

EVIDEN



# ECLAT

## Extreme Computing Lab for Astronomical Telescopes

Joint Laboratory

1<sup>st</sup> Technical Workshop, 27-29 Nov. 2024

Erwan Raffin  
Collaborative Project Leader  
HPC Distinguished Expert





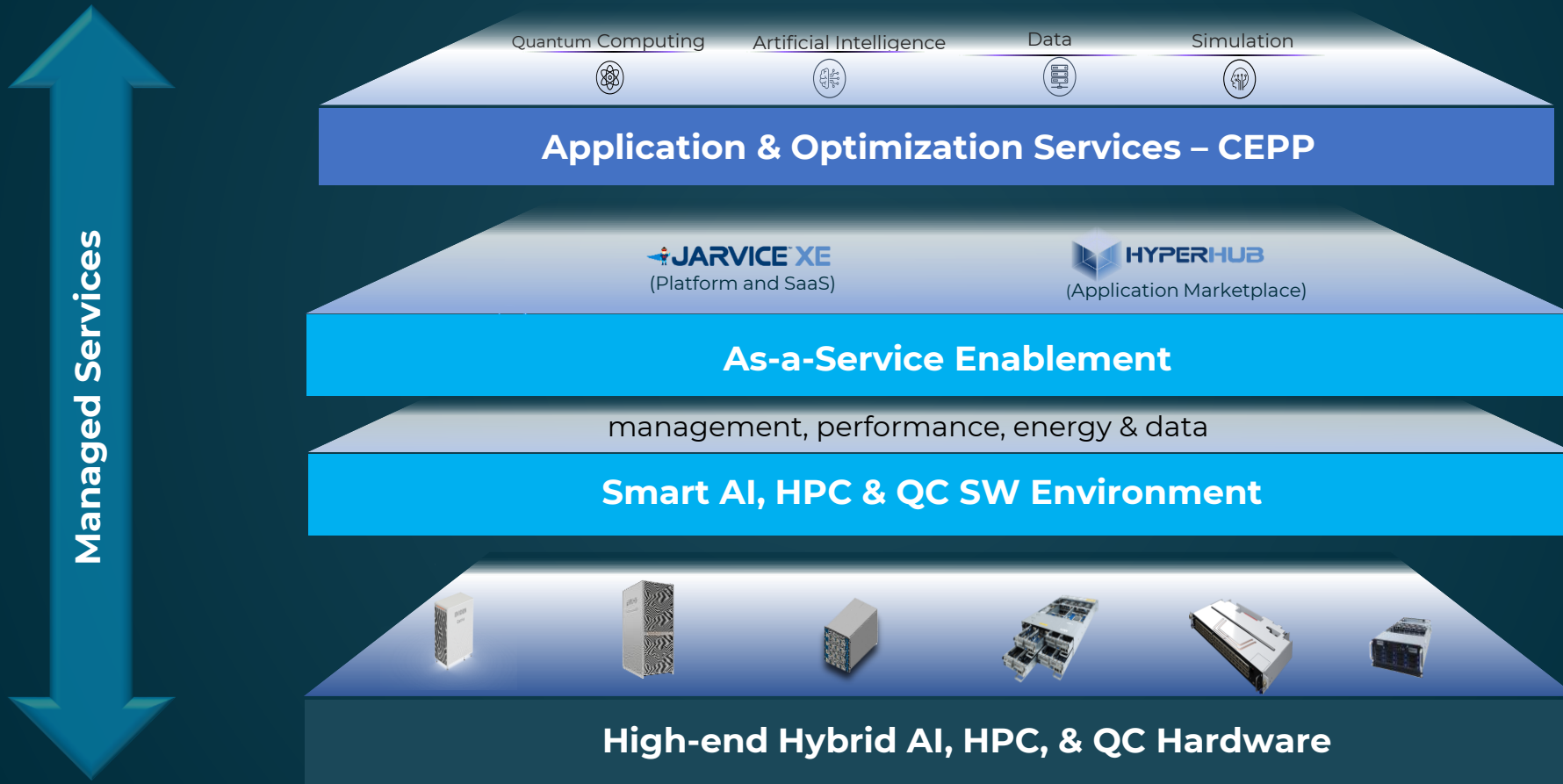


EVIDEN

# 01 Eviden CEPP – Center for Excellence in Performance Programming

# The Most Advanced High-end AI, HPC, & Quantum Portfolio

On-premises and in the cloud



# Eviden CEPP – Center for Excellence in Performance Programming

## Collaborative Ecosystem: Our Key Partnerships

### Application Optimization



**Institutions**

Science research funded projects where CEPP does HPC services & co-design

**IT Partners**

Focus on key IT partners in HPC, AI & Quantum

