

CONTENTS

- 2 SPEC RG Officers and Working Groups
- 3 Welcome to the SPEC RG Newsletter
- 3 Kaivalya Dixit Distinguished Dissertation Award
- 4 ICPE 2026: Statistics
- 4 ICPE 2027 in Gothenburg
- 5 Working Group Reports
- 11 Graduated SPEC RG Members
- 11 Selected Abstracts

SPEC KAIVALYA DIXIT DISTINGUISHED DISSERTATION AWARD 2025

The award selection committee for 2025 selected the dissertation of *Mohammad Bakhshalipour* from Carnegie Mellon University, US. The selection committee appreciated the dissertation for its strong technical depth, originality, and relevance to both academic research and industrial practice. The thesis bridges robotics and computer architecture by introducing comprehensive, open-source benchmark suites and performance studies for real-time robotic applications.

Read more on page 3

ICPE 2027 WILL BE HELD IN GOTHENBURG

The ICPE organizing committee and local general chair Philipp Leitner invite interesting high-quality submissions for the next ACM/SPEC International Conference on Performance Engineering (ICPE 2027). The conference will take place on May 24-28, 2027 in Gothenburg, Sweden.

Read more on page 5

SPEC RESEARCH WORKING GROUPS REPORT ON THEIR PROGRESS

The SPEC Research Working Groups DevOps Performance, Cloud, Security, Power, and Predictive Data Analytics report on their progress, articles, benchmarks, and technical reports published in 2025. The Working Groups are always open for new members. Feel invited to join!

Read more on pages 5-11

CONTACT



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

SPEC Research Group (SPEC RG)

Chair: Samuel Kounev (rgchair@spec.org)

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

GRADUATED LONG-TERM MEMBERS OF SPEC RG

In the past year, long-term SPEC RG member Sacheendra Talluri received his PhD. SPEC RG thanks Sacheendra for his work in the last years. Find a summary of his work in this newsletter.

Read more on page 11

SPEC RESEARCH GROUP OFFICERS

Chair:

Samuel Kounev, University of Würzburg, Germany

Vice-Chair:

Cristina Abad, ESPOL University, Ecuador

Secretary:

Diego Elias Costa, Concordia University, Canada

Steering Committee:

Cristina Abad, ESPOL University, Ecuador

Diego Elias Costa, Concordia University, Canada

Nikolas Herbst, University of Würzburg, Germany

Alexandru Iosup, VU Amsterdam, NL

Samuel Kounev, University of Würzburg, Germany

Klaus-Dieter Lange, HPE, USA

Alex Podelko, Amazon, USA

Weiyi Shang, University of Waterloo, Canada

Michele Tucci, University of L'Aquila, Italy

Petr Tůma, Charles Univ. of Prague, Czech Republic

Publicity Officer:

Weiyi Shang, University of Waterloo, Canada

Release Manager:

Michele Tucci, University of L'Aquila, Italy

Newsletter Editor:

Matthijs Jansen, VU Amsterdam, The Netherlands

Jesse Donkervliet, VU Amsterdam, The Netherlands

<https://research.spec.org>

SPEC RESEARCH WORKING GROUPS

DevOps Performance Working Group

Chair:

Heng Li, Polytechnique Montréal, Canada

Vice-Chair:

Diego Elias Costa, Concordia University, Canada

Michele Tucci, University of L'Aquila, Italy

Secretary:

Lizhi Liao, University of Guelph, Canada

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

Cloud Working Group

Chair:

Alexandru Iosup, VU Amsterdam, The Netherlands

Vice-Chair:

Nikolas Herbst, University of Würzburg, Germany

Secretary:

Sacheendra Talluri, VU Amsterdam, The Netherlands

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

Power Working Group

Chair:

Maximilian Meissner, University of Würzburg, Germany

Vice-Chair:

Klaus-Dieter Lange, HPE, USA

Secretary:

Aaron Cragin, Microsoft, USA

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

Security Working Group

Chair:

Thomas Prantl, University of Würzburg, Germany

Vice-Chair:

Aleksandar Milenkoski, SentinelLabs/Univ. Würzburg, Germany

Secretary:

Lukas Horn, University of Würzburg, Germany

Release Manager:

Simon Engel, University of Würzburg, Germany

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

Predictive Data Analytics Working Group

Chair:

André Bauer, Illinois Institute of Technology, USA
University of Würzburg, Germany

Vice-Chair:

Robert Leppich, University of Würzburg, Germany

Secretary:

Veronika Lesch, DHBW Mosbach, Germany

Release Manager:

Michael Stenger, University of Würzburg, Germany

Daniel Grillmeier, University of Würzburg, Germany

<https://research.spec.org/working-groups/rg-predictive-data-analytics.html>

WELCOME TO THE SPEC RESEARCH GROUP NEWSLETTER

With over 120 members in 30 countries and nearly two dozen benchmarks spanning highly diverse aspects of computing performance and energy efficiency, SPEC has become known as a beacon of truth for computing researchers, vendors, users, and analysts worldwide. These professionals rely on SPEC to ensure that the marketplace has a fair and useful set of metrics to differentiate computing systems. Founded in 2011, the SPEC Research Group is proud of being part of this remarkable history.

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.

Some highlights from the last year include:

- 16th ACM/SPEC **ICPE 2025** in Toronto, Canada
- 6th IEEE International Conference on Autonomic Computing and Self-Organizing Systems **ACSOS 2025** in Tokyo, Japan
- 8th Workshop on Hot Topics in Cloud Computing Performance **HotCloudPerf 2025** at ICPE 2025

We have been actively working on the preparation, planning, and organization of ICPE 2026, which will be held in Florence, Italy. We hope that a vivid exchange of ideas will be a great motivation for the next year of scientific and engineering work.

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

Samuel Kounev (SPEC Research Chair, University of Würzburg).
Matthijs Jansen (Newsletter Editor, VU Amsterdam).
Jesse Donkervliet (Newsletter Editor, VU Amsterdam).

SPEC, the SPEC logo and the names SERT, SPEC SFS, SPECjbb, SPECvirt, Chauffeur, and PTDaemon are registered 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 ©1988-2026 Standard Performance Evaluation Corporation (SPEC). Reprinted with permission. All rights reserved.

