Contribution ID: **15** Type: **Demonstrations & Tutorials**

Parallel Inference in the Edge-to-Cloud for Health Monitoring

Tuesday, 1 October 2024 18:00 (30 minutes)

In the context of Artificial Intelligence (AI), the evolution of computing paradigms from centralized data centers to the edge of the network heralds a transformative shift in how AI applications are developed, deployed, and operated. Specifically, the edge computing paradigm is characterized by processing data directly in the devices where it is collected, such as smartphones, wearables, and IoT. Edge computing significantly reduces latency, conserves bandwidth, and enhances data privacy and security, thereby catalyzing the realization of real-time, responsive AI solutions.
The AI-SPRINT (Artificial Intelligence in Secure PRIvacy-preserving computing coNTinuum) H2020 project has developed a comprehensive design and runtime framework dedicated to accelerating the development and deployment of AI applications across the computing continuum. This continuum spans from cloud data centers to edge devices and sensors, integrating AI components to operate seamlessly and efficiently. The project's core objective is to offer a suite of tools that enable an optimal balance among application performance, energy efficiency, and AI model accuracy, all while upholding rigorous security and privacy standards.

The Personalized Healthcare use case focuses on harnessing the power of AI and wearable technologies for health monitoring. By integrating quantitative data on heart functions from wearable device sensors with qualitative lifestyle information, the use case aims to develop a personalized stroke risk assessments model. This initiative is particularly critical given the prevalence of stroke among the aging population, marking it as a significant cause of death and disability globally. Leveraging the AI-SPRINT framework, this application enables efficient resource distribution and computation across the edge-to-cloud continuum, facilitating real-time, non-invasive and secure monitoring and risk assessment.
The use case has been implemented using the PyCOMPSs programming framework from BSC to develop a Machine Learning model for detecting atrial fibrillation (AF) in electrocardiogram data (ECG) implemented using the parallel dislib library built on top of PyCOMPSs.
In this demo we show how the prediction risk is calculated, using ECGs extracted from a wearable device, on a Cloud server deployed in the EGI Cloud, with resources provided on-demand by the OSCAR framework from UPV.
OSCAR is an open-source platform to support the serverless computing model for event-driven data-processing applications. It can be automatically deployed on multiple Cloud backends, thanks to the Infrastructure Manager, to create highly parallel event-driven data-processing serverless applications that execute on customized runtime environments provided by Docker containers than run on an elastic Kubernetes cluster.
In the demo, the OSCAR cluster is deployed and configured with a MinIO storage. When new data (ECG files) is sent through an HTTP request by uploading the ECG file to MinIO, OSCAR triggers a PyCOMPSs/dislib container creation to serve the execution of the inference computation in a Function as a Service mode. This is implemented through a script that starts a PyCOMPSs instance and uploads the result data back in the storage. The number of resources used in the execution can be configured dynamically through an integration with OSCAR and the COMPSs runtime through environment variables.

Topic

Needs and solutions in scientific computing: Artificial Intelligence

Primary authors: ALARCON MARIN, Caterina (Universitat Politècnica de València); LEZZI, Daniele (Barcelona Supercomputing Center); Dr CIRILLO, Davide (Barcelona Supercomputing Center); Dr LORDAN-GOMIS, Francesc (Barcelona Supercomputing Center); MOLTO, German (Universitat Politècnica de València); CABALLER, Miguel (Universitat Politècnica de València)

Presenter: LEZZI, Daniele (Barcelona Supercomputing Center)

Session Classification: Demonstrations & Posters