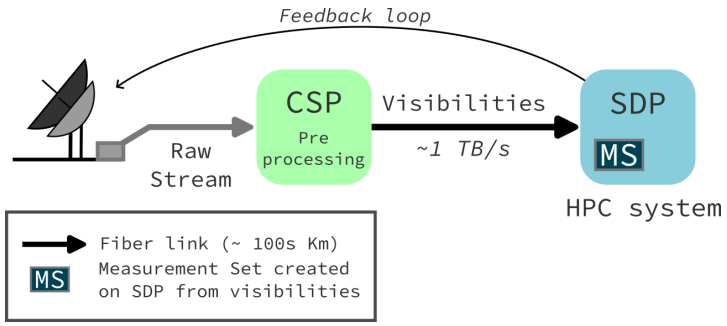
Analysis and Optimization of a File Format for Large-Scale Radio Astronomy

Advisors

- François Tessier (Inria) - francois.tessier@inria.fr
- Jakob Luettgau (Inria) - jakob.luettgau@inria.fr
- Erwan Raffin (Eviden) - erwan.raffin@eviden.com

Subject

Nowadays, there are many scientific fields where the need for computing power and data processing capacity goes beyond what current machines can provide. In radio astronomy, for example, the international SKA project aims to create the largest telescope in the world in order to observe a part of the universe. A very large volume of data is generated at the telescope level, pre-processed on local clusters (filtering, reduction) in real time and sent to a supercomputer at a rate of 1TB/s. This data feeds numerical simulation, generating 1PB of daily output data that needs to be saved. At this stage, the computing power and storage resources required are such that machines capable of reaching the **exascale** become necessary. To date, only a few supercomputers such as Frontier at Oak Ridge National Laboratory (USA) have this capability, but in the coming months, new systems will be deployed. However, the efficient use of these systems raises new challenges, especially regarding **data management**.



One clearly identified area for optimization concerns the way in which data is stored. Indeed, the data captured by the telescopes is written into **Measurement Sets (MS)** [1,2]. A MS file, which weights from a few dozen MB to several hundreds of GB, can be seen as a directory where a collection of files store diverse structured data and metadata. When such a file is read (often multiple times) for data analysis and processing, the data is distributed to processes according to several dimensions (temporal, frequency, spatial, resolution). Following a data analysis and processing phase, new data can be added to a MS although data integrity must always be guaranteed. This data format is a standard in the world of radio astronomy. Over the years, however, it has shown its **limitations in terms of scaling**. In particular, the library used to

manipulate these files, **Casacore** [3], has very limited support for parallel I/O although there have been improvements recently [4], while the access pattern is poorly optimized for multiple and sometimes concurrent accesses [5]. Ultimately, the community needs the format to evolve to accommodate very large-scale systems, particularly in view of SKA's production release.

The aim of this internship will be to **study the I/O behavior** of a key component in the data processing pipeline typical of a large-scale telescope: the Measurement Set file format. More concretely, the selected student will begin by familiarizing himself/herself with a **distributed computing environment** (computing cluster, compilation, execution of parallel applications). He or she will then **run and monitor a few Python-based components** managing MS files. Creating a simple benchmark to start with could also be an option. The student will then study the collected I/O traces, ideally using a Jupyter notebook. The aim will be to determine the potential **I/O bottleneck** of the Measurement Set format and suggest ways of optimizing. These optimizations will be implemented in the Casacore library and evaluated on SKA pipelined components.

The student will join the **KerData research team at Inria Rennes** and will be co-advised by **Eviden**. Inria is the French leading public research institute in computer science. KerData is a research team comprising six researchers, two engineers and six PhD students. The team's main theme is data management across the computing continuum, i.e. all computing resources from the Edge to HPC/Cloud infrastructures. Part of the team has a particular focus on I/O and storage issues on large-scale systems. **This work has no confidentiality requirements.**

Skills and abilities

- Programming skills (Bash, Python)
- Knowledge of computer networks and distributed systems
- Familiarity with high-performance computing or cloud computing is an advantage
- A good level of English is required

Bibliography

[1] https://casa.nrao.edu/Memos/229.html
[2] https://casadocs.readthedocs.io/en/stable/notebooks/casa-fundamentals.html#MeasurementSet-Basics
[3] G.N.J. van Diepen, Casacore Table Data System and its use in the MeasurementSet, Astronomy and Computing, Volume 12, 2015, Pages 174-180, ISSN 2213-1337, https://doi.org/10.1016/j.ascom.2015.06.002
[4] Dodson, Richard & Williamson, Alex & Gong, Qian & Elahi, Pascal & Wicenec, Andreas & Rioja, Maria & Chen, Jieyang & Podhorszki, Norbert & Klasky, Scott. (2024). Optimising the Processing and Storage of Visibilities using lossy compression. 10.48550/arXiv.2410.15683.
[5] R. Wang *et al*., "Processing Full-Scale Square Kilometre Array Data on the Summit Supercomputer," *SC20: International Conference for High Performance Computing, Networking, Storage and Analysis*, Atlanta, GA, USA, 2020, pp. 1-12, doi: 10.1109/SC41405.2020.00006.