SPEC KAIVALYA DIXIT DISTINGUISHED DISSERTATION AWARD 2025 WINNER

The SPEC Kaivalya Dixit Distinguished Dissertation Award is an annual award that aims to recognize outstanding doctoral dissertations within the scope of the SPEC Research Group in terms of scientific originality, scientific significance, practical relevance, impact, and presentation.

The winning dissertation *Bridging Real-Time Robotics and Computer Architecture* was authored by *Mohammad Bakhshalipour* from Carnegie Mellon University under the supervision of Phillip B. Gibbons. The selection committee appreciated the dissertation for its strong technical depth, originality, and relevance to both academic research and industrial practice. The thesis bridges robotics and computer architecture by introducing comprehensive, open-source benchmark suites and performance studies for real-time robotic applications, and by proposing efficient application-specific hardware accelerators and domain-specific processors that demonstrate substantial performance improvements through targeted architectural design and empirical evaluation.

The award will be presented at the 17th ACM/SPEC International Conference on Performance Engineering (ICPE), to be held in May 2026.

Given the high quality of dissertations nominated for this award, the committee decided to publicly recognize another dissertation as Runner-Up. The runner-up for the 2025 award is Yintong Huo from The Chinese University of Hong Kong. Her dissertation, “Log-Driven Automated Software Reliability Engineering,” was supervised by Michael Lyu.

The award selection committee for 2025 was chaired by Prof. Weiyi Shang of University of Waterloo.

The SPEC Kaivalya Dixit Distinguished Dissertation Award was established in 2011 to recognize outstanding doctoral dissertations within the scope of the SPEC Research Group. Contributions of interest include the design of metrics for system evaluation and the development of methodologies, techniques, and tools for measurement, benchmarking, profiling, workload characterization, dependability, and efficiency evaluation of computing systems.

Dissertations defended **between October 2024 and September 2026** will be eligible to be nominated for the 2026 award (a thesis can be nominated only once).

<https://research.spec.org/news/2026-01-21-00-00-winner-of-spec-kaivalya-dixit-distinguished-dissertation-award-2025/>

ICPE 2026: STATISTICS

The 17th ACM/SPEC International Conference on Performance Engineering (ICPE 2026) will be held in Florence, Italy, from May 4 to May 8, 2026. This year, the Research Track was highly competitive, receiving over 100 abstracts that resulted in 81 full submissions—the highest number of submissions since 2017. Following an initial screening, a small number of submissions were desk-rejected for being out of scope or violating double-anonymity requirements. Each of the remaining papers was reviewed by at least three PC members, followed by a thorough discussion phase that included two online meetings. Based on the reviewers' recommendations and these deliberations, we have accepted or conditionally accepted 23 papers (28% acceptance rate), with an additional 5 papers recommended for acceptance as short contributions. Of the 11 submissions to the Industry Track, 7 were accepted as full articles. In the emerging research track, 10 out of 17 submissions were accepted as emerging research papers. There were 9 submissions to the Artifact Track. Of these 9 submissions, 6 received the functional badge and the remaining 5 received the reusable badge. Two of the submissions requested to be included in the SPEC research tools repository, and are currently being reviewed by the respective board. For the sixth time, ICPE features a Data Challenge with 3 accepted short papers out of 4 submissions. 5 tutorials were proposed, out of which 3 were accepted. The Poster and Demo Track received 10 submissions, out of which 5 were accepted as posters and 2 as demonstrator papers. For the third time, ICPE features a Journal First track, with 1 accepted paper that has already been published in a selected journal and will be presented at the conference.

The following workshops are planned for ICPE 2026:

- The Fourth International Workshop on Performance Optimization in the LLM World (AIPerfLLM 2026)
<https://sites.google.com/view/icpe-workshop-aiperf-2026>
- The Sixth Workshop on Education and Practice of Performance Engineering (WEPPE 2026)
<https://esulabsolutions.godaddysites.com/sponsored-events>
- The Ninth Workshop on Hot Topics in Cloud Computing Performance (HotCloudPerf 2026)
<https://hotcloudperf.spec.org/>
- The Fourteenth International Workshop on Load Testing and Benchmarking of Software Systems (LTB 2026)
<https://ltb2026.github.io/>
- The Fifth Workshop on Systems and Service Quality (QualITA 2026)
<https://qualitawg.github.io/>



The city of Gothenburg, Sweden.

The following distinguished speakers will give keynotes at ICPE 2026:

- Leana Golubchik, (University of Southern California, US) **Systems for AI: Predicting Performance of Machine Learning Workloads**
- Didem Unat (Koç University, Turkey) **Illuminating Multi-GPU Communication Paths**
- Jeff Hammond (NVIDIA) **State-of-the-Art Communication Software for Supercomputers and Its Applications**

Roberto Verdecchia (University of Florence, Italy).

Catia Trubiani (Gran Sasso Science Institute, Italy).

Ana Lucia Verbanescu (University of Twente, The Netherlands).

Radu Calinescu (University of York, United Kingdom).

ICPE 2027 IN GOTHENBURG, SWEDEN 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 2027 will be held in Gothenburg, Sweden, between May 24 and May 28, 2027. Gothenburg is a port city on the west coast of Sweden, with a rich tradition in shipping and industry. With a metropolitan population of over 1 million people, Gothenburg is Sweden's second-largest city. These days, the city is known as a Scandinavian innovation hub, and the home of Volvo, SKF, and AstraZeneca. End of May, Gothenburg promises mild temperatures and long days (sunset after 10PM), perfect for an extended hike in Sweden's famed nature, visiting one of dozens of lakes in the Gothenburg region, or a boat trip to Göteborgs skärgård, the city's archipelago.

The contact person and local general chair for ICPE 2027 is Philipp Leitner, Chalmers University of Technology. The PC chairs and organizing committee will be announced later.

REPORT: DEVOPS PERFORMANCE WORKING GROUP

DevOps has become one of the main principles for engineering and operating software systems. It aims to increase the rate and velocity of releasing new software versions, which is, for instance, achieved by a high degree of automation and by integrating development and operations responsibilities. DevOps imposes immense challenges for quality assurance, e.g., concerning performance and related attributes. Key reasons are that respective activities are constrained by time and that the environment in which a software system runs is ever-changing. On the other hand, DevOps provides great opportunities because the integration between development and operations allows for a high degree of automation and a streamlined collection and analytics of performance data.

