Contribution ID: **262**   Type: **not specified**

DSLab FaaS: Fast and Accurate Simulation of FaaS Clouds

Friday 7 July 2023 12:15 (15 minutes)

Resource management in cloud computing, and especially in FaaS (Function-as-a-Service), is a very active area of research, given the rise in the popularity of serverless computing. In serverless computing users do not explicitly manage VMs or containers, instead, they upload their functions to the cloud, and the functions are executed when triggered by events, such as HTTP requests. All work regarding resource management, server provisioning, and execution is done by the cloud platform. Developing algorithms for FaaS scheduling and resource management requires a thorough evaluation, as poor decisions may result in huge financial losses. However, running large-scale experiments on a real platform is very time-consuming and expensive. Therefore researchers working in this area require a performant, accurate, and customizable simulation framework. Although there exist various tools for simulating older and more established cloud computing models (e. g. IaaS), for FaaS the choice is much more narrow, and the existing simulation tools are far from perfect.

In this report, we present a modular FaaS simulation framework, called DSLab FaaS. This framework allows for a detailed simulation of FaaS cloud platforms with customizable implementations of the main components: users can implement their own logic for the scheduler, invoker, CPU sharing, container lifetime management, and container deployment. The framework supports main workload trace formats and the users can integrate their own formats into the framework in a convenient way. The framework allows users to collect a variety of metrics related to resource management efficiency, both globally and for each app or function separately. The applicability of DSLab FaaS is shown by reproducing some of the research works related to resource management in FaaS platforms in much less time: several seconds of simulation instead of several hours of experiments on a real platform. The framework is implemented in Rust with a strong focus on performance, its runtime scales linearly to the data size and it is able to run multi-threaded experiments. The performance and scalability of the simulator are verified by computational benchmarks against other FaaS simulators on real workload trace from Microsoft Azure Functions service.

Summary

Author: SEMENOV, Yury (MIPT)

Co-author: SUKHOROSLOV, Oleg (IITP RAS)

Presenter: SEMENOV, Yury (MIPT)

Session Classification: Cloud Technologies

Track Classification: Cloud Technologies