**Customers**

Taylor-made solutions on customer KEY applications and topics

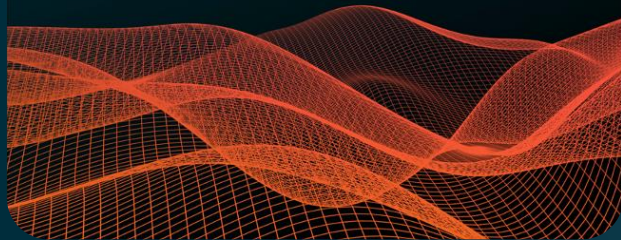
**R&D**

Co-design and co-creation between R&D for applications

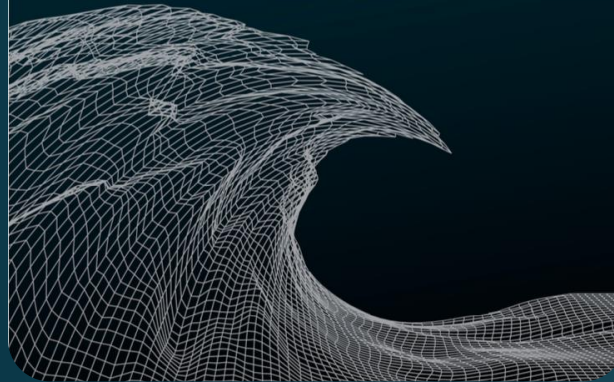
# CEPP – Modular service offers adapted to your needs

4 principal service offerings

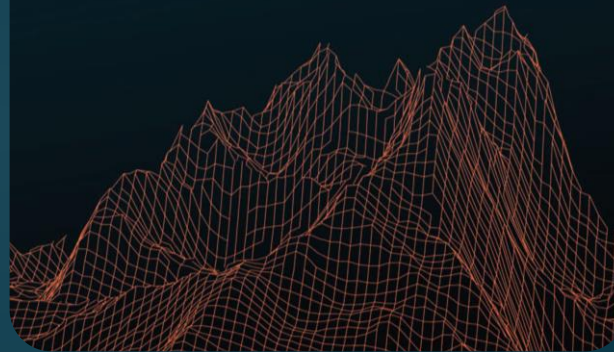
**CEPP one+**  
Accelerate workload,  
give value to simulation!



**CEPP Acceleration**  
Accelerate workload,  
give value to simulation!



**CEPP Training**  
Accelerate workload,  
give value to simulation!



**CEPP AI4SIM**  
Accelerate workload,  
give value to simulation!



**AIRBUS**

INERIS

ECMWF

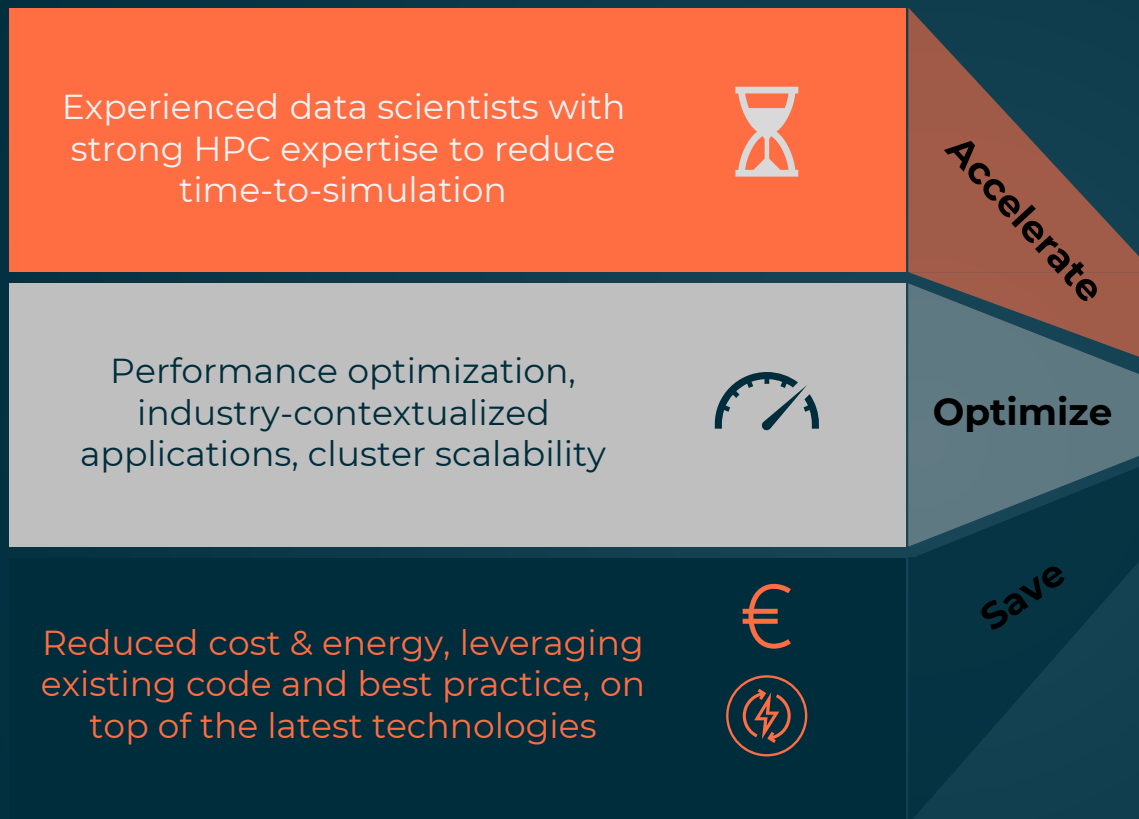
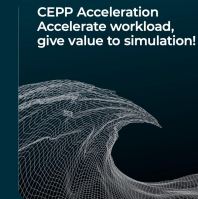