The RG DevOps Performance Working Group is a forum for individuals and organizations interested in the interplay of DevOps and performance engineering. The mission of the working group is to consolidate concepts and tools to better integrate these activities. Its membership body currently includes representatives of Amazon Web Services, BEZNext, Concordia University, MongoDB, Polytechnique Montreal, University of Alberta, University of L'Aquila, University of Guelph, University of Hamburg, University of Stuttgart, University of Waterloo, University of Würzburg, and Wuhan University.

The group as a whole meets in online meetings that are held on a monthly basis. In addition to the discussion of organizational topics, these general meetings include a technical presentation by group members or invited guests. In total, eight group meetings were held in 2025, including the following ten talks as part of our monthly lecture series:

- Jan** “ShuffleBench: A Benchmark for Large-Scale Data Shuffling Operations with Distributed Stream Processing Frameworks” by Sören Henning (Johannes Kepler University Linz)
- Feb** “Smells-sus: Sustainability Smells in IaC” by Mohammad Hamdaqa (Polytechnique Montréal)
- Mar** “Enhancing DPDK Observability: A Tracing-Based Performance Analysis Approach” by Adel Belkhir (Polytechnique Montréal)
- Apr** “Component-Based Analytical Modeling of GPU Runtime Performance: a Case-Study in Scientific Computing” by Jolly Chen (University of Twente)
- May** “Understanding the Energy Consumption of Cloud-native Software” by Understanding the Energy Consumption of Cloud-native Software (University of Groningen)
- Jun** “Using e-divisive means change detection in continuous benchmarking” by Henrik Ingo (Nyrkiö)

Sep “TracePicker: Optimization-Based Trace Sampling for Microservice-Based Systems” by Shuaiyu Xie (Wuhan University)

Nov “Seeing Beyond Traces: Code- and Data-Centric Attribution in Microservices” by Probir Roy (University of Michigan-Dearborn)

In addition to the monthly meetings, the group operates in subgroups consisting of 4-6 participants who collaborate closely on concrete topics. Collaborations include joint research papers and jointly supervised student projects. The subgroups meet biweekly and report to the whole group once a month in the regular meeting. The current subgroups are:

1. **Performance change point detection:** This subgroup focuses on the challenges of using performance detection methods in real large-systems performance data. In collaboration with three major tech companies, SAP, MongoDB, and Oracle, this subgroup has established the first curated dataset of real performance changes and assessed the quality of current change detection methods. The group is submitting a paper to the ACM Transactions on Software Engineering and Methodology (TOSEM) in March.
2. **Configuration-aware performance modeling.** This subgroup aims to investigate performance modeling methods for highly configurable systems. The current work is on studying the relationship between software workloads and configurations in order to develop a novel technique for comprehensively modeling the performance of highly configurable systems. A paper (titled “HQP: Hybrid Queueing and Configuration-Aware Performance Modeling for Microservice Systems”) reporting the current work was submitted to the ACM Transactions on Software Engineering and Methodology (TOSEM) and is currently under revision.
3. **Energy-aware DevOps.** This subgroup investigates challenges in energy-aware DevOps practices, focusing on the measurement, monitoring, and optimization of energy efficiency in SDLC. The current work focuses on evaluating the reliability of existing energy measurement tools (e.g., Intel RAPL) and developing a tool (hardware + software) for reliable measurement of software energy consumption. The subgroup is currently wrapping up a paper that reports the empirical findings.

For more information about the DevOps Performance Working Group (including our mission, activities, meetings, presentations, and projects), please visit our web page at:

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

If you are interested in following the discussions or contributing actively, please contact the working group chairs.

Heng Li (Polytechnique Montreal),
Diego Elias Costa (Concordia University),
Michele Tucci (University of L'Aquila),
Lizhi Liao (University of Guelph)

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

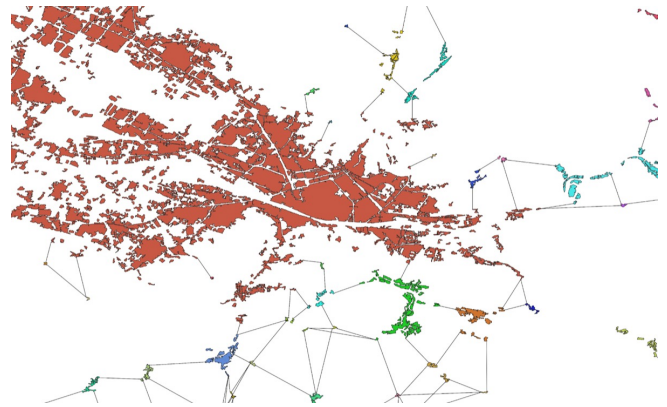
REPORT: CLOUD WORKING GROUP

In 2025, the SPEC RG Cloud Group has driven several activities aligned with its long-term **mission** of furthering cloud benchmarking, quantitative evaluation, and experimental analysis in directions relevant to both academia and industry. This year, we have focused on topics around digital performance twin, workflow scalability in the domain of earth observation, evolution and caching in the computing continuum. We were able to conclude a multi-year effort on benchmarking serverless systems with a journal publication.

The group's **scope** is to develop new methodological elements for gaining a deeper understanding of performance and operations across the compute continuum, from the cloud to the user. We develop quantitative benchmarks and performance measurement tools, enabling novel large-scale applications to be developed. We consider 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. Our work towards benchmark prototypes includes designing reference architectures, standardizing use cases, observing patterns, and methods for reproducibility.

