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, Italy*
[4] *ETH Zürich, Switzerland*

Abstract

Many RISC-V platforms and SoCs have been announced in recent years targeting the HPC sector, but only a few of them are commercially available and engineered to fit the HPC requirements. The Monte Cimone project targeted assessing their capabilities and maturity, aiming to make RISC-V a competitive choice when building a datacenter. Nowadays, RV 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 its performance. In this paper, we describe the performance assessment of the upgrade of the Monte Cimone (MCv2) cluster with the Sophgo SG2042 processor's HPC operations. The upgrade increases the attained node's performance by 127× on HPL DP FLOP/s and 69× on Stream Memory Bandwidth.

Introduction

Monte Cimone is a multi-node compute cluster built on RISC-V architecture, designed as a validation platform for High-Performance Computing (HPC) systems. The first iteration of the Monte Cimone (MCv1) [1] compute cluster utilized four E4 RV007 Server Blades, based on two boards SiFive HiFive Unmatched featuring the SiFive Freedom U740 SoC. The peak theoretical performance was 4.0 Gflop/s per node.

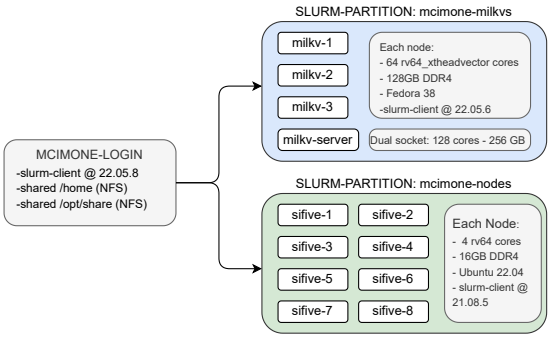

Figure 1: *Monte Cimone v1 (green) + v2 (blue) view*

The Monte Cimone software stack is built with Spack and accessible via shared modules. Nodes run Ubuntu 21.04, and mount a shared NFS, the job scheduler is SLURM and the system monitor is ExaMon. The system achieves 12.65 Gflop/s for full-machine HPL[1] and of 1.1 GB/s for STREAM DDR bandwidth[2] benchmarks. The primary focus was on assessing the maturity of RISC-V HPC system software, addressing challenges in integrating multi-node RISC-V clusters, including developing an HPC stack with interconnects, storage, and power monitoring, all built on RISC-V hardware. In this paper, we propose an upgrade of Monte Cimone named MCv2 to assess the maturity of software libraries and compilers in leveraging novel, more compute-capable RISC-V architectures featuring vector extensions and larger memory capacity. The key contributions include: (i) integrating new RISC-V-based hardware with vector extensions, (ii) setting up an updated software stack with compilation toolchains and optimized BLAS libraries, and (iii) conducting performance benchmarking using HPC-class tools.

Monte Cimone v2

The MCv2 configuration consists of four nodes, each based on the Sophgo Sophon SG2042 System-on-Chip (SoC)[3], the first RISC-V processor designed for server applications. The SG2042 features a 64-core RISC-V CPU based on the Xuantie C920 architecture. Each core includes a 128-bit wide vector unit supporting RVV 0.7.1 for vector execution. The SoC provides 64 KB of L1 instruction and data caches per core, a 1 MB L2 cache shared among four-core clusters, and a 64 MB system-wide L3 cache. It supports four channels of 3200MHz ECC DDR4 memory and provides 32 PCIe Gen4 lanes. Three of the nodes are Milk-V Pioneer Box systems, each equipped with an SG2042 processor and 128 GB of memory. The fourth node, provided by E4 Computer Engineering, is a dual-socket system with two SG2042 processors for a total of 128 cores and 256 GB of system memory based on the Sophgo SR1-2208A0 platform. The SG2042-based nodes are integrated into the Monte Cimone system (i) using the existing 1Gb/s network and (ii) as an additional SLURM partition. Similarly, the MCv2 nodes have been configured using Spack and integrated into the ExaMon monitoring infrastructure.

The MCv2 nodes run Fedora 38 as the operating system, along with the upstream GCC 13 toolchain. To enhance compatibility with the Xuantie C920 core and

[1] HPL, `www.netlib.org/benchmark/hpl/`.
[2] STREAM, `www.cs.virginia.edu/stream/`.
[3] SG2042 TRM, `github.com/milkv-pioneer/pioneer-files/blob/main/hardware/SG2042-TRM.pdf`.

its vector unit, we built and made available as shared modules two additional toolchains: the Xuantie GNU Toolchain[4], a customized GCC 10-based toolchain specifically designed by the Xuantie core developer for compiling code targeting the RVV 0.7.1 vector extension, and the GNU GCC 14 toolchain, which introduces support for the *theadvector* compilation target, as GNU GCC identifies the vector extension of the C920 core.

