UNIVERSITÀ DEGLI STUDI DI NAPOLI FEDERICO II

itee PhD information technology electrical engineering

DIETI UNINA

Gianluca Sabella

Benchmarking Approaches for Evaluating Heterogeneous Computing Resources and High-Performance Clusters Deployed via Kubernetes

Tutor: Prof. Carlo Sansone

co-Tutor: Prof. Elvira Rossi

Cycle: XXXVIII

Year: Second

itee PhD information technology electrical engineering

Candidate's information

- MSc degree in Computer Science from University of Naples "Federico II"
- Research groups: PICUS, ICSC - Spoke2 – WP2 and WP5, INFN Sezione di Napoli, Napoli's group of the ATLAS experiment at CERN
- PhD start date: 01/01/2023
- PhD end date: 31/12/2025
- Scholarship type: no funded scholarship

Summary of study activities

Courses

- Mini-Workshop on Data Management
- VHDL course

Events attended

- The IBiSCo Project and the Transition Towards the Italian National Center fot HPC, Big Data and Quantum Computing (ICSC), Napoli, Italy, 18-19 April 2024, presenting author – The User Experience in IBiSCo Data Center.
- Congresso Nazionale della Società Italiana di Fisica (SIF 2024), Bologna, Italy, 9-13 September 2024, presenting author – Benchmarking distributed-interactive HEP analysis workflows on the new Italian National Centre analysis infrastructure.

Courses	Seminars	Research	Tutorship	Total
5,6	5,3	49	0	59,9

Research area(s)

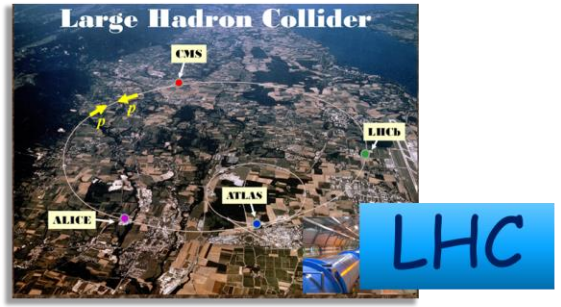
- ***High-Performance Computing (HPC) heterogeneous resources and benchmarking:*** Delving into the realm of supercomputers and computing architectures that combine CPU and GPU resources, evaluating and leveraging benchmarks to optimize scientific applications

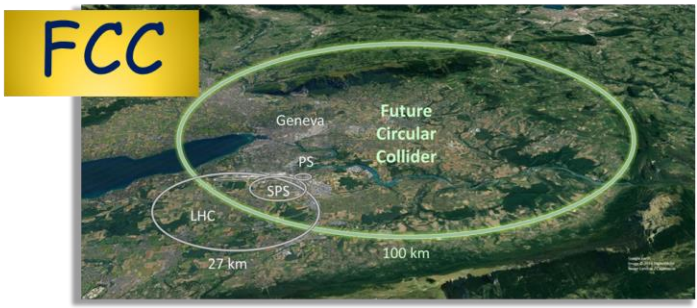
- ***High Performance Resource Containerization:*** Exploring lightweight, easily deployable container technology for efficient application delivery and efficient use of high-performance computing resources

Research results

❑ **Problem**:

Challenges of Large Hadron Collider (LHC), High Luminosity LHC (HL-LHC) and of the Future Colliders are pushing to re-think the HEP computing models having strong impact on several aspects, from software to the computing infrastructures.





❑ **Objective**:

From the software perspective, interactive/quasi-interactive analysis is a promising paradigm

- ❖ User-friendly environment
- ❖ Adopting open-source industry standards: Dask, Jupyter Notebooks, HTCondor
- ❖ Validating new frameworks (e.g. ROOT)

dask

jupyter

HTCondor Software Suite

ROOT Data Analysis Framework

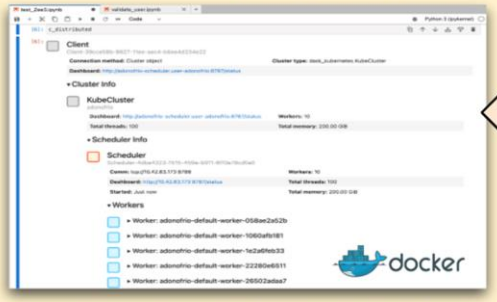
Research results (2)

