# Serverless Computing Revisited: Evolution, State-of-the-Art, and Performance Challenges

Samuel Kounev University of Würzburg Würzburg, Germany samuel.kounev@uni-wuerzburg.de

## ABSTRACT

Market analysts are agreed that serverless computing has strong market potential, with projected compound annual growth rates varying between 21% and 28% through 2028 and a projected market value of \$36.8 billion by that time [6]. Although serverless computing has gained significant attention in industry and academia over the past years, there is still no consensus about its unique distinguishing characteristics and precise understanding of how these characteristics differ from classical cloud computing [5, 6]. For example, there is no wide agreement on whether serverless is solely a set of requirements from the cloud user's perspective or it should also mandate specific implementation choices on the provider side, such as implementing an autoscaling mechanism to achieve elasticity [1, 4, 6]. Similarly, there is no agreement on whether serverless is just the operational part, or it should also include specific programming models, interfaces, or calling protocols.

In this talk, we seek to dispel this confusion by evaluating the essential conceptual characteristics of serverless computing as a paradigm, while putting the various terms around it into perspective. We examine how the term serverless computing, and related terms, are used today. We explain the historical evolution leading to serverless computing, starting with mainframe virtualization in the 1960 through to Grid and cloud computing all the way up to today. We review existing cloud computing service models, including IaaS, PaaS, SaaS, CaaS, FaaS, and BaaS, discussing how they relate to the serverless paradigm. In the second part of talk, we focus on performance challenges in serverless computing both from the user's perspective (finding the optimal size of serverless functions [2, 3]) as well as from the provider's perspective (ensuring predictable and fast container start times coupled with fine-granular and accurate elastic scaling mechanisms [8]).

#### **ACM Reference Format:**

Samuel Kounev. 2023. Serverless Computing Revisited: Evolution, State-ofthe-Art, and Performance Challenges: Keynote Talk Abstract. In *Companion* of the 2023 ACM/SPEC International Conference on Performance Engineering (ICPE'23 Companion), April 15–19, 2023, Coimbra, Portugal. ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3578245.3584856

ICPE'23 Companion, April 15-19, 2023, Coimbra, Portugal

© 2023 Copyright held by the owner/author(s).

ACM ISBN 979-8-4007-0072-9/23/04.

https://doi.org/10.1145/3578245.3584856

### **BIOGRAPHY**

Samuel Kounev is a Professor of Computer Science holding the Chair of Software Engineering at the University of Würzburg. His research is aimed at the engineering of software for building dependable, efficient, and resilient distributed systems, including cloudbased systems, cyber-physical systems, and scientific computing applications. Research topics of the Chair of Software Engineering span the areas of: (1) Software Architecture, focussing on the design, modeling, and simulation of distributed system architectures, (2) Systems Benchmarking, focussing on experimental analysis of performance, scalability, energy efficiency, dependability, and resilience properties, (3) Cyber Security, focussing the design, testing, and evaluation of adaptive security architectures and homomorphic computing techniques, and (4) Predictive Data Analytics, focussing on the software engineering of workflows and tools for time series forecasting, anomaly detection, and critical event prediction. Kounev's research is inspired by the vision of self-aware computing systems, to which he has been one of the major contributors shaping its development [7].

Samuel Kounev studied Mathematics and Computer Science at the University of Sofia from which he holds a MSc degree with distinction (2000). He moved to TU Darmstadt (Germany) in 2001 starting a PhD (Dr.-Ing.) in computer science, which he completed in 2005 with distinction (summa cum laude) and an award for outstanding scientific achievements.He was a research fellow at the University of Cambridge (2006-2008) and Visiting Professor at UPC Barcelona (2006, 2007). In 2009, Kounev received the DFG Emmy-Noether-Career-Award (1MM EUR) for excellent young scientists, establishing his research group "Descartes" at Karlsruhe Institute of Technology. Since 2014, Samuel Kounev is a Full Professor holding the Chair of Software Engineering at the University of Würzburg, where he has served in various roles including Dean and Vice Dean of the Faculty of Mathematics and Computer Science, Managing Director of the Institute of Computer Science.

