboration, and to set the path for long-lasting collaborations towards performance-aware DevOps.

As the major joint activity in 2017, we performed a survey on the current state of performance in DevOps. Our study reveals that automatic performance evaluations are usually not integrated into automatic delivery pipelines and not performed regularly. In addition, performance modeling is not applied in most companies. The results are currently under review for publication in a renowned software engineering magazine. Moreover, we continued the development of our reference platform for performance evaluation [1].

Several sub-projects among subsets of the members have been conducted in 2017. Selected technical efforts include: (i) the evaluation of SLAs applying measurement and model-based evaluation within a single framework [2], (ii) automated decision support to select performance evaluation approaches that is adaptable on evolution of approaches [3], and (iii) and performance engineering for microservices [4]. Moreover, a report on the GI-Dagstuhl seminar 16394, "Software Performance Engineering in the DevOps World" co-organized and attended by several group members has been published [5]. Moreover, collaborations included jointly supervised student projects.

In April 2017, the group met for its third face-to-face meeting, co-located with the 8th ACM/SPEC International Conference on Performance Engineering (ICPE 2017) in L'Aquila, Italy and had a joint discussion with the other SPEC RG working groups there. The next meeting is scheduled for the 9th ACM/SPEC International Conference on Performance Engineering (ICPE 2018) in Berlin, Germany.

The working group co-organized the third edition of the International Workshop on Quality-Aware DevOps (QUDOS 2017) [6]. QUDOS 2017, technically supported by the SPEC RG, was co-located with ICPE 2017. The workshop comprised 12 talks (including an industrial and an academic keynote), 10 scientific papers, and around 25 registered participants. The fourth QUDOS edition will be co-located with ICPE 2018. Also, the Symposium on Software Performance (SSP 2017) [7] was co-organized by members of the working group. The program included 16 talks (including an industrial and an academic keynote), 10 papers, and more than fifty registered participants.

The group met on a biweekly basis to discuss current activities and new project ideas, including invited presentations. A number of new members and member organizations joined the group. In 2017, researchers from more than ten organizations worldwide attended regularly, including Canada, Ecuador, Germany, Switzerland, and United States. In total, 18 meetings were held, including talks about survivability modeling of mission critical systems, benchmark requirements for microservices architectures [8], Docker-based performance benchmarking and load testing using BenchFlow [9], accomplishments and challenges for performance regression analysis, and joint projects of the group. As a result of discussions, a research visit for 2018 is planned that will apply BenchFlow in the context of continuous and declarative analysis at the University of Stuttgart.

Targeted topics for 2018 include efficient continuous regression testing, automated (and incremental) performance model learning, declarative performance engineering, and performance visualization.

For more information about the DevOps Performance Working Group (including our mission, activities, meetings, presentations, and projects), please visit our web page [10]. If you are interested in following the discussions or contributing actively, please get in touch with the working group chairs.

Jürgen Walter (Uni. of Würzburg), André v. Hoorn (Uni. of Stuttgart)
<https://research.spec.org/en/working-groups/rg-devops-performance.html>

[1] T. F. Düllmann, R. Heinrich, A. v. Hoorn, T. Pitakrat, J. Walter, and F. Willnecker. Caspa: A platform for comparability of architecture-based software performance engineering approaches. IEEE ICSA 2017. IEEE,

2017.

[2] J. Walter, D. Okanovic, and S. Kounev. Mapping of Service Level Objectives to Performance Queries. WOSP-C 2017 at ACM/SPEC ICPE 2017. ACM, 2017.

[3] J. Walter, A. v. Hoorn, and S. Kounev. Automated and Adaptable Decision Support for Software Performance Engineering. EAI Valuetools 2017.

[4] R. Heinrich, A. v. Hoorn, H. Knoche, F. Li, L. E. Lwakatare, C. Pahl, S. Schulte, and J. Wettinger. Performance engineering for microservices: Research challenges and directions. ACM/SPEC ICPE 2017 Companion, pages 223–226, New York, NY, USA, 2017. ACM.

[5] Report from GI-Dagstuhl Seminar 16394: Software Performance Engineering in the DevOps World. <https://arxiv.org/pdf/1709.08951.pdf>.

[6] International Workshop on Quality-aware DevOps (QUDOS). <http://qudos-workshop.org/>.

[7] Symposium on Software Performance (SSP). <http://www.performance-symposium.org/>.

[8] C. M. Aderaldo, N. C. Mendonça, C. Pahl, and P. Jamshidi. Benchmark requirements for microservices architecture research. IEEE/ACM ECASE 2017, pages 8–13, 2017.

[9] Benchflow software. <https://github.com/benchflow>.

[10] RG DevOps Website. <https://research.spec.org/devopswg>.

