Workflow vs. Dataflow: Concepts, challenges, and simulation
for High Performance Computing

Internship/Ph.D. proposal

This internship requires solid skills in C/C++ programming, Linux, and concurrent programming. The main objective is to combine the respective strengths of two renowned simulation toolkits: PREESM [1] and SimGrid [2]. The former is dataflow-oriented [3] and able to accurately simulate the performance of a single heterogeneous node that combines CPUs, GPUs, and FPGAs. The latter is workflow-oriented [4] and can accurately simulate the communications between heterogeneous nodes.

Workflows and dataflows are conceptually close. Both model applications that exhibit task and data parallelism, but they also show some fundamental differences. A first milestone will be to identify the common concepts and main differences between these two application models. Then, the main part of the internship will be to adapt, extend, and/or modify the workflow-oriented programming interface of the SimGrid toolkit to enable the simulation of dataflow-oriented applications.

Depending on the capacity of the candidate to achieve these two milestones within the duration of the internship, the development of a full-fledged simulator and the performance evaluation of scheduling algorithms will be considered.

This internship will indeed be done in the context of the funded ANR DARK-ERA (2021-2024) project and can lead to a Ph.D program. This project is connected to the international Square Kilometer Array (SKA) radio-telescope that will generate data at an unprecedented rate (several Terabits per second). A chain of complex algorithms will be executed on this data in real time, without any available storage capacity, and with an energy budget of less than a MWatt for 250 Petaflops.

Such constraints require to design an innovative and highly heterogeneous architecture. To help at this design, the Dark-Era project aims at developing a rapid prototyping tools of dataflow applications to quickly assess memory usage, latency, throughput, or energy consumption and explore a large design space.

To apply to this internship, please send your resume and cover letter to martin.quinson@ens-rennes.fr and fred-eric.suter@cc.in2p3.fr

References


