

Sheshadri K R

PhD candidate, CSL, IISC, Bangalore

Advised by: Dr. J Lakshmi, CSL, SERC, IISc

Research topic

QoS aware FaaS platform, FaaS, serverless computing, cloud computing, IoT

Research Summary

Cloud computing offers virtualized computing resources as services to the users. It is used extensively by many businesses for building and hosting their products.

Popular offerings include IaaS and CaaS. However these services often require personnel with specific skills to manage. They have slow start-up times and cannot quickly scale to changing workloads. Though these computing services offer high degree of control, the complete responsibility of managing them is left to the users which can be hard to undertake.

In a shift towards offering ease-of-use, Function as a Service(FaaS) is positioned to be the service offering that reduces the burden on the user and shifts the burden of operational management of servers to the platform. FaaS is a form of serverless computing, and it is currently one of the most popular cloud computing service offerings. FaaS adopts serverless computing features to provide agility in deployment and sub-second billing for payment on use (and also provides no idle resource billing).

The programming model of FaaS expects an application to follow certain characteristics by virtue of its design. In this paradigm, developers design their applications as a workflow of functions, which are stateless (function state is not maintained between executions) and event-driven (Function executions are triggered by events such as document uploads, database updates, HTTP events, etc.).

Though the adoption of the FaaS service model has happened at a high pace, it comes with many shortcomings. The existing FaaS platforms (AWS lambda, Google Cloud Functions, Apache Openwhisk, and Openlambda) are insensitive to the Quality of Service (QoS) requirements of an application. QoS is used to specify non-functional requirements; it can be aspects such as time of completion, resource specification, number of concurrent resource instances of a particular function, etc. Existing platforms lack a mechanism to specify and capture the application's QoS requirements flexibly. The QoS requirements are derived from applications of various domains (Object Analytics, Web applications, NFV) which have adopted FaaS or show characteristics that are a good fit for FaaS. It is shown that the insufficiency in specifying application specificities in the FaaS platform is a problem that impacts applications from many domains.

A QoS aware FaaS platform tries to address this problem by proposing a QoS framework as the additional component to the existing FaaS framework. It contains two major components QoS specification framework and QoS aware resource manager. The QoS specification framework addresses the problem of flexibly expressing the application's requirements to the FaaS platform; it increases the application's expressibility. The QoS aware resource manager is the central component in the FaaS platform. It oversees many decisions ranging from resource instance creation to instance placement decisions. Importantly, the QoS aware resource manager binds each function execution request with the QoS specification of the function and provisions appropriate resources for its deployment.

Publication

* QoS aware FaaS platform, Sheshadri K R and J Lakshmi - STEERS-2021, held in conjunction with CCGrid 2021 (Apr-2021)

Awards

* Best paper award, QoS aware FaaS platform, Sheshadri K R and J Lakshmi - STEERS-2021