Contribution ID: **4** Type: **Talk (15min + 5min)**

Taskfarm: A Client/Server Framework for Supporting Massive Embarrassingly Parallel Workloads

Thursday, March 7, 2024 11:20 AM (20 minutes)

Task farms can be used to solve embarrassingly parallel workloads where a number of independent tasks need to be performed. This presentation introduces *taskfarm*, a python client/server framework that was designed to manage a satellite data processing workflow with hundreds of thousands of tasks with variable compute costs. The server uses flask to hand out tasks via a REST API and a database to track the progress of tasks. The client is also implemented in python. The presentation will focus on the software design process, the pitfalls and dead ends encountered when dealing with big data and how they were resolved.

Slot length

Primary author: HAGDORN, Magnus Karl Moritz (Charité Universitätsmedizin Berlin)

Presenter: HAGDORN, Magnus Karl Moritz (Charité Universitätsmedizin Berlin)

Session Classification: Workflowmanagement for Parallel Computing

Track Classification: Research Software: Research Software for Science