RÉPUBLIQUE FRANÇAISE
Liberté Égalité Fraternité

PROGRAMME DE RECHERCHE
NUMÉRIQUE POUR L'EXASCALE

WP4: Applications

GYSELA code: PLasma turbulence in tokamaks

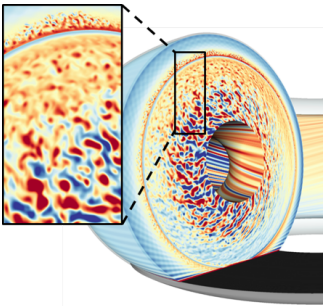
Virginie Grandgirard

September 18,2024

IRFM/CEA Cadarache

GYSELA code for exascale plasma turbulence simulations

- Gyrokinetic plasma turbulence simulations to understand turbulent transport that mainly governs confinement in Tokamaks

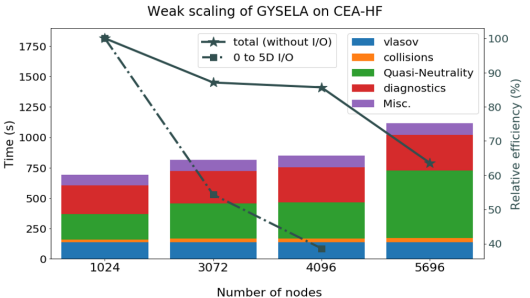


- GYSELA = Fortran 90 code with hybrid MPI/OpenMP parallelisation optimized up to 730,000 cores
  - Relative efficiency of 85% on more than 500k cores and 63% on 730k cores on CEA-HF (AMD EPYC 7763)
- Intensive use of petascale resources: $\sim$ 150 millions of hours / year (GENCI + PRACE + HPC Fusion resources)
- **Exascale needs for ITER plasma turbulence simulations**
  - Rewritting of the code in C++ via Kokkos (EoCoE-III project)
  - **I/O optimisation + In-situ diagnostic (Exa-Dost project)**

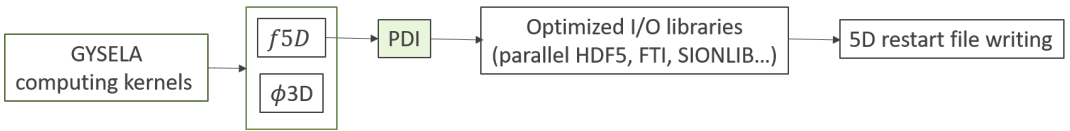
cea cnrs Inria

GYSELA: I/O a strong bottleneck for exascale (1/2)

- I/O scalability is an issue: $\sim$ 50% for 3072 nodes and 38% for 4096 nodes. Crash on 5696 nodes.
- 22784 MPI process $\rightarrow$ 22784 HDF5 files written at the same time equiv. to 16.2 TBytes.



- Use PDI to let I/O optimization to I/O specialists



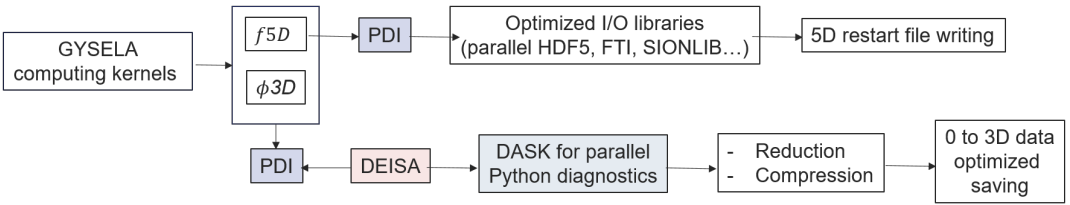
GYSELA: I/O a strong bottleneck for exascale (2/2)

- **Work in Progress with WP1**
  - Contact persons for GYSELA: Dorian Midou + Kevin Obrejan

- **PDI installed both in Fortran code and C++ prototype**
- Meeting in April 2024 with François Tessier (INRIA Rennes) + Francieli Zanon Boito (INRIA Bordeaux) + Jean-Thomas Acquaviva (DDN)
- IRFM Internship: Méline Trochon (03/2024-07/2024)
  - Checkpointing optimization via PDI + parallel HDF5
  - PhD starting in oct. 2024 with François Tessier + Francieli Zanon Boito.

