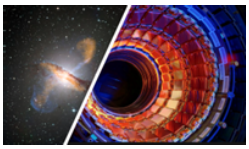


PhD course of National Interest in Technologies for
Fundamental Research in Physics and Astrophysics

Annual report

Name and surname: Saim Ali
Cycle and a.a.: 39th, 2023/24
Supervisor: Dr. Fabio Roberto Vitello, Dr. Eva Sciacca

Title: Analysis of astrophysical phenomena using efficient and parallelized models on HPC computing systems.

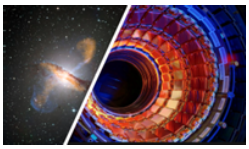
Research:
The research project aims to design and develop efficient, parallelized models for big data analysis, leveraging advanced visualization techniques and algorithmic solutions on modern HPC computing systems. The primary objective is to uncover astrophysical patterns from cosmological simulations.

Moreover, the project focuses on creating innovative solutions for processing large datasets and accelerating computation. This involves developing portable algorithms based on multi-platform paradigms and optimizing image reduction pipelines, potentially utilizing GPU platforms and AI libraries.

The project is built upon **VisIVO Server**, an open-source platform for data analysis and scientific visualization, designed for the fast rendering of 3D views of astrophysical datasets. **VisIVO Server** consists of three core components: **VisIVO Importer**, which converts user datasets into a simple and highly efficient internal data format used by VisIVO Filters; **VisIVO Filter**, a collection of data processing modules that construct data tables from the tables produced by VisIVO Importer; and **VisIVO Viewer**, which generates 3D visualizations of astrophysical datasets.

The VisIVO Server provides a variety of filters, with my research focusing on the **Point Distribute Filter**. This filter generates a table representing a volume derived from selected fields of the input table, distributed using the NGP (Nearest Grid Point), CIC (Cloud-In-Cell, default), or TSC (Triangular Shaped Cloud) algorithms. The filter exclusively accepts the VBT (VisIVO Binary Table) data format for processing.

The data for this research is provided in Gadget format from DEMNUni simulations. These simulations use the Tree Particle Mesh-Smoothed Particle Hydrodynamics (TreePM-SPH) code **GADGET-3**. The DEMNUni suite consists of large-scale cosmological N-body



simulations, referred to as the "Dark Energy and Massive Neutrino Universe" (DEMNUni). They track the evolution of Cold Dark Matter (CDM) and Hot Dark Matter (HDM) neutrino particles, treating them as two separate collisionless species.

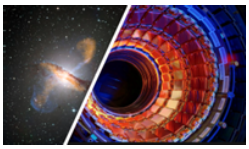
To enhance the efficiency of the Point Distribute Filter, a strategy for parallelization using MPI (Message Passing Interface) and OpenMP (Open Multi-Processing) will be developed.

1. **Data Distribution with MPI**
    - The data will be divided into smaller subgroups using MPI. This involves breaking the set of points into manageable chunks.
    - **MPI_Scatter** will be used to distribute these subgroups of points to different MPI processes. Each process will handle a portion of the dataset and perform the filtering operations.
2. **Parallel Processing with OpenMP**
    - Within each MPI process, OpenMP will be employed to parallelize the filtering tasks.
    - The **#pragma omp parallel** directive will create multiple threads to manage the workload efficiently.
    - The loop that processes the points will be parallelized using OpenMP to ensure that each thread handles a segment of the points.
3. **Data Collection and Communication**
    - After filtering, the results from each MPI process need to be gathered. **MPI_Gather** can be used to collect the processed points back to a single process, if necessary, or to distribute them to all processes as required.
    - Effective load balancing, minimizing communication overhead, and managing data dependencies are crucial for optimizing performance.

By leveraging both MPI for data distribution and OpenMP for parallel processing within each process, the Point Distribute Filter's efficiency can be significantly improved.

- **List of attended courses and passed exams**

1. Machine learning for Physics,	3 CREDITS	Preparing for exam
2. High energy physics detectors in space,	2,5 CREDITS	Exam done
3. Deep network and structure learning.	2 CREDITS	Exam done



UNIVERSITÀ DEGLI STUDI DI PADOVA

PhD course of National Interest in Technologies for Fundamental Research in Physics and Astrophysics

- **List of attended conferences, workshops and schools, with mention of the presented talks**
  - Course on computing and HPC for Astronomy and Astrophysics (24 June 2024 to 5 July 2024 - Bologna https://indico.ict.inaf.it/event/2785/)
- **List of published papers/proceedings**

  N/A
- **Thesis title (even temporary)**

  Analysis of astrophysical phenomena using efficient and parallelized models on HPC computing systems.

Date, 09/09/2024 **Signature**

Seen

the supervisor

Vitello Fabio Roberto

Eva Sciacca