Access and security

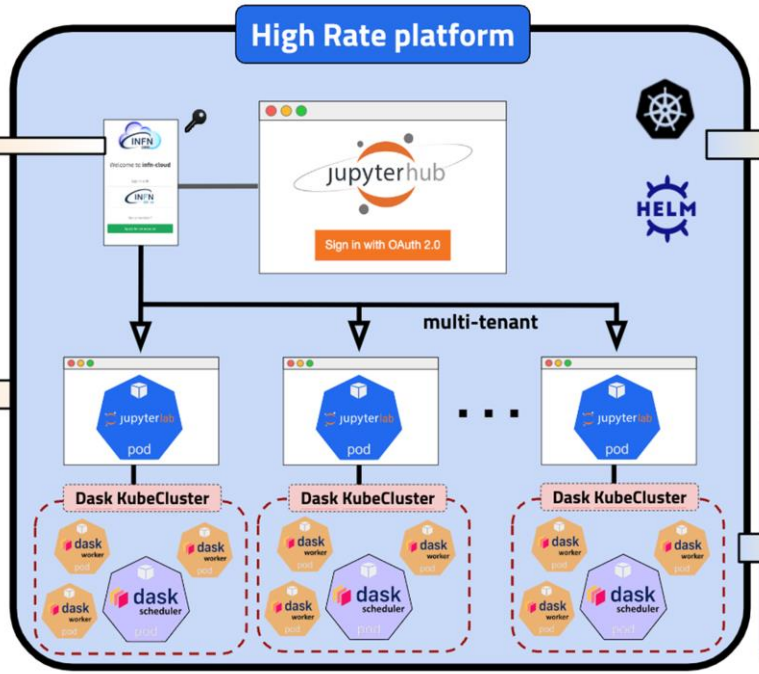
After connecting to an entrypoint URL, the user reaches a **Jupyterhub** instance that, after authentication and authorization via INDIGO-IAM, allocates the required resources for the user's working area

User Interface

The user interface is based on **Jupyterlab**, customised with specific plugins for specific purposes (e.g. Dask).



The working environment is highly customizable, using tailored Docker containers. This is important when analyses require specific software (collaboration-wise)



Deployment

The deployment of the **Kubernetes** resources needed for the spawning of this platform, is handled via **HELM charts** available in the GitHub organization.

Check the docs!

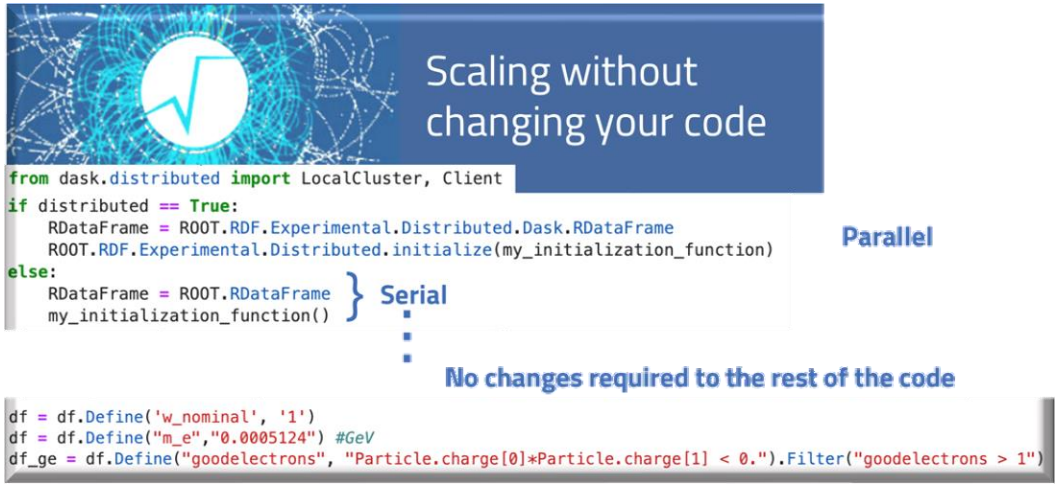
This allows a seamless, flexible, scalable and fault-tolerant deployment on the available resources, with a limited impact on the admin's work time

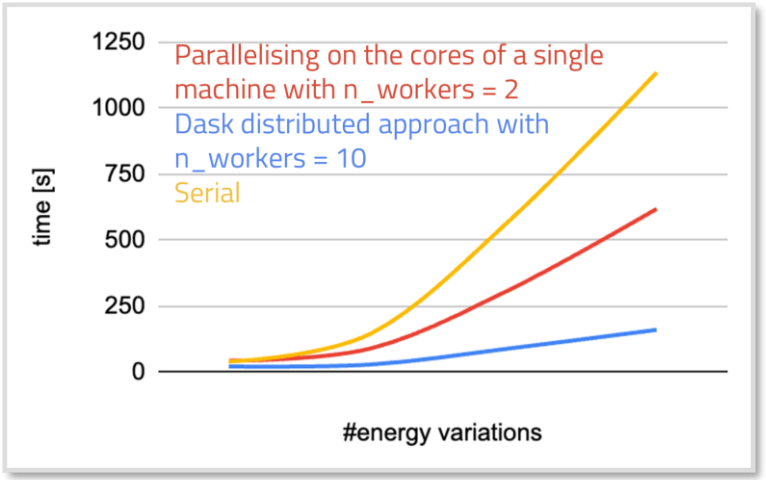
Software

From the software perspective, interactive/quasi interactive analysis is a promising paradigm

