Finanziato dall'Unione europea NextGenerationEU
Ministero dell'Università e della Ricerca
Italiadomani PIANO NAZIONALE DI RIPRESA E RESILIENZA
ICSC Centro Nazionale di Ricerca in HPC, Big Data and Quantum Computing

RISC-V4ASTRO

Jenny Almerol, E4 Computer Engineering S.p.A.

Spoke 3 Progetti Bandi a Cascata, 24/09, 2024

E4 in a Nutshell

E4 Computer Engineering is an Italian Company founded in 2002, E4 designs and manufactures highly technological solutions for **HPC** Clusters, **Cloud**, **Data Analytics**, **Artificial Intelligence** and **Hyper-Converged infrastructure** for the Academic and Industrial markets. We have been collaborating for years with the main research centers at national and international level (CINECA, CERN, ECMWF, LEONARDO) and we are involved in national and **European projects** in the HPC, AI and Quantum fields.

OUR MISSION

We anticipate the ever-accelerating disruptive transformation of our era,
providing mature and reliable solutions
in sophisticated technological environments with an innovative approach

OUR VISION

We explore future scenarios to find solutions to particularly complex computational demands
in application areas that were previously unimaginable

Project Overview

- High energy physics, astroparticles, and gravitational wave computations currently have the **highest computational demands .**
- **Energy consumption** has become a primary concern making it the largest component of TCO. Reducing energy consumption is also crucial for **CO_2 footprint** reasons.
- **New architectures designed from scratch offer better energy efficiency** (e.g., in Joules per operation) but are rarely used in scientific computing mainly for two reasons:
  - Large scientific experiments have extensive code bases
  - Low-power architectures generally have lower computational power density per rack unit

Thus, the development of scalable code for HPC systems using RISC-V CPUs, new GPUs, and FPGA accelerators is now a critical issue.

Project Overview

- **Scientific computing on RISC-V architectures is essential for fundamental research**, especially for next-generation astrophysical experiments requiring vast computational resources and energy.
- A major challenge is simulating the gravitational formation and evolution of dense stellar environments (e.g., star clusters) to understand gravitational wave sources.
- Accurate, direct N-body simulations of these systems are necessary but are **computationally complex with scalability of $O(N^2)$**.
- **Current direct N-body simulations are limited to smaller clusters** ($N<\sim10^5$), as larger clusters would require excessive time and power on modern GPU clusters.

New hardware/software combinations are needed to explore larger systems efficiently.

Objective

- The **RISC-V4ASTRO project aims to investigate the use of RISC-V (new hardware) for direct N-body simulations (new software)** to assess the applicability of these technologies for the astrophysical interpretation of gravitational waves and open a new window on the exploration of the densest stellar systems.
- The project aims to demonstrate with reliable and precise data derived from research activities on available or simulated data the current capacity of RISC-V-based green computing systems for scientific computing.
- Showing:
  - the scientific validity of the codes when executed on RISC-V machines
  - the advantage in terms of time reduction (time-to-solution) compared to other architectures and/or accelerators
  - the energy and TCO advantage of using such machines and accelerators instead of more standard architectures
  - documenting the porting processes

Objective

- The activities of the RISC-V4ASTRO project will have a significant impact both on **fundamental research activities** (impact on the scientific community) and on the **development of High-Performance Computing techniques on innovative hardware** (impact on industry). Specifically
  - A detailed analysis of the benefits obtained by switching platforms and/or accelerators,
  - Precise and realistic metrics can be condensed into dissemination messages (e.g., webinars, documents, workshops) for both scientific and industrial entities,
  - A push to design compatibility with RISC-V for future experiments as a fundamental feature of both scientific and industrial codes,
  - Make portions of these platforms accessible to external users via collaboration agreements.

Methodology

Validation tests:

- High-energy physics, astroparticles, and gravitational wave physics utilizes extensive simulation, analysis, reconstruction, and interpretation codes developed over decades and comprising of millions of codes.
- Porting of scientific components from key parts of these codes (a) to measure the usability for scientific production use, and (b) to assess the economic and energy impacts of these platforms

Main scientific methodology:

- Solve the classic gravitational N-body problem using Hermite integration methods
  - Direct N-body software is crucial to accurately study the evolution of dense stellar systems (e.g., star clusters)
  - Developing efficient direct N-body codes is crucial for the astrophysical interpretation of gravitational wave sources.
  - Exploring new efficient hardware/software combinations will allow to evaluate the feasibility of investigating large and complex parameter space

Methodology

Development plans:

- The new N-body code will be optimized on RISC-V hardware made available on dedicated machines at E4 facility. The machines will be accessible locally and externally through VPN for monitoring and control.

