The Arm architecture in HPC: from mobile phones to the Top500

Filippo Mantovani, Marta Garcia-Gasulla

Barcelona Supercomputing Center

The Arm architecture is gaining momentum in the HPC community as evidenced by several projects leveraging it including the Japanese Post-K, European Mont-Blanc, and the UKs GW4/EPSRC. For the first time during November 2018 the Astra supercomputer powered by Marvell's Cavium ThunderX2, assembled by HPE and installed at the Sandia National Laboratories (US) has been ranked 204 in the Top500 list.

Since more than six years several research projects in collaboration with industry evaluated Arm-based parallel systems for HPC and scientific computing advocating the higher efficiency of the technology mutated from the mobile and the embedded market. Several papers have been published with the preliminary analysis of benchmarks and performance projections of Arm SoC coming from mobile and embedded markets [1]. More recently also tests on Arm-based server SoC appeared in the literature [2].

In this presentation, we summarize the outcome of the third phase of the European project, Mont-Blanc, recently concluded, while evaluating the Dibona test platform. Dibona is the Arm-based parallel system developed and deployed within the Mont-Blanc 3 project and based on the same Marvell's CPUs housed in the Astra supercomputer.

The talk is structured with a bottom-up approach, presenting contributions with an increasing level of complexity. We will introduce the hardware features and the software configuration of the Dibona test platform. As the second step, we will present the results of a set of simple micro-benchmarks, exposing the basic architectural features such as the floating point throughput of the CPU, the structure of the memory subsystem and the bandwidth and the latency of the network interconnecting the computational nodes. Once clarified the architecture of the system, the results of the most relevant HPC benchmarks, LINPACK and HPCG, will be reported. Finally, since most of the scientific community is interested in the performance of production codes, we will dedicate the last part of the talk presenting the results obtained running on Dibona the Alya computational fluid- and particle-dynamics code, a real production scientific application combined with runtime optimizations [3].

The ultimate goal of the presentation is to review the performance and the efficiency of modern Arm-based systems that are de-facto part of the HPC market.

References

[1] Rajovic, Nikola, Alejandro Rico, Filippo Mantovani, Daniel Ruiz, Josep Oriol Vilarrubi, Constantino Gomez, Luna Backes et al. "The Mont-Blanc prototype: an alternative approach for HPC systems." In High Performance Computing, Networking, Storage and Analysis, SC16: International Conference for, pp. 444-455. IEEE, 2016.

[2] Calore, Enrico, Filippo Mantovani, and Daniel Ruiz. "Advanced performance analysis of HPC workloads on Cavium ThunderX." In 2018 International Conference on High Performance Computing & Simulation (HPCS), pp. 375-382. IEEE, 2018.

[3] Garcia-Gasulla, Marta, Marc Josep-Fabrego, Beatriz Eguzkitza, and Filippo Mantovani. "Computational Fluid and Particle Dynamics Simulations for Respiratory System: Runtime Optimization on an Arm Cluster." In Proceedings of the 47th International Conference on Parallel Processing Companion, p. 11. ACM, 2018.