Experimental Results

To evaluate MCv2, we carried out a series of tests similar to those conducted on MCv1, focusing on memory performance and FP64 scalar and vector execution. These tests utilized the STREAM and HPL benchmarks. Stream and HPL were compiled with GCC 13 with the latter linked against two different sets of BLAS libraries. The first one is OpenBLAS[5] vanilla. The second one is an optimized OpenBLAS with assembly kernels compiled via Xuantie GNU toolchain.

STREAM The MCv2 single socket node saturates its memory bandwidth with 64 OpenMP threads, achieving a bandwidth of 41 GB/s. Interestingly, the MCv2 dual socket node achieves a memory bandwidth of 83 GB/s, still using 64 OpenMP threads but pinned symmetrically in the two sockets - increasing the number of OpenMP threads reduces the attained bandwidth. In contrast, an MCv1 node achieves a memory bandwidth of 1.1 GB/s with 4 OpenMP threads.

HPL Figure 2 reports the performance characterization for the HPL benchmark for the MCv2 compute node while scaling the number of cores and for different BLAS libraries. From the Figure, we can see that the vanilla OpenBLAS libraries are still lagging behind the SG2042-optimized ones, with a relative efficiency of 68% with one core, which increases to 89% of the optimized one. Both of them experience a degradation in relative performance when all the cores are used. This suggests that the optimized OpenBLAS suffers from SoCs bottlenecks as the unoptimized one.

These results confirm previous works comparison between the SiFive Freedom U740 SoC and the Sophgo Sophon SG2042[2], though obtained using a different software stack. Furthermore, we extend previous works to a multi-node cluster and dual-socket nodes, providing further insights into performance scaling across multiple MCv2 nodes.
Figure 3 shows the performance results from HPL runs with different node configurations. The MCv1 32-cores case refers to the HPL executed in parallel on all 8 MCv1 compute nodes, which achieves 13Gflop/s. The MCv2 64-cores case refers to the MCv2 single-socket HPL run, while the MCv2 128-cores case refers to

[4] Xuantie GNU Toolchain repository, `github.com/XUANTIE-RV/xuantie-gnu-toolchain`.
[5] OpenBLAS repository, `github.com/OpenMathLib/OpenBLAS`.

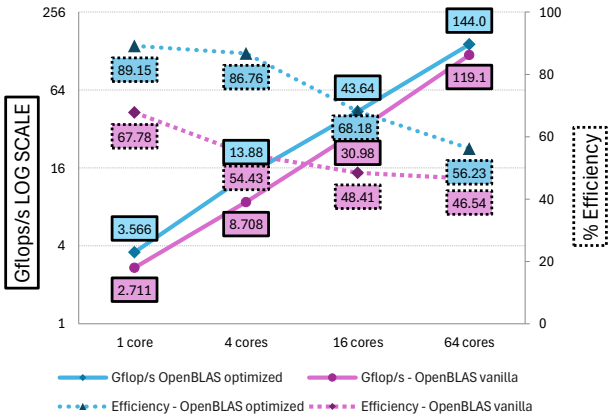

Figure 2: *MCv2 HPL w. OpenBLAS (opt. & vanilla)*

both (i) a dual MCv2 single socket nodes configuration and (ii) a single MCv2 dual socket node configuration. From it, we can notice that while 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× of the performance of the MCv2 single-socket node and 127× more performance of an MCv1 compute node.

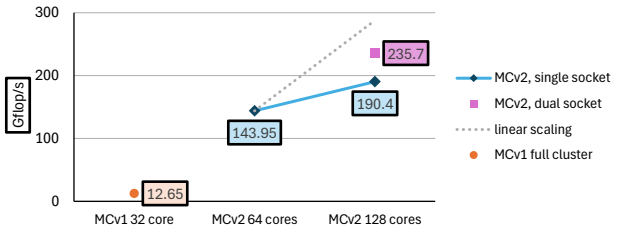

Figure 3: *HPL on different node's configurations*

Our comparative analysis of MCv1 and MCv2 shows that the RISC-V ecosystem has rapidly evolved in 2 years. For context, the Top500 list sees a 127× performance improvement over an average of eight years [6]. This acceleration suggests that RISC-V hardware is becoming as a viable platform for HPC systems.

Acknowledgements This activity has been supported by the HE EU Graph-Massivizer (g.a. 101093202), DECICE (g.a. 101092582), and DARE (g.a. 101143421) projects, as well as the Italian Research Center on High Performance Computing, Big Data, and Quantum Computing.

References

[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*.

[6] Top 500 statistics, `top500.org/statistics/perfdevel/`