ECMWF



# CEPP – Center for Excellence in Performance Programming

Accelerate workload, give value to simulation



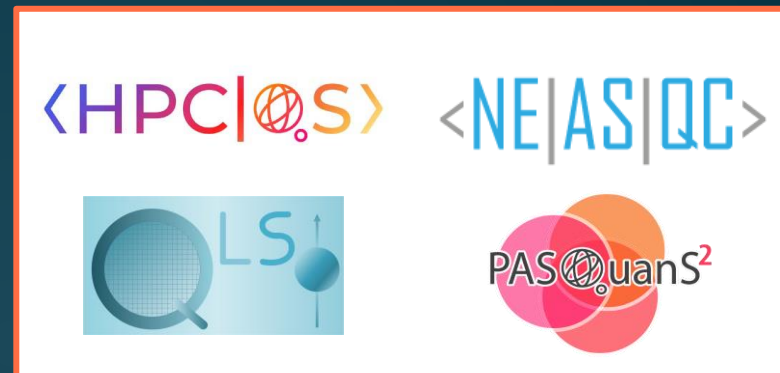
# Center for Excellence in Performance Programming (CEPP)

CEPP Collaborations - HPC AI and Quantum collaboration projects

European technology for exascale



Quantum computing



Cloud



Centre of Excellence for computing applications



Customer Centre of Excellence



International collaboration





EVIDEN

## 02 Eviden CEPP and the SKA project

# SKA and CEPP

## A long terms collaboration

- Bull/Atos/Eviden is collaborating for the SKA since 2013
  - 4 co-authors from Eviden of the “French SKA White Book - The French Community towards the Square Kilometre Array” for section “Industrial perspectives and solutions”
  - Participation in the 3 SKA France Day events
  - Imager DDFacet benchmarking, profiling and parallelization
  - Exa-SKA PhD thesis co-funding
    - DDFacet new parallelization implementation
    - New algorithm for Gridder/Degradder named G2G, CPU and GPU implementation
  - IDG porting from CUDA to HIP
  - HPC architecture exploration for calibration and imaging in line with SKAO teams SCHAAP & SCOOP
  - G2G: separation of the CPU version (public) and the GPU one (on demand)
  - Data Management: IO profiling and optimization of DDFacet



### CEPP team:

- David Guibert
- Loris Lucido
- Erwan Raffin

### R&D team:

- Sylvie Lesmanne (Hw)
- Grégoire Pichon (Sw)
- Philippe Couvée (Sw)

And many others

# SKA and CEPP

## Ongoing and Future collaboration

- Ongoing
  - ECLAT: Eviden partie contributrice & membre du Comité de Pilotage
  - DARK ERA (ANR): Eviden part of the advisory board of the project
- Future
  - to be found!
  - Please wait until the end of this presentation!!!





