

James Hutton Institute / Biomathematics and Statistics Scotland

JHI

Staff

Paolo Ribeca (PhD)
Lucio Marcello (PhD)
Tyler Chafin (PhD)
Iain Milne (Engineer)
Ayman Bourramouss (Student at URV)
Francisco Damián Maleno (Student at URV)
Xavier Roca Canals (Student at URV)
Gil Arasa Vergé (Student at URV)

Summary

JHI developed the genomics use case. We started the first phase of the project as a collaboration with Imperial, whereby we studied possible ways of integrating bioinformatics workflows with Faasm. Due to practical issues with recruitment exacerbated by COVID-19, during the last phase of the project we switched to a collaboration with URV based on Lithops and benefitting from the contribution of URV students.

Open source repository

Bioinformatic components and genomic workflow:
https://gitlab1.bioss.ac.uk/lmarcello/serverless_genomics

Collaborations (Total: 2)

JHI collaborated with Imperial and URV on the implementation of the genomics use case.

In detail:

1. At the beginning of the project, JHI and Imperial collaborated on the porting to Faasm/WebAssembly of complex, existing bioinformatics code written in C (the GEM mapper) previously developed by the JHI PI's lab.
2. During the last months of the project, JHI and URV collaborated closely on the implementation of Serverless Genomics pipelines with Lithops. We were able to pursue a multi-pronged program of work based on Lithops, that explored:

1. The creation of needed components to port to the cloud existing bioinformatics workflows
2. A full demonstration of the technique, consisting of a variant-calling pipeline that uses algorithms previously developed by the JHI PI's group being ported to Lithops and serverless. Such a pipeline provides results that are competitive with existing commercial alternatives, both in terms of execution wall-clock time and cost
3. How to adapt to serverless in the most transparent way an existing bioinformatics code base written in OCaml. That was accomplished by designing a wrapper framework based on Lithops that minimises the need for code rewrites.

The results will be published after CloudButton finishes.