

Dataflow Algorithm aRchitecture co-design of SKA pipeline for Exascale Radio Astronomy

Daniel Charlet, Karol Desnos, Mickaël Dardaillon, André Ferrari, Chiara Ferrari, Nicolas Gac, Jean Francois Nezan, François Orieux, Simon Prunet, Martin Quinson, et al.

▶ To cite this version:

Daniel Charlet, Karol Desnos, Mickaël Dardaillon, André Ferrari, Chiara Ferrari, et al.. Dataflow Algorithm aRchitecture co-design of SKA pipeline for Exascale Radio Astronomy. ISC High Performance 2021, Jun 2021, Virtual conference, France. pp.1. hal-03233205v1

HAL Id: hal-03233205 https://hal.science/hal-03233205v1

Submitted on 24 May 2021 (v1), last revised 27 Aug 2021 (v2)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers. L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.





Dataflow Algorithm aRchitecture co-design of SKA pipeline for Exascale Radio Astronomy

Daniel Charlet^{**5} (IJCLab), Karol Desnos¹, Mickael Dardaillon³, André Ferrari⁴, Chiara Ferrari⁴, Nicolas Gac³, Jean-François Nezan¹, François Orieux³, Simon Prunet⁴, Martin Quinson², Frédéric Suter^{**2}(IN2P3 Computing Center), Cyril Tasse^{**5} (GEPI), Cédric Viou⁵

> ¹IETR (INSA), ²IRISA (ENS), ³L2S (CS), ⁴Lagrange (UCA), ⁵Nançay (Obs Paris) contact: nicolas.gac@l2s.centralesupelec.fr

> > ISC Project Poster Session, June 30 2021

SKA computing, an HPC challenge

The exascale radio telescope Square Kilometre Array (SKA) [?] will require supercomputers with high technical demands. The Science Data Processor (SDP) pipeline in charge of producing the multidimensional images of the sky will have to execute in **realtime** a **complex** algorithm chain with data coming from telescopes at an incredible rate of several Tb/s and limited storage possibilities. The SDP will also have to be as green as possible with an energy budget of only 1 MWatt for 250 Petaflops.

SimSDP, an exascale simulation tool

SimSDP purpose is to provide early analyses in terms of mem-



The **SDP** supercomputer will be based on a standard HPC system combined with FPGA or application-specific architectures like GPU or the manycore Kalray Massively Parallel Processor Array (MPPA). One crucial challenge is to assess the performance both in time and energy of new complex scientific dataflow algorithms on not-yetexisting complex computing infrastructures. It will be hardly possible without efficient **co-design methods** and **rapid prototyping tools**.

ory usage, latency, throughput, and energy consumption through an original mixed approach based on execution and simulation. Following an Algorithm Architecture Matching (AAM) approach, SimSDP will rely on a dataflow model of the algorithm and a model of the target architecture. SimSDP will be based on two existing tools: PREESM and SimGrid. PREESM accurately evaluates **heterogeneous single node** performance; SimGrid accurately simulates **inter-node communications**. Their association will allow for reliable simulations of **large scale heterogeneous HPC systems**.



Dark-Era tasks

T1 – Radio astronomy requirements

Dark-Era, a 4-year project starting in 2021

The **consortium** gathers complementary skills in computer science, signal processing, and astronomy with twelve permanent members from the **SimGrid** [?] development Team at IRISA, the **PREESM** [?] development team at IETR, the **inverse problem** team at L2S, and two radio astronomy teams at Observatories of Paris and Côte d'Azur. Dark-Era has the support of SKA-France and works in collaboration with Atos-Bull.

The **PhD and PostDoc positions** within Dark-Era will be published on https://dark-era.pages.centralesupelec.fr :

- One post-doc at LS2 on HLS FPGA prototype (autumn 2021)
- Two PhDs at IETR and IRISA respectively on PREESM and SimGrid extensions (autumn 2022)



T1 will provide reference datasets and an SDP computing state of the art in terms of programming models, architectures and algorithms.

T2 is the core development task of the SimSDP rapid prototyp-New clustering and architecture model extensions in ing tool. PREESM, and dataflow and complex HPC node extensions in SimGrid will be developed to improve the performance of both tools, enable seamless communication between them and, finally improve the quality of the large scale simulations.

one post-doc at Lagrange on algorithm/architecture exploration for radioastronomy (spring 2024)

Dark-Era goals

- Building **SimSDP**, a rapid prototyping tool providing exascale simulations from dataflow algorithm descriptions.
- 2 Exploring low power accelerators like FPGA or Kalray MPPA as alternatives to mainstream GPU architecture.
- 3 Being source of proposals for SKA computing and promoting French contributions such as **ddfacet** [?].

In **T3**, SDP prototypes on MPPA and FPGA designed through High Level Synthesis (HLS) tools and will be developed and profiled on small scale datasets set up at the NenuFAR radio telescope. It will allow evaluating the potential of these low power accelerators. The SDP profiling feedbacks on GPU, MPPA and FPGA obtained will provide SimSDP with annotations on the dataflow graph.

In **T4**, algorithm and architecture spaces will be explored in the SKA context through the large scale simulations offered by SimSDP. The radio astronomy imaging pipelines will be described at a high-level of abstraction suitable for targeting any heterogeneous multinode HPC system SKA may choose in the future. Several SDP architecture configurations (number of nodes, type of accelerators) and several SDP algorithm configurations will be then explored.