- User-friendly environment
- Adopting open-source industry standards: *Dask, Jupyter Notebooks* and *HTCondor*
- Validating new frameworks (e.g. *ROOT RDataFrame* with multi-threading)

Research results (Use case 1)





How to compare the performance?

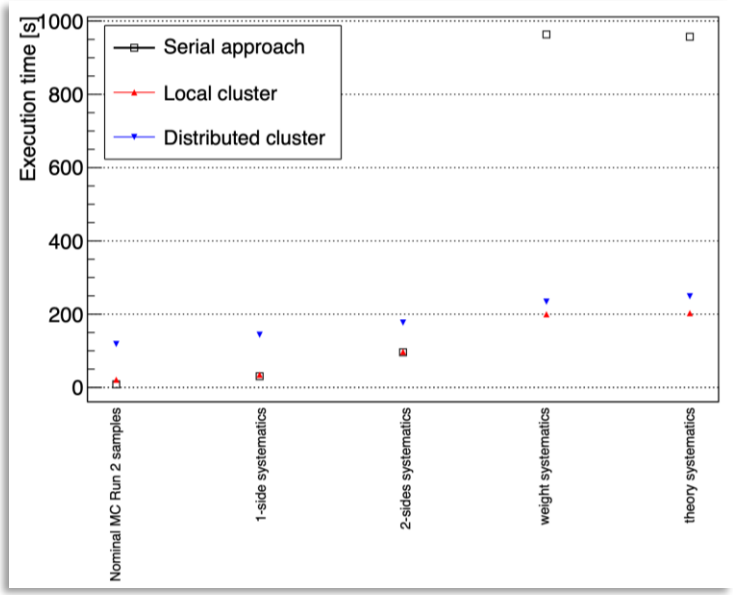
Defined Metric	
Overall execution time	Time elapsed from the start of the execution (execution triggered) to the end of execution

- ✓ Exploiting the local client approach, the execution time improves *wrt* the standard/serial approach if we iterate over a significative number of energy variations (> 10)
- ✓ Moving to a distributed Dask model and **scaling resources, the performance improves**

Research results (Use case 2)

Exploiting the distributed approach, the execution time improves *wrt* the standard/serial approach if we iterate over a significative number of systematic variations

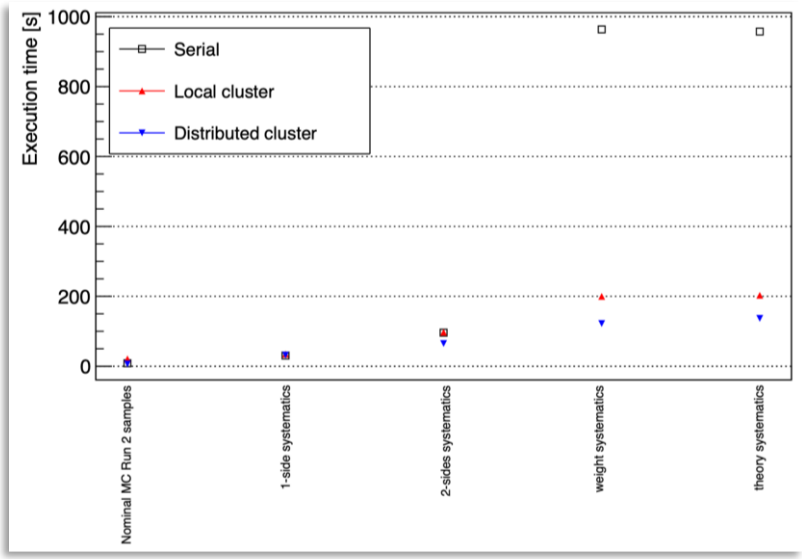
Execution time including spent to copy the inputs to the worker nodes (~120 s)



Limiting factor for the distributed approach: time spent to copy the inputs to the worker nodes (~120 s) -> **TO BE IMPROVED**

Defined Metric	
Overall execution time	Time elapsed from the start of the execution (execution triggered) to the end of execution

Execution time not including spent to copy the inputs to the worker nodes (~120 s)



Next year

- **Replication on bare metal**
  Development of a bare-metal infrastructure to directly compare performance with the cloud configuration.

