Exploring the Feasibility of Performance Regression Testing for Serverless Applications

Simon Eismann
simon.eismann@uni-wuerzburg.de
Universität Würzburg, Germany

Diego Costa
diego.costa@concordia.ca
Concordia University, Canada

Lizhi Liao, llizhi@encs.concordia.ca
Concordia University Canada

Weiyi Shang,
shang@encs.concordia.ca
Concordia University, Canada

Cor-Paul Bezemer
cpbezemer@gmail.com
University of Alberta, Canada

André van Hoorn
van.hoorn@informatik.uni-stuttgart.de
University of Stuttgart, Germany

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

Serverless applications combine Function-as-a-Service offerings (e.g., AWS Lambda, Google Cloud Functions or Azure Functions) with Backend-as-a-Service offerings (e.g., managed storage, databases, pub/sub, queueing, streaming or workflows) to create applications that require no resource management [1, 2]. The cloud provider opaque handles resource management tasks, such as deployment, resource allocation or auto scaling and bills the user on a pay-per-use basis [3, 5].

While the cloud provider takes care of the resource management, managing the performance of the serverless application remains a developer concern [5, 6]. Executing performance tests as part of a CI/CD pipeline to monitor the impact of code changes on system performance (also known as performance regression testing), is a common and powerful approach to manage system performance [7, 8]. One of the key requirements for reliable performance tests is ensuring that an identical resource environment is used for all tests [9].

However, with serverless applications, developers have no control over the resource environment. Worse yet, cloud providers expose no information to developers about the resource environment [10]. Therefore, information such as the number of provisioned workers, worker utilization, worker version, virtualization stack, or underlying hardware is unavailable to developers. This begs the question, "Is it possible to conduct accurate performance regression testing for serverless applications?".

Existing work focuses on performance regression testing for traditional systems, the performance variability of IaaS offerings, and the performance analysis of managed cloud services. However, there are very few studies on the performance of realistic serverless applications [11] and to the best of our knowledge, no studies on the performance regression testing of serverless applications.

In this study, we explore the feasibility of performance regression testing for serverless applications. We conduct 180 performance experiment runs spanning over 45 hours of total measurement time and more than 25 million requests using the serverless airline application, a representative, production-grade serverless application. Additionally, we conduct a longitudinal study with continuous daily measurements spanning 50 days at the time of submission.

References

[4] Van Eyk, Erwin, Johannes Grohmann, Simon Eis-