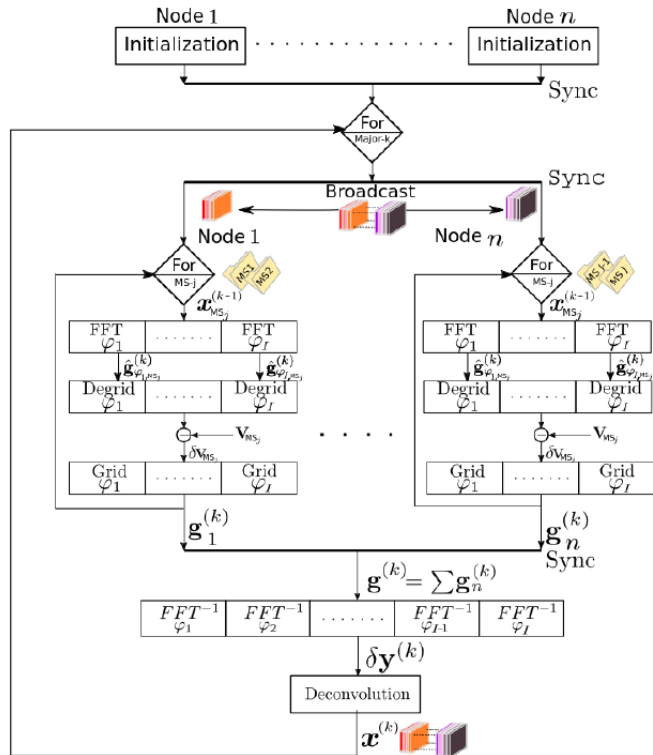
EVIDEN

## 03 CEPP's contributions

# Collaboration on DDFacet parallelization

## Overview

### Big picture



### Multi-node parallelization with data distribution

This multi-node implementation<sup>1</sup> makes two levels of parallelism of DDFacet.

- 1 the first one is based on facets parallelization and shared memory
- 2 the second one is based on **Measurement Set parallelization** which allows for a distributed memory system.

- if  $I_{facets} < nb_{cores}$ : under-utilization of the computational resources.
- if we put more MPI processes per node, we can reach more optimal usage of computational resources when

$$I_{facets} \geq \frac{n_{cores}}{n_{mpi}}$$

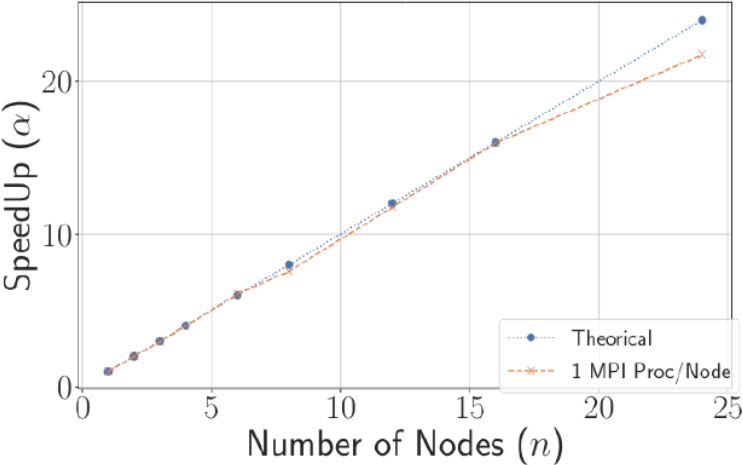
<sup>1</sup> SIPS IEE 2022 paper: Multi-core multi-node parallelization of the radio interferometric imaging pipeline DDFacet

# Collaboration on DDFacet parallelization

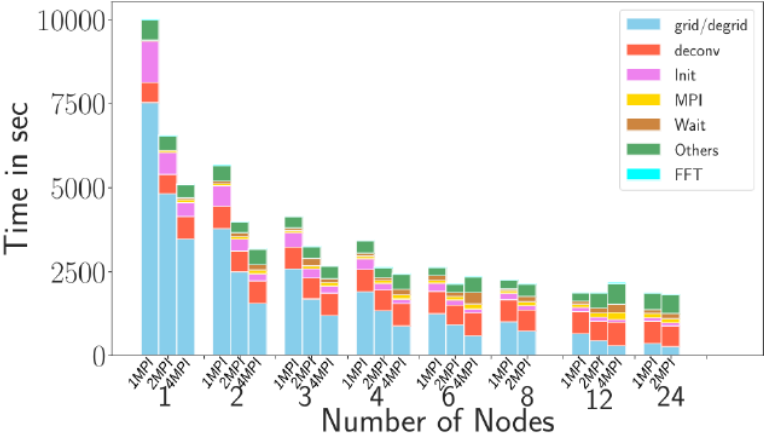
## Results

### Performance of the MPI parallelisation

- scaling of Gridding/Degridding with 1 MPI process per node



- Profiling of DDFacet for a complete execution with 1, 2 or 4 MPI Proc per node and from 1 to 24 nodes