In 2025, through monthly online meetings facilitated by SPEC's Zoom and meetings focusing on furthering specific activities, and through continuous discussion via a Slack workspace, we have advanced work on the following main topics:

1. **Digital Performance Twin:** Datacenters are essential to our modern digital societies, but simultaneously raise important operational concerns. Specifically, designing and operating datacenters to meet SLOs such as performance and sustainability goals are important challenges. This activity addresses these challenges through the design of datacenter digital twins: datacenter simulators that take real-world telemetry data as input and control datacenter operations (with a human in the loop) as output. Over the past year, we made two important steps towards addressing this challenge. First, we have started work on OpenDT, our open-source digital twin, with promising initial results. Our system calibrates predictions in real time based on differences between predictions and real-world measurements,



The SDA settlement network around the city of Florence

reducing prediction error (MAPE) from 7.86% in state of the art to 4.39% in our work [1]. Second, we design and implement STEAM, a datacenter simulator that helps stakeholders investigate the individual and combined impact of sustainability techniques on datacenter operational and embodied carbon emissions [2]. Using STEAM, we obtain surprising findings: for example, we find that temporal shifting is less effective than previously reported

2. **Serverless Execution of the SDA Workflow:** Earth Observation (EO) workflows are often developed as a single software artifact encapsulating its processing steps and requiring a monolithic deployment. With the Settlement Delineation and Analysis (SDA) workflow [3] we developed a component-based EO workflow which computes, clusters and visualizes a settlement network from a World Settlement Footprint (WSF) raster image originally inferred from satellite data. Figure 2 features the SDA settlement network around the city of Florence which hosts the ICPE conference this year. With the component-based nature of the SDA workflow, it is also applicable for FaaS deployment in addition to the traditionally monolithic setup since every workflow step can be deployed as a serverless function.

By integrating the SDA workflow into the Serverless Benchmarking Suite (SeBS) [4] we can deploy the workflow's processing steps as AWS Lambda functions and orchestrate the workflow using a state machine (Amazon Step Functions). We compare the monolithic deployment on a local machine orchestrated by a custom scheduler with the new FaaS deployment in AWS. While the monolithic deployment could achieve significantly shorter time-to-results, the FaaS is more cost-effective. Further details regarding the deployment, experimental setup and measurements can be found in our contribution to HotCloudPerf 2026: From Monolith to Serverless: Running a Component-Based Scientific Workflow Across Execution Paradigms.

3. **Cloud Continuum and Edge Computing:** The activity has been working on a white paper retracing the evolution of the computing continuum, which will be available soon. In this direction, we aim to classify the capabilities of the continuum, evaluate the existing challenges across various application domains, and identify the disparities between current capabilities and future requirements. By drawing insights from previous computing models, we will recommend research directions aimed at addressing these challenges. Another current topic of interest is data caching strategies in the computing continuum. Efficient cache placement is critical for optimizing performance in computing continuum environments, where edge, fog, and cloud nodes collaborate to deliver content. In light of this, we are developing a cache placement strategy that leverages edge and fog nodes to alleviate complex optimization tasks from network nodes to intermediate fog nodes. This strategy gathers request statistics along the data delivery pathway and utilizes a multi-criteria cost function to determine the optimal content for caching. In the future, we plan to start exploring novel methods for performance analysis of Generative AI in the computing continuum.
4. **Serverless Benchmarking Evolution:** Over the past decade, serverless computing has evolved from simple function execution to complex distributed applications. Despite this growth, the field still lacks a standardized benchmarking approach, which leads to fragmented evaluations and hinders consistent cross-platform performance analysis. To address this, our activity is currently finalizing a vision paper that examines the evolution of serverless benchmarks to identify key limitations and major successful lessons. We are proposing design principles for a comprehensive, future-proof benchmark suite that emphasizes modularity, platform agnosticism, and adjustability to the system's size. We advocate for an evolutionary benchmark design that can accommodate emerging technologies while maintaining comparability across platforms, ultimately aiming to unite the community to drive reproducible research.
5. **Concluding ServiBench/Let's trace it FGCS article [5]:** ServiBench was one of RG Cloud's early projects to benchmark serverless systems. We are happy to report its conclusion and acceptance for publication in FGCS. We investigated serverless applications which involved multiple services such as serverless functions, databases, cloud storage, and ML services. Measuring the performance of these applications required us to augment the cloud providers's distributed tracing framework to combine disjoint tracing spans. This work uncovered the performance decrease due to inefficient coordi-

nation mechanisms. It also revealed that high cold start time was due to application startup and not container startup.

Besides these focused activities, Cloud WG has been successfully organizing and growing the yearly workshop HotCloudPerf 2026. On May 5, the workshop will be held in Florence, Italy, co-located with ICPE 2026, and feature two keynotes, three accepted full workshop paper presentations and two talks.

Keynote presentations:

- Predictable and energy-efficient execution of distributed Cloud services: a 2-decade research overview by **Tommaso Cucinotta**, Associate Professor, Sant'Anna School of Advanced Studies, Pisa, Italy
- Osmotic Computing in Practice: Adoption Challenges and Emerging Reference Scenarios by **Antonino Galletta**, Assistant Professor, University of Messina, Italy

Paper presentations:

- OpenDT: Exploring Datacenter Performance and Sustainability with a Self-Calibrating Digital Twin by R. Nicolae, J. van der Toorn, H. Liu, S. Kraniti, A. Iosup
- From Monolith to Serverless: Running a Component-Based Scientific Workflow Across Execution Paradigms by L. Gruber, N. Herbst, M. Copik, S. Kounev
- Performance Optimization in Stream Processing Systems: Experiment-Driven Configuration Tuning for Kafka Streams by D. Chen, S. Henning, K. Matteussi, R. Rabiser

Talk presentations:

- Exploring Power and Latency Trade-offs in DPDK Packet Forwarding by J. Carlon, T. Cucinotta, G. Lettieri
- An Analysis of HPC and Edge Architectures in the Cloud by S. Santillan, C. Abad

The Cloud working group congratulates Sacheendra Talluri, a multi-year active group member, on successfully defending his doctoral studies "On the Design and Characteristics of Cloud Computing Systems Along the Performance-Failure Continuum". He is now at Clickhouse. For more insights on his work, please check the Graduated PhDs section.

To conclude, this was a successful year for the RG Cloud Group. We started new activities in 2025. We look forward to them succeeding in 2026. For this, we are actively seeking new participants. You can also propose new activities that would be of interest to members.