Samuel Kounev is Founder and Elected Chair of the SPEC Research Group within the Standard Performance Evaluation Corporation (SPEC). This group has over 50 member organizations from around the world and serves as a platform for collaborative research efforts in the area of quantitative system evaluation and analysis, fostering the interaction between academia and industry. His research has lead to over 300 publications and multiple scientific and industrial awards including 7 Best Paper Awards, SPEC Presidential Award for "Excellence in Research", Google Research Award, ABB Research Award, and a VMware Academic Research Award.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

ICPE'23 Companion, April 15-19, 2023, Coimbra, Portugal

## REFERENCES

- André Bauer, Nikolas Herbst, Simon Spinner, Ahmed Ali-Eldin, and Samuel Kounev. 2019. Chameleon: A Hybrid, Proactive Auto-Scaling Mechanism on a Level-Playing Field. *IEEE Transactions on Parallel and Distributed Systems* 30, 4 (April 2019), 800–813.
- [2] Simon Eismann, Long Bui, Johannes Grohmann, Cristina Abad, Nikolas Herbst, and Samuel Kounev. 2021. Sizeless: Predicting the Optimal Size of Serverless Functions. In Proceedings of the 22nd International MIDDLEWARE Conference. 248–259. Best Student Paper Award, ACM Artifacts Evaluated – Functional.
- [3] Simon Eismann, Johannes Grohmann, Erwin van Eyk, Nikolas Herbst, and Samuel Kounev. 2020. Predicting the Costs of Serverless Workflows. In Proceedings of the 2020 ACM/SPEC International Conference on Performance Engineering (ICPE) (ICPE'20). Association for Computing Machinery (ACM), New York, NY, USA, 265–276. https://doi.org/10.1145/3358960.3379133
- [4] Nikolas Roman Herbst, Samuel Kounev, and Ralf Reussner. 2013. Elasticity in Cloud Computing: What it is, and What it is Not. In Proceedings of the 10th International Conference on Autonomic Computing (ICAC 2013). USENIX.
- [5] Samuel Kounev, Cristina Abad, Ian T. Foster, Nikolas Herbst, Alexandru Iosup, Samer Al-Kiswany, Ahmed Ali-Eldin Hassan, Bartosz Balis, Andre Bauer, Andre B. Bondi, Kyle Chard, Ryan L. Chard, Robert Chatley, Andrew A. Chien,

A. Jesse Jiryu Davis, Jesse Donkervliet, Simon Eismann, Erik Elmroth, Nicola Ferrier, Hans-Arno Jacobsen, Pooyan Jamshidi, Georgios Kousiouris, Philipp Leitner, Pedro Garcia Lopez, Martina Maggio, Maciej Malawski, Bernard Metzler, Vinod Muthusamy, Alessandro V. Papadopoulos, Panos Patros, Guillaume Pierre, Omer F. Rana, Robert P. Ricci, Joel Scheuner, Mina Sedaghat, Mohammad Shahrad, Prashant Shenoy, Josef Spillner, Davide Taibi, Douglas Thain, Animesh Trivedi, Alexandru Uta, Vincent van Beek, Erwin van Eyk, Andre van Hoorn, Soam Vasani, Florian Wamser, Guido Wirtz, and Vladimir Yussupov. 2021. Toward a Definition for Serverless Computing. In *Serverless Computing (Dagstuhl Seminar 21201)*, Cristina Abad, Ian T. Foster, Nikolas Herbst, and Alexandru Iosup (Eds.). Vol. 11. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl, Germany, 34–93. Issue 4. https://doi.org/10.4230/DagRep.11.4.34

- [6] Samuel Kounev, Nikolas Herbst, Cristina L. Abad, Alexandru Iosup, Ian Foster, Prashant Shenoy, Omer Rana, and Andrew A. Chien. 2023. Serverless Computing: What It Is, and What It Is Not? *Communications of the ACM (CACM)* (2023). Under publication.
- [7] Samuel Kounev, Jeffrey O. Kephart, Aleksandar Milenkoski, and Xiaoyun Zhu. 2017. Self-Aware Computing Systems. Springer Verlag, Berlin Heidelberg, Germany.
- [8] Martin Straesser, André Bauer, Robert Leppich, Nikolas Herbst, Kyle Chard, Ian Foster, and Samuel Kounev. 2023. An Empirical Study of Container Image Configurations and Their Impact on Start Times. In 2023 23rd IEEE International Symposium on Cluster, Cloud and Internet Computing (CCGrid).