# Collaboration on DDFacet parallelization

## Ongoing activity

- Merge Request under investigation by DDFacet developers

The screenshot shows a GitHub pull request titled "[WIP] distributed implementation using mpi rebased on master #888". The pull request is from the branch "dg/mpi-rebased-on-master" to the "master" branch, containing 22 commits. The author is "dguibert". The pull request is currently in a "DO NOT MERGE" state. A comment from "dguibert" dated August 23 explains that the patch allows running DDFacet on multiple nodes using mpi4py and provides a terminal command: `mpirun -n $number_of_processes DDF.py ....`. The comment also notes that the pull request is rebased on top of master from a specific GitHub commit. The right sidebar shows the reviewer "bennahugo" with a message: "Requested changes must be addressed to merge this pull request." The sidebar also includes sections for "Reviewers", "Assignees", and "Labels".

[WIP] distributed implementation using mpi rebased on master #888 Edit <> Code

Open dguibert wants to merge 22 commits into `master` from `dg/mpi-rebased-on-master`

Conversation 43 Commits 22 Checks 0 Files changed 33 +469 -198

**dguibert** commented on Aug 23

This patch allows to run DDFacet on multiple nodes using mpi4py.  
It can be launched as classical mpi program as:

```
mpirun -n $number_of_processes DDF.py ....
```

NB: this is rebased on top of master from [https://github.com/cyriltasse/DDFacet/tree/mpi\\_mergeMaster\\_mergeAPPkMS](https://github.com/cyriltasse/DDFacet/tree/mpi_mergeMaster_mergeAPPkMS).

**dguibert** added 7 commits 4 months ago

Reviewers

**bennahugo**

Requested changes must be addressed to merge this pull request.

Still in progress? [Learn about draft PRs](#)

Assignees

No one—[assign yourself](#)