Alexandru Iosup (Vrije Universiteit Amsterdam, Chair),
Nikolas Herbst (University of Würzburg, Vice-Chair)

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

-
- [1] R. Nicolae, J. van der Toorn, S. Kraniti, H. Liu, and A. Iosup, "OpenDT: Exploring datacenter performance and sustainability with a self-calibrating digital twin," in *Companion of the 2026 ACM/SPEC International Conference on Performance Engineering, ICPE 2026, Florence, Italy, May 4-8, 2026*, ACM, 2026.
- [2] D. Niewenhuis, S. Talluri, A. Iosup, and T. de Matteis, "STEAM: Realistic modeling and systematic exploration of composable techniques for sustainable datacenters," in *Proceedings of the 17th ACM/SPEC International Conference on Performance Engineering, ICPE 2026, Florence, Italy, May 4-8, 2026*, ACM, 2026.
- [3] L. Gruber, N. Herbst, P. Friedl, T. T. Nguyen, T. Esch, and S. Kounev, "The FISHNET case study on implementing and scaling a complex earth observation workflow," in *IEEE International Conference on eScience, eScience 2025, Chicago, IL, USA, September 15-18, 2025*, IEEE, 2025, pp. 159–168. doi: [10.1109/ESCIENCE65000.2025.00027](https://doi.org/10.1109/ESCIENCE65000.2025.00027). [Online]. Available: <https://doi.org/10.1109/eScience65000.2025.00027>.
- [4] M. Copik, G. Kwasniewski, M. Besta, M. Podstawski, and T. Hoefler, "Sebs: A serverless benchmark suite for function-as-a-service computing," in *Middleware '21: 22nd International Middleware Conference, Québec City, Canada, December 6 - 10, 2021*, K. Zhang, A. Gherbi, N. Venkatasubramanian, and L. Veiga, Eds., ACM, 2021, pp. 64–78. doi: [10.1145/3464298.3476133](https://doi.org/10.1145/3464298.3476133). [Online]. Available: <https://doi.org/10.1145/3464298.3476133>.
- [5] J. Scheuner et al., "Let's trace it: Fine-grained serverless benchmarking for synchronous and asynchronous applications," *Future Gener. Comput. Syst.*, vol. 179, p. 108336, 2026. doi: [10.1016/J.FUTURE.2025.108336](https://doi.org/10.1016/J.FUTURE.2025.108336). [Online]. Available: <https://doi.org/10.1016/j.future.2025.108336>.

REPORT: SECURITY WORKING GROUP

The security group looks back on a year of upheaval. With changes in membership and participation, the group is shifting increasingly from the intrusion detection and hypervisor security domain to the privacy-preserving machine learning domain. More on that later.

In April 2025, our chair Thomas Prantl successfully defended his PhD thesis. The remainder of the working group congratulates him on this outstanding achievement. He continues to work at the University of Würzburg as a postdoctoral researcher.

Building on the success of the 2024 publication of "De Bello Homomorphico" [1] the group published the follow-up paper "Quo Vadis Homomorphico" in the Journal of Information Security and Applications [2]. It compares the previously proposed methods against state-of-the-art polynomials and explains when to use each technology.

Given the importance of privacy-preserving technologies, our group members Tobias Schneider and Thomas Prantl developed a lecture for Master's students on Homomorphic Encryption. Following the principle of design-based research, the lecture was presented and evaluated at three different universities, and the material was published at the EAI Conference on Computer Science and Education in Computer Science [3].

In the process of no longer solely focusing on homomorphic encryption, the group applied secure multi-party computation to a medical CNN. The evaluation showed

that the approach is feasible, with inference times in the domain of minutes, thus providing a reasonable trade-off between privacy and performance. The results were published at ICSC [4].

Nevertheless, we remain active in the traditional topics of the group and published the article "Benchmarking Hyper-Breakpoints for Efficient Virtual Machine Introspection" in the Journal of Electronics [5]. With it, we proposed the bpbench benchmarking tool to quantify the performance of breakpoints in Virtual Machine Introspection (VMI). The benchmark measures the overhead that executing and hiding a breakpoint introduces, as well as the overhead that toggling breakpoint states on context switches adds.

Using this tool, we published "EPT Switching vs. Instruction Repair vs. Instruction Emulation: A Performance Comparison of Hyper-Breakpoint Variants" in the Journal of Engineering [6]. We utilized bpbench to measure four different VMI breakpoint implementations across 20 Intel CPUs from the 4th to the 13th generation. The fastest processor achieved a breakpoint execution time of 15 μ s using instruction emulation.

The group developed two new visions for its future research directions. The first is "Self-Aware Performance Optimization" While there is now a large set of privacy-preserving technologies, each has its strengths and weaknesses, and there is no clear "best" solution. Furthermore, it requires significant domain knowledge to find the most suitable one for a project. We therefore envision self-awareness: the technologies themselves should be able to decide the right implementation variant to reach optimal performance. There are case studies where switching between homomorphic encryption and secure multi-party computation at the right times significantly improves performance.

The second vision is a component of the first. We aim to develop a domain-specific language for privacy-preserving neural networks, serving two purposes. The first is to lower the hurdles for applying such technologies and make them available to non-security researchers. The second is to automatically leverage technology synergies. The language shall be able to generate several solutions for a given PPML problem so that the user can then select the one that suits them best.

To summarize, the SPEC RG Security Benchmarking Working Group continues to push forward in several research directions, including:

- Proposing a representative benchmark approach to evaluate mathematical functions in homomorphic cryptosystems
- Developing new metrics to classify the performance of function approximation in homomorphic encryption

- Implementing a framework to perform homomorphic neural network inference
- Exploring switching capabilities for transitioning between homomorphic encryption and secure multi-party computation, as well as within these technologies
- Identifying challenges and applications in the area of boolean and integer homomorphic cryptosystems
- Exploring the foundations of homomorphic training of neural networks
- Working towards a comparative benchmark of current PPML techniques like homomorphic encryption, secure multi-party computation, and functional encryption
- Working towards a domain-specific language for inference of neural networks using homomorphic encryption, secure multi-party computation, and functional encryption
- Exploring if homomorphic encryption and functional encryption can be combined in a suitable way

The SPEC RG Security WG is looking forward to its biggest goal for the coming year: establishing a new benchmark for homomorphic encryption. The group is happy to accept new members who would like to contribute or bring new visions to the general area of security.

