

Serverless Services for Scientific Cloud Computing

Thursday 13 October 2022 16:15 (30 minutes)

Serverless is an application development paradigm that uses Cloud-based managed services for automated resource allocation. Functions as a Service (FaaS) is a computing model that executes functions in response to events in the infrastructure to perform highly elastic computations. These are being adopted for scientific computations, to hide the operational procedures of resource allocations in the Cloud.

This contribution focuses on the development of innovative serverless-based open-source software developments carried out by the GRyCAP/I3M research group at the Universitat Politècnica de València (Spain).

First, SCAR (<https://github.com/grycap/scar>), which supports Docker containers in AWS Lambda years before this functionality was natively introduced in AWS Lambda. SCAR is integrated with API Gateway and AWS Batch to support event-driven execution of Docker-based computing applications that can execute in a hybrid approach both using a FaaS approach, in AWS Lambda, and via auto-scaled computing clusters, even with GPU support, in AWS Batch.

Second, OSCAR (<https://github.com/grycap/oscar>) provides a platform for event-driven data processing on auto-scaled Kubernetes clusters, which can be dynamically deployed across multiple Clouds using the Infrastructure Manager (<https://www.grycap.upv.es/im>). It supports both asynchronous executions via Kubernetes jobs and synchronous ones via Knative to fit different use case requirements, such as performing the inference of pre-trained AI/ML models. OSCAR can be executed in low-power devices such as Raspberry Pis, useful for edge computing, and shares the same Functions Definition Language (FDL) as SCAR and, therefore, it allows to create event-driven workflows that can span along the computing continuum.

Third, MARLA (<https://github.com/grycap/marla>) provides a completely serverless environment for MapReduce executions on top of AWS Lambda environments. MARLA uses the S3 service to receive data files to be processed via an event-driven workflow. Therefore, all the data processing is done automatically from data loading to obtaining the final result. In addition, the user can define their specific map and reduce functions thanks to the builtin python environment included in the AWS Lambda service, allowing any type of analysis on the data.

Fourth, TaSaaS (<https://github.com/grycap/TaSaaS>) is a serverless service to distribute and balance work executions among independent infrastructures. TaSaaS is focused on scientific executions with huge computational cost, which usually must be split into hundreds or thousands of smaller partitions to be computed in parallel on different nodes and computational infrastructures. With this aim, TaSaaS divides automatically the works in partitions to be executed on the available resources. It also balances the load of all partitions belonging to the same job and monitors their status to detect node failures. TaSaaS is deployed on AWS but its architecture can be extended to other cloud providers.

In summary, a range of serverless-based innovative software developments have been created which aim to simplify the adoption of this novel computational paradigm for scientific computing.

Grant PID2020-113126RB-I00 funded by MCIN/AEI/10.13039/501100011033. Project PDC2021-120844-I00 funded by MCIN/AEI/10.13039/501100011033 funded by the European Union NextGenerationEU/PRTR

Primary authors: RISCO, Sebastián (Universitat Politècnica de València); Dr GIMÉNEZ, Vicent (UPVLC); Dr CABALLER, Miguel (UPVLC); MOLTÓ, Germán (Universitat Politècnica de València)

Presenter: MOLTÓ, Germán (Universitat Politècnica de València)

Session Classification: IBERGRID Contributions

Track Classification: Development of innovative software services