MERIC: energy-efficient approach for your EuroHPC project

EuroHPC Summit Week 2022

Ondrej Vysocky, Matej Spetko, and Lubomir Riha
IT4Innovations national supercomputing center
VSB - Technical University of Ostrava
Ostrava, Czech Republic

Extended abstract

The power consumption of supercomputers is one of the critical problems for the upcoming exascale supercomputing era. Despite the power efficiency of the High-Performance Computing (HPC) hardware consistently rising, the efficiency of the greenest HPC centers is around 30 GFlops/Watt accoring top500 list. Just ten years ago the community accepted that the future exascale system should not consume more than 20 MW, which means power efficiency of 50 GFops/Watt. Anyway, first exascale computers are going to start execution soon, exceeding the overall power consumption, which causes high demands to power resources, as well as transmission network, producing a large amount of CO_2 emissions.

The energy efficiency can be further improved only if there is power awareness at both the hardware and software side. The scientific applications are developed to optimize its performance however it is hard to fully utilize any given hardware, that is available in modern HPC centers. The under-utilized hardware gives a possibility to exploit energy savings with none or limited performance degradation.

We present our user-friendly runtime system MERIC for dynamic hardware parameters tuning during your parallel application run with the goal of minimal energy consumption. The MERIC development started under the Horizon 2020 READEX project, which comes with an approach to exploit the dynamic behavior of HPC applications, which means that different parts of an application utilize different resources. This approach's goal is in the detection of the optimal configuration for each part of the application and dynamic tuning during the production runs.

The MERIC can control CPU core and uncore frequencies, concurrency throttling, GPU streaming multiprocessor, and memory frequency as well as CPU and GPU power capping. Since GPUs produced by Nvidia only are supported, the CPUs of Intel, AMD, IBM, and ARM vendors can be used to execute MERIC analysis and tuning. The multi-platform support is limited to architectures, that provide a power monitoring system, while Runnnig Average Power Limit (RAPL), High Definition Energy Efficiency Monitoring (HDEEM), and Dwarf in Giant (DiG) are currently supported by MERIC.

Alongside the MERIC we also develop RADAR visualizer, which visualizes the application resources consumption in various configurations, its dynamic behavior that can be exploited, and reachable energy savings with corresponding performance trade-offs. In comparison to other parallel application performance visualization tools, it is designed to use data from several application runs, potentially in a different parameters configuration, and compare these measurements.

Using the MERIC can bring major energy savings over 30 %, moreover in some cases improve scalability of the tuned application. The MERIC and the RADAR visualizer have been used for application performance analysis and tuning in several projects including Horizon 2020 projects READEX, SCALABLE, POP2, and EUPEX. From these, we will show an example of application analysis and the impact of the MERIC dynamic tuning.