GYSELA: Optimized Data workflow (1/2)

- Main idea : Decouple I/O from computing kernels
- Development of in-situ diagnostics framework based on PDI + DEISA + DASK
  - PDI Data Interface for handling I/O (developed at MDLS)
    `https://pdi.julien-bigot.fr/master/`
  - DEISA (dask-enabled in situ analytics) library (developed at MDLS)
  - DASK a flexible library for parallel computing in Python
    `https://docs.dask.org/`



GYSELA: Optimized Data workflow (2/2)

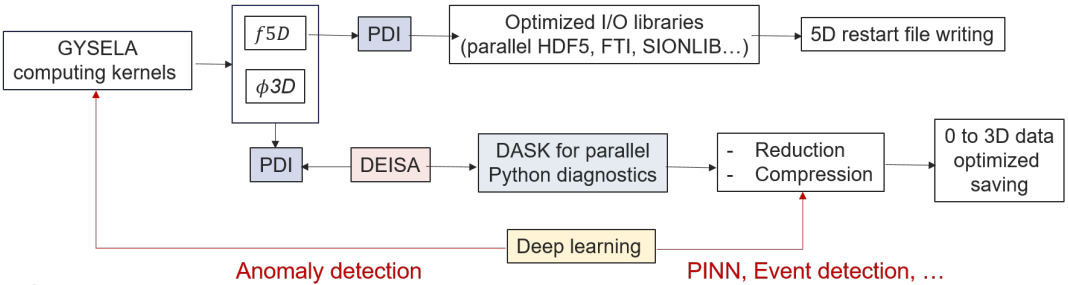
- **Work in Progress with WP2**
  - Contact persons for GYSELA: Virginie Grandgirard + Kevin Obrejan

- GYSELA meeting at MdlS on Tuesday 17th September 2024 - PM
  - CEA MdlS: Yushan Wang + Bruno Martin + Julien Bigot.
  - INRIA Grenoble: Bruno Raffin + Andres Bermeo Marinelli (EoCoE-III Engineer starting in Sept. 2024)
  - CEA IRFM: Virginie Grandgirard + Dorian Midou.

- **GYSELA Fortran + PDI + DEISA : proof of concept done** during Amal Gueroudji's PhD.
- C++ prototype + PDI + DEISA: work for Andres via EoCoE-III

How can AI plays a role to optimize I/O ? (1/2)

- Data Compression optimization
- In-situ AI diagnostics to optimize exascale simulations:
  - Automatic anomaly detection → Automatic stop of simulation → CPU or GPU consumption optimization
  - Automatic rare event detection → Optimisation of diagnostic saving → Memory storage reduction

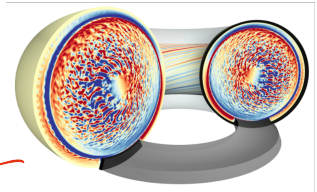


How can AI plays a role to optimize I/O ? (2/2)

- **Work in progress with WP3**
  - Contact persons for GYSELA: Virginie Grandgirard + Feda Almuhisen

- Collaboration with MdlS via PTC-simulation CEA (Mathieu Lobet + Martial Mancip + Hiba Taher)
- Collaboration with NTU Singapore university via SAFE (Robin Varennes)

- GYSELA meeting at MdlS on Tuesday 17th September 2024 - AM
  - INRIA Saclay: Thomas Moreau + Mansour Benbakoura (Post-Doc starting in Oct. 2024)
- **First objective: Turbulence dynamics detection with Tokam2D fluid code**



Analogy with
2D fluid turbulence