Thomas Prantl (University of Würzburg),
Aleksandar Milenkoski (SentinelLabs),
Lukas Horn (University of Würzburg),
Simon Engel (University of Würzburg)

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

- [1] T. Prantl et al., “De Bello Homomorphico: Investigation of the extensibility of the OpenFHE library with basic mathematical functions by means of common approaches using the example of the CKKS cryptosystem,” *International Journal of Information Security*, vol. 23, no. 2, pp. 1149–1169, Apr. 2024, ISSN: 1615-5270. DOI: [10.1007/s10207-023-00781-0](https://doi.org/10.1007/s10207-023-00781-0). Accessed: Aug. 5, 2025.
- [2] T. Prantl, L. Horn, S. Engel, A. Bauer, and S. Kounev, “Quo Vadis CKKS: Comparison of the realization of basic mathematical functions for the homomorphic cryptosystem CKKS using De Bello and polynomial approximations,” *Journal of Information Security and Applications*, vol. 93, p. 104 110, Sep. 2025, ISSN: 2214-2126. DOI: [10.1016/j.jisa.2025.104110](https://doi.org/10.1016/j.jisa.2025.104110). Accessed: Aug. 5, 2025.
- [3] T. Prantl et al., “Oratio homomorphico: Creating a lecture for homomorphic encryption by applying design-based research,” in *Proceedings of the 4th EAI Conference on Computer Science and Education in Computer Science*, ser. EAI CSECS 2025, Springer, Jul. 2025.
- [4] T. Prantl et al., “Privacy-preserving training and inference of cnns: A medical performance case study,” in *2025 5th Intelligent Cybersecurity Conference (ICSC)*, IEEE, May 2025, pp. 303–311. DOI: [10.1109/ICSC65596.2025.11140461](https://doi.org/10.1109/ICSC65596.2025.11140461).
- [5] L. Beierlieb, A. Schmitz, R. Springer, C. Dietrich, and L. Iffländer, “Benchmarking hyper-breakpoints for efficient virtual machine introspection,” *Electronics*, vol. 14, no. 3, 2025, ISSN: 2079-9292. DOI: [10.3390/electronics14030534](https://doi.org/10.3390/electronics14030534).
- [6] L. Beierlieb, A. Schmitz, A. Karazon, A. Leinweber, and C. Dietrich, “Ept switching vs. instruction repair vs. instruction emulation: A performance comparison of hyper-breakpoint variants,” *Eng.*, vol. 6, no. 10, 2025, ISSN: 2673-4117. DOI: [10.3390/eng6100278](https://doi.org/10.3390/eng6100278).

REPORT: PREDICTIVE DATA ANALYTICS WORKING GROUP

The mission of the group is to bridge the missing links between the facets involved in data analytics, namely big data storage and provisioning, data versioning, and performance evaluation. The group’s ambition is to standardize and benchmark the entire data lifecycle, i.e., the analytics and prediction methods and especially pipelines for data analytics ranging from big data storage and pre-processing to analytics and assessment, as well as to provide heuristics for the selection of tools, patterns, and infrastructure. Furthermore, the group members are interested in promoting interaction between industry and academia by contributing research towards standardization and benchmarking of the different aspects of data analytics. For this, the group members investigate data analytics-related methodologies, systems, and metrics. Another important goal is to support open data and promote the reproducibility of experiments and benchmarking of data analytics methods.

In 2025, our group continued its work on improving the methodological foundations for time series analytics, with a particular focus on evaluation, compression, and the structure of forecasting pipelines. In “TACO Revisited: An In-Depth Analysis of the Lightweight Tree-based Approximate Compression for Time Series” [1], accepted in *Communications in Computer and Information Science* and published in February 2026, we present TACO approach for approximate time series compression. The study revisits the design and performance characteristics of the lightweight tree-based compression technique and evaluates its effectiveness for efficiently storing and processing time series data. In “STEB: In Search of the Best Evaluation Approach for Synthetic Time Series” [2], we investigate different strategies for assessing the quality of synthetic time series and analyze how evaluation choices influence conclusions about synthesis methods. The work aims to provide guidance on selecting appropriate evaluation approaches and highlights the need for consistent and reproducible evaluation practices. In “Decomposing the Time Series Forecasting Pipeline: A Modular Approach for Time Series Representation, Information Extraction, and Projection” [3], we propose a modular perspective on forecasting pipelines that separates representation, information extraction, and projection stages. This decomposition provides a systematic framework for analyzing forecasting approaches and facilitates more transparent comparisons between different methods. The last two works are currently under review, but we have made them publicly available on arXiv.

Besides the above-mentioned activities, the interests of the group lie in but are not limited to:

- Performance modeling, analysis, testing, and prediction

- Performance analysis of ML systems
- Resource autoscaling and reconfigurable systems
- Performance behavior in resource-constrained environments
- Change point and anomaly detection
- Time series analysis, forecasting, and synthesis
- Federated learning
- Streamlining the data science process (DataOps)
- Benchmarking of big data infrastructure

André Bauer (Illinois Institute of Technology, University of Würzburg),
Michael Stenger (University of Würzburg)

<https://research.spec.org/working-groups/rg-predictive-data-analytics/>

- [1] A. Bauer, R. Leppich, and M. Stenger, "Taco revisited: An in-depth analysis of the lightweight tree-based approximate compression for time series," *Communications in Computer and Information Science*, 2026.
- [2] M. Stenger, R. Leppich, A. Bauer, and S. Kounev, "STEB: In Search of the Best Evaluation Approach for Synthetic Time Series," *arXiv preprint arXiv:2505.21160*, May 2025. [Online]. Available: <https://arxiv.org/abs/2505.21160>.
- [3] R. Leppich, M. Stenger, A. Bauer, and S. Kounev, "Decomposing the time series forecasting pipeline: A modular approach for time series representation, information extraction, and projection," *arXiv preprint arXiv:2507.05891*, 2025. [Online]. Available: <https://arxiv.org/abs/2507.05891>.

REPORT: POWER WORKING GROUP

Group Members and Their Mission

The SPEC Research Power Working Group [1] (AMD, Ampere, Dell, HPE, IBM, IEIT, Intel, Microsoft, NVIDIA, University of Wuerzburg; Chair: Maximilian Meissner, Vice-Chair: Klaus-Dieter Lange) has been closely aligned with the SPECpower Committee [2] since its launch in 2016, and later with the SPEC International Standards Group (ISG) [3], launched in 2020.

