Monte Cimone v2: Down the Road of RISC-V High-Performance Computers

Emanuele Venieri[1], Simone Manoni[1], Giacomo Madella[1], Federico Ficarelli[2], Daniele Gregori[3], Daniele Cesarini[2], Luca Benini[1,4], Andrea Bartolini[1]

[1]University of Bologna, Italy - [2]CINECA, Italy - [3]E4 Computer Engineering Spa - [4]ETH Zürich, Switzerland

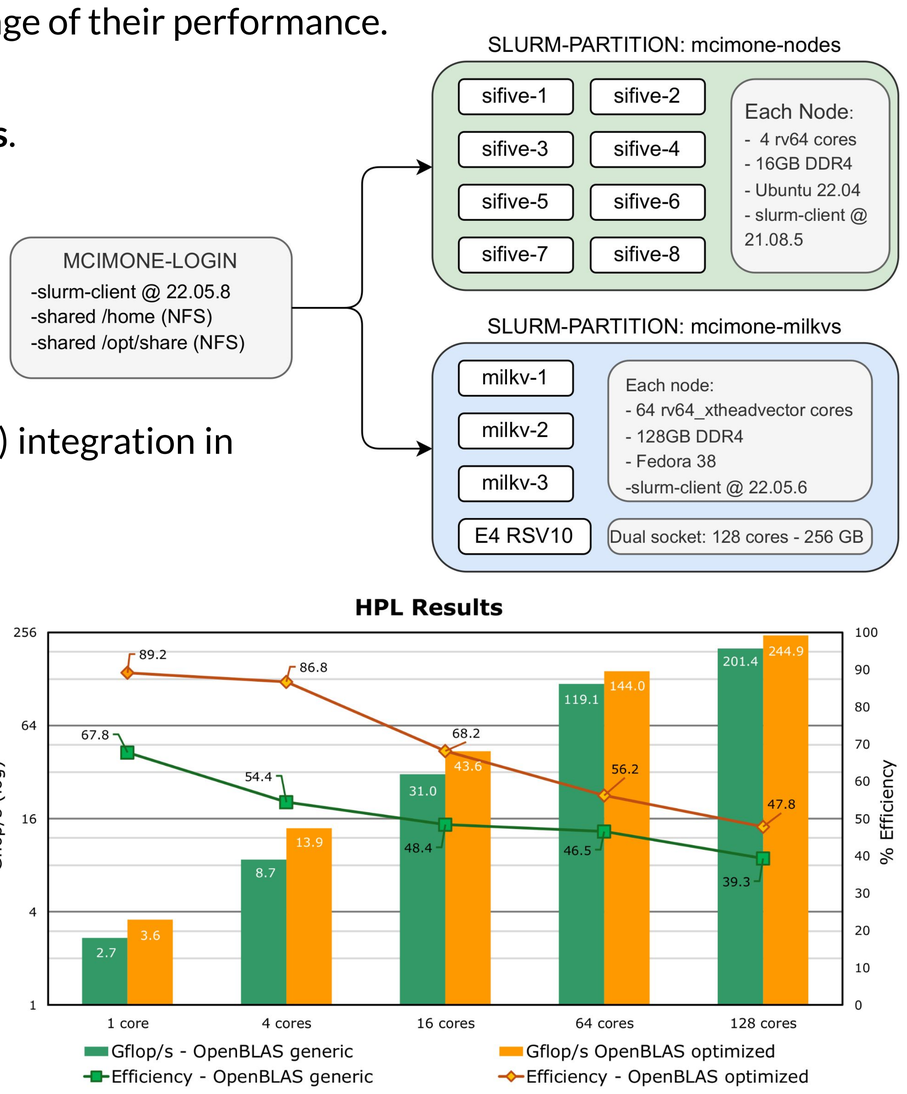
`emanuele.venieri2@unibo.it`

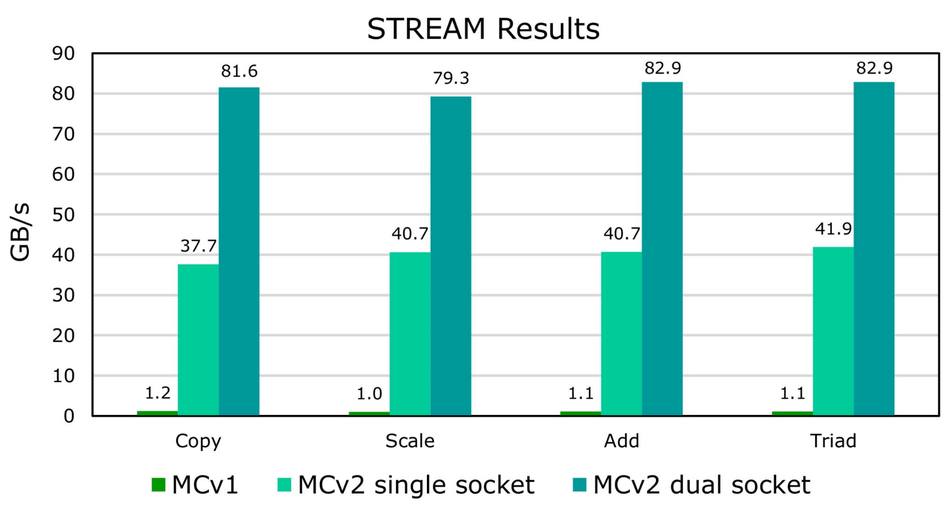
Motivation and Contributions

- Nowadays, RISC-V SoCs with vector extension, form factor and memory capacity suitable for HPC applications are available in the market, but it is unclear how compilers and open-source libraries can take advantage of their performance.
- The key contributions include:
  - Integrating new RISC-V-based hardware with **vector extensions**.
  - Setting up an **updated software** stack.
  - Conducting **performance benchmarking** using HPC-class tools.
  - Producing **new optimized kernels** for the SG2042 processor

Monte Cimone v2

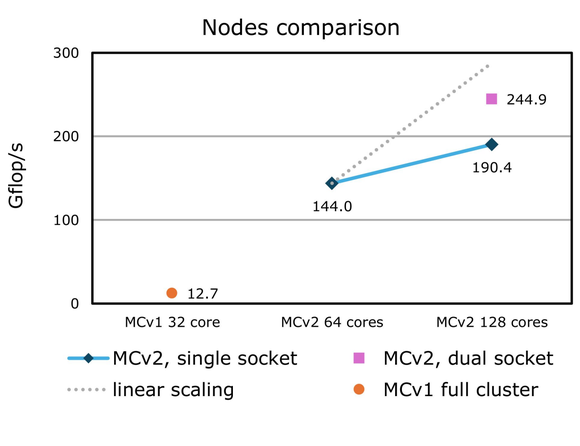
- **SG2042 based nodes** (three Milk-V Pioneer + one E4 dual socket) integration in Monte Cimone cluster in a dedicated **SLURM** partition
- **Examon** + graphical monitoring
- Dedicated and optimized libraries and toolchains via **Spack**

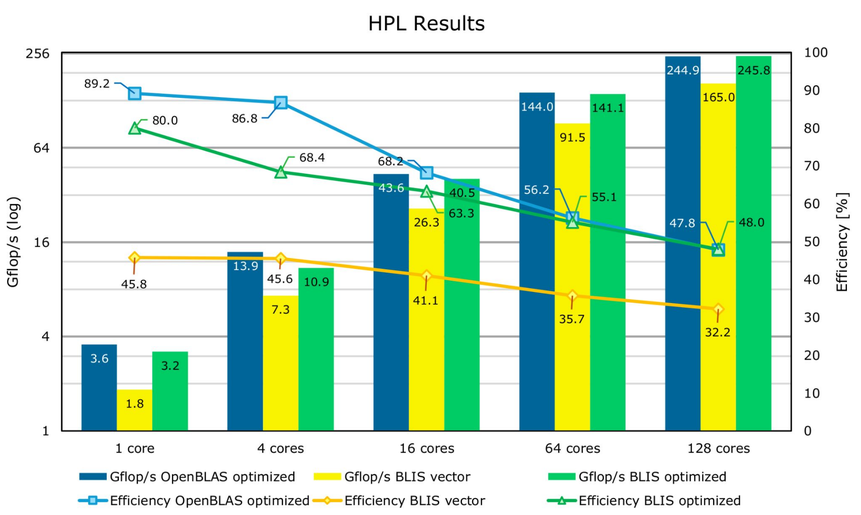




Methods