CONTINUITY: NEW RESEARCH PROJECT

Modern businesses require their application systems to perform in a responsive and cost-effective way. To develop such systems, performance regressions have to be identified early during the software development process, and especially before deployment. However, practice shows that load testing, as a relevant form of performance (regression) testing, is either not performed adequately or not at all. Load tests are usually not representative enough for the actual usage profile in production. This situation is caused by the fact that they require additional effort and necessary competence to create load test scripts. In addition, these scripts become obsolete very quickly due to system and usage profile changes, and their maintenance is associated with high costs. In the context of emerging continuous software development (DevOps), in which software changes are being put into operation ever faster and more frequently, this becomes even a greater problem. Furthermore, because the execution of these tests requires a longer time, it is not practical to carry out a load test for each change, e.g., to detect performance regressions. However, if regressions are detected in these aggregated changes, manual diagnosis is necessary to determine the responsible change in the software or the usage profile.

The Continuity research project, executed by NovaTec Consulting GmbH and the University of Stuttgart, aims to ensure automated and efficient load testing by integrating continuously recorded measurement data from production, and to integrate it into the continuous software

development environment. Load tests are automatically extracted from APM (application performance management) data, and evolved taking into account changes in the usage profile. To specify load tests, we use a modular description language, which can be extended with additional contextual information, e.g., with test type and objectives. As a part of the automation of the software development process (continuous delivery), relevant load tests are selected to be executed in testing. Obtained results are used for the detection of regressions and their diagnosis. ContinulTy examines regressions to find whether they are caused by changes in the implementation or the usage profile, and provides results to the DevOps team.

The 2-year project started in September 2017. It is sponsored by the German Federal Ministry of Education and Research (BMBF), and conducted in the context of the SPEC RG DevOps Performance Working Group. Case studies for the evaluation of project results are provided by two associated application partners: Daimler FleetBoard GmbH and Alberto Avritzer as an independent expert. For more information about the project visit our web page and get in touch with us.

Dušan Okanović, André van Hoorn (University of Stuttgart), Henning Schulz, Christoph Heger (NovaTec Consulting GmbH)

ContinulTy web site: <http://continuity-project.github.io>

H. Schulz, T. Angerstein, A. van Hoorn: Towards automating representative load testing in continuous software engineering. LTB 2018 at ACM/SPEC ICPE 2018, ACM (2018)

C. Heger, A. van Hoorn, M. Mann, D. Okanović: Application performance management: State of the art and challenges for the future. ACM/SPEC ICPE 2017, ACM

TOWARDS A SERVERLESS BENCHMARK

Following a durable trend of miniaturization and commodification of software services, there is an emergence of an architecture (model) for cloud-based software that focuses on executing arbitrary functions without much server- and resource-management burden put on the cloud developer or customer. This model, which is now commonly associated with emerging terms such as serverless and Function-as-a-Service (FaaS), addresses needs for which PaaS and SaaS leave many aspects unanswered or insufficiently addressed.

There are several reasons for the rapid emergence of serverless computing. The cloud user is unburdened of the need to manage resources, designing autoscaling policies, and other operational logic. This allows them to focus on their business logic; decreasing the costs of development, removing the need for extensive distributed systems expertise, and improving the time-to-market for applications. On the other hand, the cloud provider has

more control (controlling the entire lifecycle) and insight into the users workload. This provides the provider with options to improve scheduling and resource management, minimizing the resource costs of workloads.

In May 2017, the SPEC RG Cloud formed a new activity, lead by Erwin van Eyk and Alexandru Iosup, to investigate this emerging paradigm. The activity initially focused on answering two primary questions: (1) What challenges and issues exist in the current state of serverless computing? (2) What role can the SPEC RG Cloud have in answering (some of) these challenges? Following an initial period of pitching the activity, an initial team was formed, lead by Erwin van Eyk and Alexandru Iosup. After these initial nine months, the inter-disciplinary team has grown to span both industry and academia, with active participants from the TU Delft, VU Amsterdam, University of Würzburg, ESPOL (Equador), IBM, and SAP.

To answer these questions, we first surveyed the landscape in serverless computing, formulating a consistent terminology (e.g., what is serverless? What is Function-as-a-Service?), learning about the current state-of-the-art in both industry and academia, and determining the current perspectives and challenges [1]. With this clear view of the serverless field, we found that the SPEC RG CLOUD should focus primarily on performance evaluation, creating a much-needed benchmark of serverless platforms. To, potentially, other related performance (engineering) challenges, for which we formed a separate, more specific vision [2].

Our main goal for the coming year is to design and release an initial version of the 'serverless benchmark'; a benchmark of the numerous open-source and closed-source FaaS platforms. As we are progressing towards a serverless benchmark, we are currently working on a reference architecture for FaaS platforms to help identify common components that can be compared across platforms and providers and provide a terminology for the internals of these platforms. To this end, we invite the SPEC Newsletter audience to join this effort; your input, insight, workloads, and knowledge of closed-source implementations could help us shape this field.

Erwin van Eyk (Platform9 & Vrije Universiteit Amsterdam)

[1] E. v. Eyk, A. Iosup, S. Seif, and M. Thömmes. The SPEC Cloud Group's Research Vision on FaaS and Serverless Architectures, International Workshop on Serverless Computing, 1–4. ACM, 2017.

[2] E. v. Eyk, A. Iosup, C. Abad, S. Eismann, J. Grohmann. A SPEC RG Cloud Group's Vision on the Performance Challenges of FaaS Cloud Architectures, ACM/SPEC ICPE 2018. ACM.