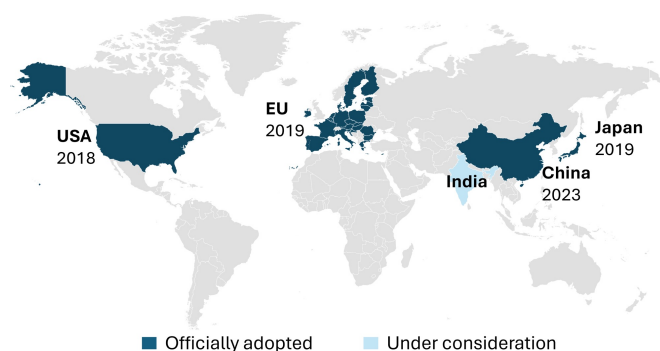
Our mission is to research the energy and resource efficiency of computing devices and software. Our close collaboration strengthens the interaction between industry and academia by contributing research that enhances and promotes methods and tools for evaluating energy and resource efficiency, addressing this essential concern for both sectors as well as regulatory institutions.

10 Year Anniversary – A Decade of Shaping Standardized Benchmarking

To address rising concerns about server power consumption, SPEC members founded the **SPECpower Committee** in 2006. In 2007 the committee released the SPECpower_ssj 2008 benchmark and the SPEC PTDaemon interface as the basis for reliable power measurement. Work on the Chauffeur WDK framework, which became the core of the SERT suite, began in 2009. The initial SERT 1 suite launched in 2013 and received four subsequent upgrades to improve usability and plat-

form support. To broaden engagement and strengthen the research pipeline, the **SPEC RG Power Working Group** was established in 2016 to bridge industry and academia, align academic work with industry needs, foster collaboration and contribute to the evolution of standards, thereby making findings more actionable and impactful. The SPECpower Committee and the RG Power Working Group conducted large-scale experiments to establish and validate the server-efficiency metric used in the SERT 2 suite; key improvements included reduced run time and automated hardware and software discovery to lower test time and cost. SERT 2 suite was officially launched in 2017. In 2020 the **SPEC International Standards Group (ISG)** was formed to oversee the development of standardized benchmarks intended for use in government regulations and programs, and it assumed responsibility for SERT development. The three groups continue to work together to ensure the benchmarks remain accurate, reliable, and relevant.

Development follows a rigorous process of specification, extensive testing across diverse environments, and cross-vendor collaboration to preserve fairness and transparency. Submitted results undergo thorough review, and strict run and reporting rules prevent ambiguities and ensure compliance. As a result of this continuous cycle of refinement and verification, SPEC tools have seen widespread adoption: there are now over 1,000 official SPECpower_ssj 2008 results from over 50 server vendors available, and industry-wide server energy efficiency has improved 63-fold since the benchmark's inception. The SERT suite is included in regulation and certification programs in multiple jurisdictions, including the U.S., the EU, and Japan, and SPEC has reviewed and accepted more than 100 client configurations for SERT to ensure reproducible, low variance, and fair official runs.



SERT 2 Suite adoption in national and regional government regulation and certification programs.

Today, the SERT 2 suite is the only server energy-efficiency tool that is compliant with ISO/IEC 21836:2020, an internationally recognized ISO/IEC standard that facilitates adoption across national regulatory frameworks.

This success was supported by our group's research. During its 10 years, our group published more than 30 conference papers, journals, book chapters, demos, and tutorials, receiving 2 best paper awards, 1 best demo award, and over 1000 citations.

Current and Future Focus

The SPECpower Committee and the SPEC ISG are eagerly working on the next generation of SPEC efficiency benchmarks. Further advancing these industry-standard efficiency benchmarks is critical to account for new trends in hardware and software development.

At ICPE 2026, we present our paper "The Impact of Memory Configuration on Server Efficiency", in which we analyze and demonstrate the extent to which benchmark results, using the SPEC SERT 2 suite, depend on a servers' memory configuration. This study provides important feedback to various stakeholders using efficiency benchmarks:

- **Government regulatory programs and policy-makers** seeking to certify servers with unpopulated memory channels.
- **Hardware manufacturers and customers** aiming to understand the efficiency implications of different memory configurations.
- **Benchmark designers** interested in refining methodologies for representative server testing.
- **Researchers** focused on the impact of server configuration on experimental outcomes, which is critical for both experiment design and interpretation.

In addition, also at ICPE 2026, we present a tutorial on the challenges associated with standardized energy efficiency benchmarking of GPU-servers, an important focus during the development of the upcoming benchmarks. The group's current focus is on expanding the SPEC Power and Performance Benchmark Methodology [4]. As general-purpose servers adopt new technologies—such as liquid cooling, direct current power supplies, accelerators (GPUs, FPGAs, ASICs), and advanced memory features—we examine key factors influencing efficiency and their relevancy for the next generation of efficiency benchmarks and standards [5].

Maximilian Meissner (University of Würzburg),
Klaus-Dieter Lange (Hewlett Packard Enterprise)
<https://research.spec.org/working-groups/rg-power/>

- [1] <https://research.spec.org/working-groups/rg-power/>.
- [2] <https://www.spec.org/power/>.
- [3] <https://www.spec.org/isg/>.
- [4] https://www.spec.org/power_ssj2008/docs/SPECpower-Methodology.pdf.
- [5] <https://www.spec.org/sert3/>.

GRADUATED LONG-TERM MEMBERS OF THE SPEC RG

Sacheendra Talluri



Cloud computing enables essential societal services, including communication, commerce, entertainment, and healthcare. It aims to be easy to use and low-cost so that as many fields as possible can reap its benefits. However, it is challenging to provide a highly performant and reliable service that is both easy to use and low-cost. We aim to help cloud system designers address this problem by providing a set of tools that enable reasoning about performance and reliability together.