Measurement and Monitoring:

- Software: IPMI systems will be used to collect hardware level data to monitor system's performance and energy efficiency
- Hardware: Monitoring with PDUs and serial measurement systems for which dedicated scripts will be developed for collection and graphic representation

Benchmarking:

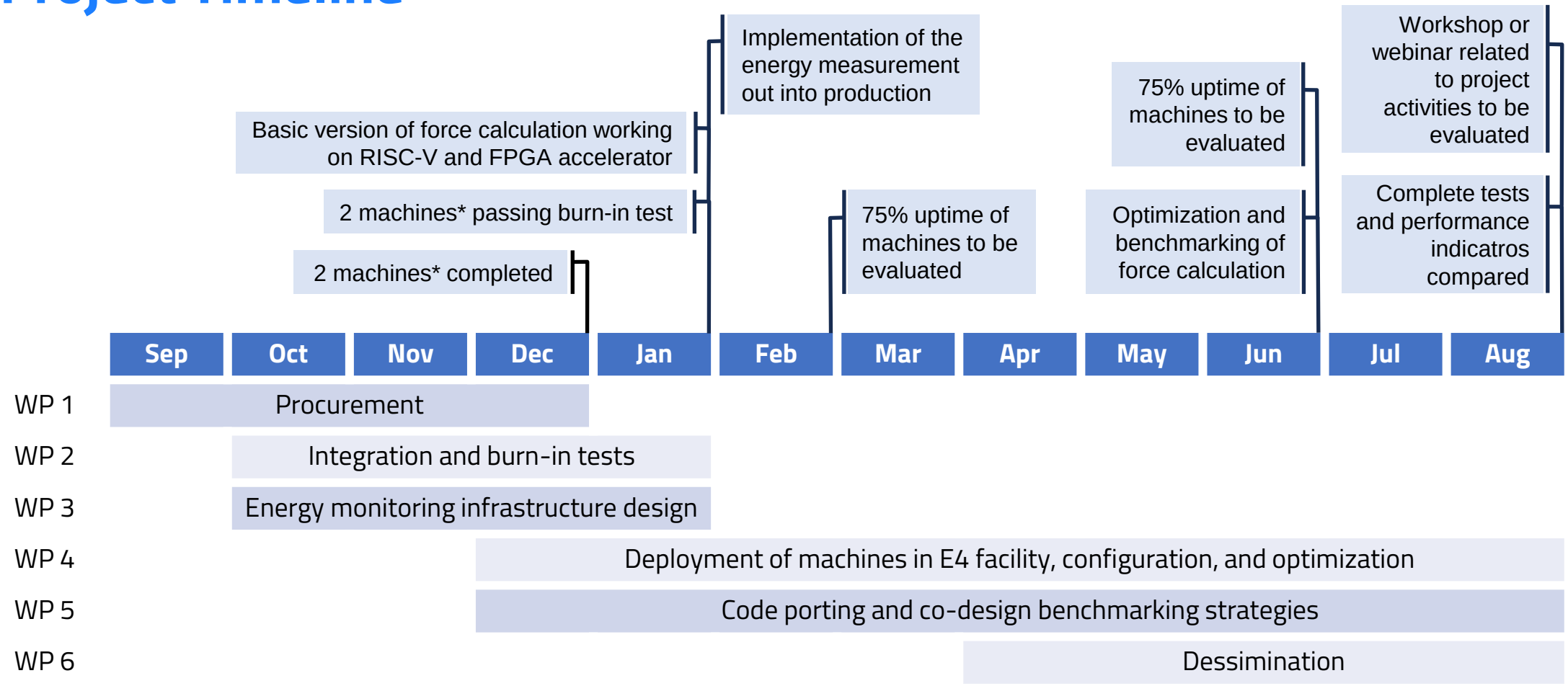
- Use synthetic HPC Benchmarks such as Linpack or Stream (for bandwidth measurement)
- Spoke 3 community specific software, with direct contact with the developer / user

Staff

For this activity:

- **Daniele Gregori** (Project Manager)
- **Elisabetta Boella** (Activity Supervision)
- **Jenny Lynn Almerol** (Research activity manager)
- **Mattia Paladino** (Dissemination Officer)

Project Timeline



*Machines with dual socket Sophgo 2042 processors including FPGA accelerators such as AMD Xilinx U55C and U280