- ✓ **Goal**
  Identify the strengths and weaknesses of both systems.

- ✓ **Specific benchmarks**
  Exploring the CORAL-2 benchmark to evaluate scalability, response times, and resource management.

Next year (2)

- **Use-case diversification**

  Extending the infrastructure to fields beside high-energy physics

- ✓ **Biomedical and engineering fields**

  Areas with large-scale analysis and distributed computing needs

- ✓ **Goal**

  Demonstrate the platform's adaptability and versatility across diverse sectors.

Research products (1)

[P1]	Carracciuolo L, Hegde P R, Sabella G, Spisso B, Tortora A., *The User Experience in IBiSCo Data Center,* **The IBiSCo Project and the Transition Towards the Italian National Center for HPC, Big Data and Quantum Computing (ICSC),** cap. 15, pp. 156 – 161, Zenodo, 2024, doi:10.5281/zenodo.13120590
[P2]	Barone G B, Bottalico D, Carracciuolo L, Michelino D, Sabella G., *Design, Implementation and Validation of a Heterogeneous Resource for High Performance Computing,* **The IBiSCo Project and the Transition Towards the Italian National Center for HPC, Big Data and Quantum Computing (ICSC),** cap. 12, pp. 109 – 122, Zenodo, 2024, doi:10.5281/zenodo.13120590
[P3]	Bottalico D, Izzo A, Michelino D, Pardi S, Sabella G, Sociale R, Tortora A., *The Monitoring System in the IBiSCo Data Center,* **The IBiSCo Project and the Transition Towards the Italian National Center for HPC, Big Data and Quantum Computing (ICSC),** cap. 16, pp. 163 – 174, Zenodo, 2024, doi:10.5281/zenodo.13120590
[P4]	Russo G, Barone G B, Bottalico D, Carracciuolo L, Izzo A, Michelino D, Piccolo C, Sabella G., *The IBiSCo infrastructure realized at UNINA,* **The IBiSCo Project and the Transition Towards the Italian National Center for HPC, Big Data and Quantum Computing (ICSC),** cap. 5, pp. 43 – 53, Zenodo, 2024, doi:10.5281/zenodo.13120590

Research products (2)

[P5]	Spisso B, D'Onofrio A, Sabella G, Cirotto F, Doria A, Carlino G., *The HPC IBiSCo seeds for INFN experiments workflow based cluster,* **The IBiSCo Project and the Transition Towards the Italian National Center for HPC, Big Data and Quantum Computing (ICSC),** cap. 14, pp. 141 – 153, Zenodo, 2024, doi:10.5281/zenodo.13120590
[P6]	Battista, B.G., Gianluca, S. (2024), *Optimization of University Services Through Cloud-Edge Integration: Study of a Use Case at the University of Naples "Federico II"* **Advanced Information Networking and Applications, Lecture Notes on Data Engineering and Communications Technologies, AINA 2024** vol 203, Springer, Cham. Doi:10.1007/978-3-031-57931-8_12
[P7]	D'Onofrio A, Cirotto F, Gravili F G, Rossi E, Sabella G, Spiga D, Spisso B, Tedeschi T., *Benchmarking distributed-interactive HEP analysis workflows on the new Italian National Centre analysis infrastructure*, **42nd International Conference on High Energy Physics**, Prague, Czech Republic, July, 18-24, 2024, ICHEP
[P8]	Diotalevi T, Gravili F G, Anwar M, Bartolini M, Cagnotta A, D'Onofrio A, Mastrandrea P, Sabella G, Simone F M, Spisso B, Tarasio A, Tedeschi T., *Quasi interactive analysis of High Energy Physics big data with high throughput*, **22nd International Workshop on Advanced Computing and Analysis Techniques in Physics Research,** Stony Brook University, USA, March 2024, ACAT

Research products (3)

[P9]	Sabella G., Cirotto F, D'Onofrio A., Gravili F.G., Loffredo S, Rossi E., Spiga D., Spisso B., Tedeschi T., *Benchmarking distributed-interactive HEP analysis workflows on the new Italian National Centre analysis infrastructure*, **Congresso Nazionale della Società Italiana di Fisica,** Bologna, Italy, Sep., 9-13, 2024, SIF
[S1]	Battista, B.G., Gianluca, S., *Innovating Cloud-Based Academic Services: Advancing the University of Naples Federico II's Infrastructure with Terraform and Oracle GoldenGate* **Advanced Information Networking and Applications, Lecture Notes on Data Engineering and Communications Technologies, AINA 2025**