We posit that performance and reliability are part of the same performance-failure continuum. The techniques used to address low performance and unreliability share similarities. For example, when an RPC request is slow to respond, the application retries. The application also retries when a computation task fails in a data-processing workflow. Despite these similarities, the two concepts are studied separately. This separation leads to wasted work and even conflicting design objectives. We believe that high-performance and reliability techniques, when studied together, yield systems that better achieve their objectives. We provide conceptual and technical tools, as well as data, to enable this joint study.

Sacheendra now works as a Software Engineer at ClickHouse [2]. Have a look at his personal website [3] for more information about his achievements and activities.

- [1] S. Talluri, "On the design and characteristics of cloud computing systems along the performance-failure continuum," PhD Thesis, Vrije Universiteit Amsterdam, Feb. 2026. DOI: [10.5463/thesis.1430](https://doi.org/10.5463/thesis.1430).
- [2] <https://clickhouse.com/>.
- [3] <https://sacheendra.github.io/>.

SELECTED ABSTRACTS

DistWalk: a Distributed Workload Emulator

This paper introduces DistWalk, a flexible, distributed, scalable, and open-source toolkit designed to emulate

compute, network, and storage workloads across a networked infrastructure, and measure the resulting end-to-end latency. DistWalk provides fine-grained control over the workload behavior, which consists of a graph-like sequence of operations spanning multiple servers. It supports several communication protocols and traffic patterns, and enables the customization of several factors, such as the duration and parallelism of compute-intensive operations, and the I/O data access and synchronization mode, among others. The proposed toolkit may be used to experiment with a variety of deployment models, from bare-metal to virtualized or containerized environments, e.g., using Cloud/Edge infrastructures, OpenStack, Kubernetes, or other orchestrators, allowing for experimental comparisons of the achievable latency across a wide range of system-level configurations.

R. Andreoli and T. Cucinotta, "Distwalk: A distributed workload emulator," in *2025 IEEE 25th International Symposium on Cluster, Cloud and Internet Computing (CCGrid)*, IEEE, May 2025, pp. 1–10. DOI: [10.1109/ccgrid64434.2025.00042](https://doi.org/10.1109/ccgrid64434.2025.00042). [Online]. Available: <http://dx.doi.org/10.1109/CCGRID64434.2025.00042>.

An Empirical Study on Transient Phases of Microservice Applications

Microservice applications in practice often face runtime adaptations, like autoscaling or updates. These adaptations cause these applications to leave their steady-state performance, leading to so-called transient phases. There is a lack of understanding and quantitative data about the transient phases of microservice applications. Several use cases, including microbenchmarking and autoscaling, can profit from empirical data on transient phases. To this end, we present the most comprehensive empirical study of transient phases to date. We investigated 92 microservices from 10 reference applications, analyzing the influence of programming languages, workloads, and container properties on the duration of transient phases.

I. Rohwer, M. Straesser, Y. Lubas, S. Kounev, and A. Bauer, "An empirical study on transient phases of microservice applications," in *33rd International Symposium on Modeling, Analysis and Simulation of Computer and Telecommunication Systems, MASCOTS 2025, Paris, France, October 21-23, 2025*, IEEE, 2025, pp. 1–8. DOI: [10.1109/MASCOTS67699.2025.11283322](https://doi.org/10.1109/MASCOTS67699.2025.11283322). [Online]. Available: <https://doi.org/10.1109/MASCOTS67699.2025.11283322>.

STEAM: Realistic Modeling and Systematic Exploration of Composable Techniques for Sustainable Datacenters

The need to reduce datacenter carbon-footprint is urgent. While many sustainability techniques have been proposed, they are often evaluated in isolation, using limited setups or analytical models that overlook real-world dynamics and interactions between methods. This makes it challenging for researchers and operators to understand the effectiveness and trade-offs of combining such techniques. We design STEAM, an open-source customiz-

able datacenter simulator, to investigate the individual and combined impact of sustainability techniques on datacenter operational and embodied carbon emissions, and their trade-off with performance. Using STEAM, we systematically explore three representative techniques—horizontal scaling, leveraging batteries, and temporal shifting—with diverse workloads, datacenter configurations, and carbon-intensity traces. Our analysis highlights that datacenter dynamics can influence their effectiveness and that combining strategies can significantly lower emissions, but introduces complex cost emissions-performance trade-offs that STEAM can help navigate. STEAM supports the integration of new models and techniques, making it a foundation framework for holistic, quantitative, and reproducible research in sustainable computing.

D. Niewenhuis, S. Talluri, A. Iosup, and T. de Matteis, "STEAM: Realistic modeling and systematic exploration of composable techniques for sustainable datacenters," in *Proceedings of the 17th ACM/SPEC International Conference on Performance Engineering, ICPE 2026, Florence, Italy, May 4-8, 2026*, ACM, 2026.

MapReplay: Trace-Driven Benchmark Generation for Java HashMap

Hash-based maps, particularly `java.util.HashMap`, are pervasive in Java applications and the JVM, making their performance critical. Evaluating optimizations is challenging because performance depends on factors such as operation patterns, key distributions, and resizing behavior. Microbenchmarks are fast and repeatable but often oversimplify workloads, failing to capture the realistic usage patterns. Application benchmark (e.g., DaCapo, Renaissance) provide realistic usages but are more expensive to run, prone to variability, and dominated by non-HashMap computations, making map-related performance changes difficult to observe. To address this challenge, we propose MapReplay, a benchmarking methodology that combines the realism of application benchmarks with the efficiency of microbenchmarks. MapReplay traces HashMap API usages generating a replay workload that reproduces the same operation sequence while faithfully reconstructing internal map states. This enables realistic and efficient evaluation of alternative implementations under realistic usage patterns. Applying MapReplay to DaCapo-Chopin and Renaissance, the resulting suite—MapReplayBench—reproduces application-level performance trends while reducing experimentation time and revealing insights difficult to obtain from full benchmarks. We publish MapReplay as open-source software on GitHub¹ and Zenodo.²

F. Schiavio, A. Rosà, J. Löff, L. Bulej, P. Tüma, and W. Binder, "MapReplay: Trace-driven benchmark generation for Java HashMap," in *Proceedings of the 17th ACM/SPEC International Conference on Performance Engineering, ICPE 2026, Florence, Italy, May 4-8, 2026*, ACM, 2026.

¹<https://github.com/usi-dag/MapReplay>

²<https://zenodo.org/records/18475007>