Gregor Engels, Regina Hebig, Matthias Tichy (Hrsg.): SE 2023,
Lecture Notes in Informatics (LNI), Gesellschaft für Informatik, Bonn 2023 59

Benchmarking Scalability of Cloud-Native Applications

Sören Henning[1] Wilhelm Hasselbring[2]

Abstract: This contribution has been published in the journal *Empirical Software Engineering* (Springer Nature) in 2022 [HH22], https://doi.org/10.1007/s10664-022-10162-1.

Keywords: Scalability; Benchmarking; Performance engineering; Cloud-Native

1 Introduction

Software architectures [Ha18] significantly influence the quality characteristics of the resulting software systems. Scalability is such a quality characteristic that is in particular relevant for data stream processing systems [HH20; HH21b; HHM19]. In the context of analyzing IoT sensor data [HWD21], we study scalable architectures for power consumption monitoring [He21]. Cloud-native applications constitute a recent trend for designing large-scale software systems, with a focus on *scalability* (https://www.cncf.io/).

In this paper, we present the Theodolite benchmarking method, allowing researchers and practitioners to conduct empirical scalability evaluations of cloud-native applications, frameworks, and deployment options. Although scalability is often mentioned as a key driver for adopting cloud-native architectures and microservices [KH19], we found that research is lacking a commonly accepted method to empirically assess and compare the *scalability* of cloud-native applications. In empirical software engineering, benchmarks are used as a measuring instrument for comparing different technologies or configurations [Ha21]. Thus, we designed the Theodolite method for benchmarking scalability of cloud-native applications [HH22]. Our benchmarking method consists of scalability metrics [HH21a], measurement methods, and an architecture for a scalability benchmarking framework, particularly suited for cloud-native applications. To balance usability and reproducibility, our benchmarking method provides configuration options, controlling the trade-off between overall execution time and statistical grounding. We performed an extensive experimental evaluation of our method's configuration options for data stream processing applications. We find that, independent of the cloud platform, it only takes a few repetitions (≤ 5) and short execution times (≤ 5 minutes) to assess whether SLOs are achieved. Combined with our findings from evaluating different search strategies, we conclude that our method allows to benchmark scalability in reasonable time.

[1] Kiel University, Software Engineering Group, 24098 Kiel, Germany soeren.henning@email.uni-kiel.de
[2] Kiel University, Software Engineering Group, 24098 Kiel, Germany hasselbring@email.uni-kiel.de

Data Availability A replication package is available at Zenodo (`https://doi.org/10.5281/zenodo.5596982`). The source code is available at GitHub (`https://github.com/cau-se/theodolite`) and the software documentation at `https://www.theodolite.rocks/`.

Literatur

[Ha18] Hasselbring, W.: Software Architecture: Past, Present, Future. In: The Essence of Software Engineering. Springer, S. 169–184, 2018.

[Ha21] Hasselbring, W.: Benchmarking as Empirical Standard in Software Engineering Research. In: International Conference on Evaluation and Assessment in Software Engineering (EASE 2021). ACM, S. 365–372, Juni 2021.

[He21] Henning, S.; Hasselbring, W.; Burmester, H.; Möbius, A.; Wojcieszak, M.: Goals and measures for analyzing power consumption data in manufacturing enterprises. Journal of Data, Information and Management 3/1, S. 65–82, 2021.

[HH20] Henning, S.; Hasselbring, W.: Scalable and Reliable Multi-Dimensional Sensor Data Aggregation in Data-Streaming Architectures. Data-Enabled Discovery and Applications 4/1, S. 1–12, 2020.

[HH21a] Henning, S.; Hasselbring, W.: How to Measure Scalability of Distributed Stream Processing Engines? In: Companion of the ACM/SPEC International Conference on Performance Engineering. ACM, S. 85–88, Apr. 2021.

[HH21b] Henning, S.; Hasselbring, W.: Theodolite: Scalability Benchmarking of Distributed Stream Processing Engines in Microservice Architectures. Big Data Research 25/100209, S. 1–17, Juli 2021.

[HH22] Henning, S.; Hasselbring, W.: A Configurable Method for Benchmarking Scalability of Cloud-Native Applications. Empirical Software Engineering 27/143, S. 1–42, 2022, URL: `https://doi.org/10.1007/s10664-022-10162-1`.

[HHM19] Henning, S.; Hasselbring, W.; Möbius, A.: A Scalable Architecture for Power Consumption Monitoring in Industrial Production Environments. In: 2019 IEEE International Conference on Fog Computing (ICFC). IEEE, Prague, Czech Republic, S. 124–133, Juni 2019.

[HWD21] Hasselbring, W.; Wojcieszak, M.; Dustdar, S.: Control Flow Versus Data Flow in Distributed Systems Integration: Revival of Flow-Based Programming for the Industrial Internet of Things. IEEE Internet Computing 25/4, S. 5–12, 2021.

[KH19] Knoche, H.; Hasselbring, W.: Drivers and Barriers for Microservice Adoption – A Survey among Professionals in Germany. Enterprise Modelling and Information Systems Architectures (EMISAJ) – International Journal of Conceptual Modeling 14/1, S. 1–35, 2019.