- **STREAM** benchmark for memory performance
- **HPL** benchmark for HPC FP64 performance
  - Vanilla + optimized BLIS
  - Vanilla + our optimization of BLIS





Results

- In MCv1, the 1Gb/s network was sufficient for obtaining almost an HPL linear scaling, in the case of the performance of the MCv2 nodes, it is no longer sufficient and increasing the number of parallel processes reduces the HPL efficiency (only the 1.33× w.r.t single node performance). The MCv2 dual-socket compute node, in contrast, achieves almost **1.76×**
- MCv2 dual-socket blade achieves **127×** more performance of an MCv1 compute node, these system are **two years** apart in release. For context, the Top500 list sees the same 127× performance improvement over an average of eight years.
- Our efforts in porting and optimizing BLAS libraries, specifically BLIS, demonstrate the feasibility of enhancing the RISC-V software ecosystem, ultimately benefiting the broader HPC community.

1] A. Bartolini et al. “Monte Cimone: Paving the Road for the First Generation of RISC-V High-Performance Computers”. In: IEEE 35th International System-on-Chip Conference. 2022.
[2] N. Brown et al. “Performance Characterisation of the 64-Core SG2042 RISC-V CPU for HPC”. In: ISC High Performance 2024 International Workshops.
[3] Top 500 statistics, top500.org/statistics/perfdevel/