Excerpt of event, available at:
https://www.fsb.unizg.hr/parnum2019/abs_book_web.pdf

Reduction of Radio Astronomical Observations to Sky Maps as an Example of Exascale Data Service

Ladislav Hluchý, Jason Maassen, and Hanno Spreeuw

Abstract. The reduction of radio astronomical observations to sky maps is one of the use cases of the PROCESS project[1] which brings together high performance computing and exascale data management (in cooperation with U-COMP project[2]).

Its goal is to deliver a comprehensive, mature and modular service-based set of solutions and tools, all available to the community as open source packages and specially developed to enable extreme scale data processing in both, scientific research and advanced industry settings. The final results of PROCESS will be a set of extreme data focused services for exascale systems, driven by the requirements of five representative pilot use cases.
For each of them, a service prototype will be assembled to demonstrate the usefulness of the PROCESS solutions in real-world settings.
The main requirement coming from the reductions of radio astronomical observations to sky maps is a portable approach to support workflow infrastructure for analysis of large volumes of data. The amount of data produced by Low-Frequency Array (LOFAR) is now expanding at a rate of approximately 5-7 PB/year. The current volume of LOFAR Long Term Archive (LTA) is about 43 PB. For Square Kilometre Array (SKA) this problem will increase significantly, as the amount of data produced will be at least ten times larger.
Reduction workflows should scale both horizontally and vertically to be able to cope with extremely large (potentially streaming) data sizes. Within the PROCESS project, we explore a mechanism to run containerized workflows, thereby improving the portability and easy of use. Running the same workflow in parallel on different dataset provides the horizontal scalability required for distributed scientific instruments such as LOFAR or SKA. Vertical scalability will be achieved by applying multi- and many-core techniques.

[1] "PROviding Computing solutions for ExaScale ChallengeS" (PROCESS) project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 777533.

[2] Project APVV-17-0619 (U-COMP) "Urgent Computing for Exascale Data"