NEWS LETTER

3

OCTOBER | 2023

CLOUD STARS

First experiences from the University of Murcia in Yorktown

by Aurora González Vidal (University of Murcia)

The University of Murcia (UM) has started its secondemnts within the Cloudstards project since 1 postdoctoral researcher has finished her 3 months externship in IBM Thomas J. Watson Research Center, Yorktown, NY.

The researched topic was serving Artificial Intelligence Models for inference, which is a critical aspect of deploying AI systems into real-world applications. Inference is the process where a trained AI model applies its learned knowledge to make predictions or classifications based on new data inputs. To effectively serve AI models for inference, a robust and scalable infrastructure is required. This infrastructure often involves deploying models on high-performance servers or cloud-based platforms, ensuring low-latency responses, and handling concurrent requests efficiently while deadlines guarantees are achieved. More in detail, the research consisted of monitoring the status of the system and measuring time under different cyrcumstances in order to propose a mathematical model that optimises the decisions that are taken when performing image classification.

During this period, the secondments from researchers from the Technical University of Munich (TUM) and Technical University of Vienna (TU Wien) also took place in the same facility, fostering an environment conducive to open dialogue, collaboration, and the exchange of experiences. Given the proximity of their expertise, there were numerous synergies among the teams. In fact, UM and TUM worked together on crafting and submitting a paper to the 9th International Workshop on Serverless Computing (WOSC9), held in conjunction with Middleware'23. This paper presented preliminary findings on the topic of serving Artificial Intelligence models for inference and was co-authored by two mentors from IBM Yorktown as well.

Overall, it proved to be a learning and productive period that has opened new research lines for the UM team.

cloudstars.eu | twitter.com/Cloudstars_2023 | github.com/cloudstars-eu

CLOUDSTARS project has received funding from the European Union's Horizon research and innovation programme under grant agreement No 101086248