[Extended Abstract] 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, l_lizhi@encs.concordia.ca
Concordia University Canada

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

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

[1] Castro, P., Ishakian, V., Muthusamy, V., & Slominski, A. (2019). The rise of serverless computing. Communications of the ACM, 62(12), 44-54.
[2] Eismann, S., Scheuner, J., van Eyk, E., Schwinger, M., Grohmann, J., Herbst, N., ... & Iosup, A. (2020). A review of serverless use cases and their characteristics. arXiv preprint arXiv:2008.11110.
[3] Baldini, Ioana, Paul Castro, Kerry Chang, Perry Cheng, Stephen Fink, Vatche Ishakian, Nick Mitchell et al. "Serverless computing: Current trends and open problems." In Research Advances in Cloud Computing, pp. 1-20. Springer, Singapore, 2017.
[4] Van Eyk, Erwin, Johannes Grohmann, Simon Eis-

mann, André Bauer, Laurens Versluis, Lucian Toader, Norbert Schmitt, Nikolas Herbst, Cristina L. Abad, and Alexandru Iosup. "The SPEC-RG Reference Architecture for FaaS: From Microservices and Containers to Serverless Platforms." IEEE Internet Computing 23, no. 6 (2019): 7-18.
[5] Leitner, P., Wittern, E., Spillner, J., & Hummer, W. (2019). A mixed-method empirical study of Function-as-a-Service software development in industrial practice. Journal of Systems and Software, 149, 340-359.
[6] Van Eyk, E., Iosup, A., Abad, C. L., Grohmann, J.,& Eismann, S. (2018, April). A SPEC RG cloud group's vision on the performance challenges of FaaS cloud architectures. In Companion of the 2018 ACM/SPEC International Conference on Performance Engineering (pp. 21-24).
[7] Daly, D., Brown, W., Ingo, H., O'Leary, J., & Bradford, D. (2020, April). The Use of Change Point Detection to Identify Software Performance Regressions in a Continuous Integration System. In Proceedings of the ACM/SPEC International Conference on Performance Engineering (pp. 67-75).
[8] Leitner, Philipp, and Cor-Paul Bezemer. "An exploratory study of the state of practice of performance testing in java-based open source projects." Proceedings of the 8th ACM/SPEC on International Conference on Performance Engineering. 2017.
[9] Eismann, S., Bezemer, C. P., Shang, W., Okanović, D., & van Hoorn, A. (2020, April). Microservices: A Performance Tester's Dream or Nightmare?. In Proceedings of the ACM/SPEC International Conference on Performance Engineering (pp. 138-149).
[10] Wang, L., Li, M., Zhang, Y., Ristenpart, T., & Swift, M. (2018). Peeking behind the curtains of serverless platforms. In 2018 USENIX Annual Technical Conference (USENIXATC 18) (pp. 133-146).
[11] Scheuner, J. & Leitner, P. (2020). Function-as-a-Service performance evaluation: A multivocal literature review. Journal of Systems and Software, 110708.