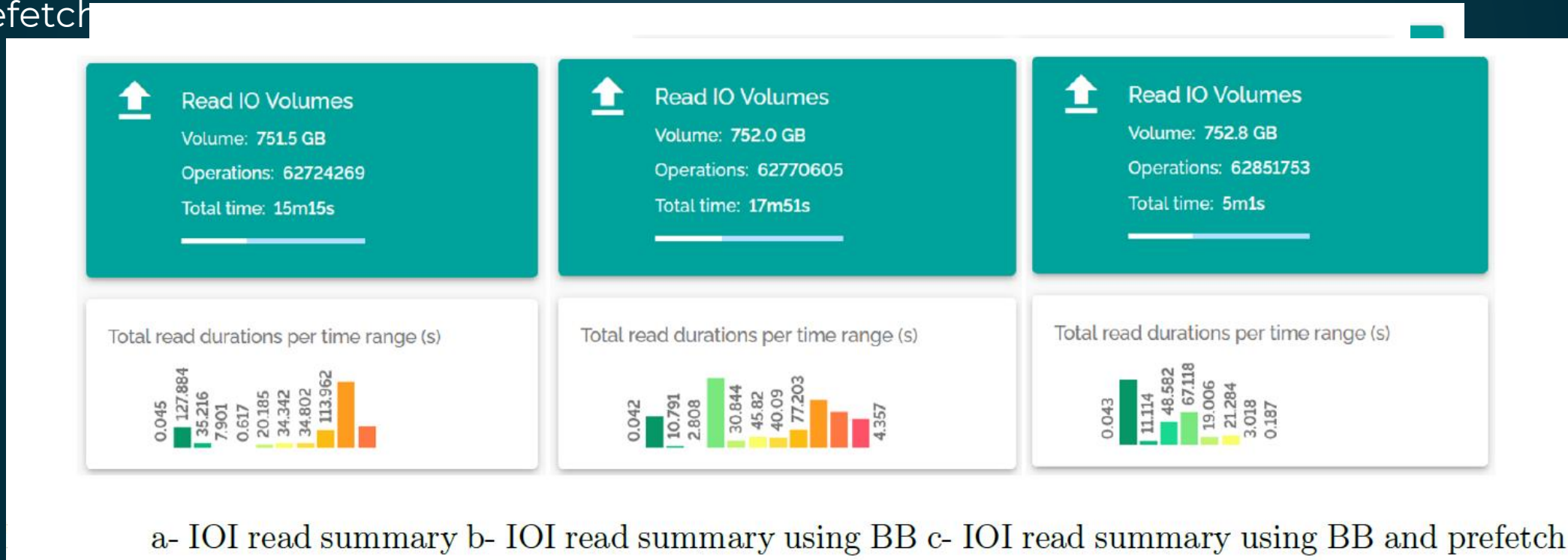
Labels

**DO NOT MERGE**

# DDFacet IO characterization

Data Management SW R&D – Internship of Armand Grenier

- Characterization of the DDFacet Input/Output using IOI (IO Instrumentation) tool
- Burst Buffer
- + Prefetch



a- IOI read summary b- IOI read summary using BB c- IOI read summary using BB and prefetch

# DDFacet IO characterization SKAO imager pipeline study and optimization

## DP3 et WSClean – Internship of Clément Devatine

- DP3 and WSClean chosen by SKAO as imager pipeline candidate
- Difficult installation on Eviden internal benchmark supercomputer
- Docker image,
- Singularity image,
- By hand
- With Spack
- Performance analysis of DP3  Low use of the computational resource
- Proof-of-concept with OpenMP replacing the custom implementation of the internal task-based parallelism
  - 10% to 30% gain according to the test case
  - Estimation of a speedup of 2,5x with the proposed optimizations
- Presentation to SKAO in summer 2023.



# Gridder/Degridder kernel

## IDG and G2G – GPU approach

- IDG - Image Domain Gridding
  - Porting CUDA implementation to HIP CUDA and HIP ROCm
  - Performance comparison Nvidia V100 vs AMD MI100
- G2G
  - Following the work of Nicolas Monnier on the Fast grid to grid interpolation for radio interferometric imaging and its CPU and GPU implementation
  - Code takeover by CEPP
    - Test and comparison of the CPU and GPU version
    - Extraction of the CPU version to make it publicly available
    - GPU version is available on-demand for partners

EVIDEN

## 04 New collaboration

# SKA and CEPP

## New collaborations

- Eviden is open for collaborations
- Our goal is teaming up for exploratory work toward helping at co-designing and co-sizing the SPCs
  - Application:
    - DDFacet, Grid/Degrid, KillIMS, Casacore & Dask MS
    - Dataflow model and tool: SimSDP
  - Hardware architecture
    - BXI v3 Interconnect
    - Smart Data Management
  - Exploration:
    - Interconnect & data movement solution
    - CPU: ARM based platforms, HBM
    - GPU: portability and performance

### CEPP team:

- David Guibert
- Loris Lucido
- Erwan Raffin

### R&D team:

- Sylvie Lesmanne (Hw)
- Grégoire Pichon (Sw)
- Philippe Couvée (Sw)

And many others

# SKA and CEPP

## New collaborations

- Application
  - Pipeline DDFacet parallelization, benchmarking, profiling and optimization
    - Vectorization is crucial, is Gridder/Degridder well vectorized?, any GPU version envisioned?
    - KillMS parallelization, benchmarking, profiling and optimization?
    - IO parallelization?
  - G2G
    - GPU version portability and optimization
  - Investigate non-uniform FFT methods and implementations and compare it to DDFacet?



# SKA and CEPP

## New collaborations

- Interconnect
  - Exploiting BXI offload : update applications to take benefit of BXI offload features (non-blocking MPI calls, collective communications) and optimize compute/communication overlap
  - Improve AI communications : characterize the communication profile of AI workloads to optimize next BXI versions
  - Near real-time communications : how to manage hi-bandwidth continuous communications between Stream Processing and Science Data Processor ?

# SKA and CEPP

## New collaborations

- Data Management
  - Multimetrics profiling of the SKA pipeline in different processing configurations, efficiency evaluation and if relevant, memory optimisation
  - Energy efficiency of the SKA pipeline, power optimization & capping opportunities
  - Ephemeral IO Services/Datanodes interest in the SKA Science Data Processor

# EVIDEN

## Thank you!

For more information please contact:  
[erwan.raffin@eviden.com](mailto:erwan.raffin@eviden.com)



2019, SKAO HQ, Jodrell Bank



2022, SKA France Day



2023, South Africa

Confidential information owned by Eviden SAS, to be used by the recipient only. This document, or any part of it, may not be reproduced, copied, circulated and/or distributed nor quoted without prior written approval from Eviden SAS.

© Eviden SAS