Biomolecular Simulations on the LUMI Supercomputer

Rasmus Kronberg [1]

[1]CSC – IT Center for Science Ltd., Finland

ICT Solutions for Brilliant Minds

LUMI is an HPE Cray EX Supercomputer

LUMI is a pan-European pre-exascale supercomputer co-funded by the EuroHPC Joint Undertaking and a consortium of ten European countries. LUMI is located in Kajaani, Finland, and operated by CSC – IT Center for Science, the national competence center for high-performance computing in Finland. LUMI is currently the fastest supercomputer in Europe and the fifth fastest globally [1].

Figure 1. Cabinets of the LUMI supercomputer.

Modern Architecture

LUMI is composed of eight hardware partitions targeting various use cases. All compute nodes are connected by a 200 Gb/s HPE Slingshot 11 high-speed interconnect.

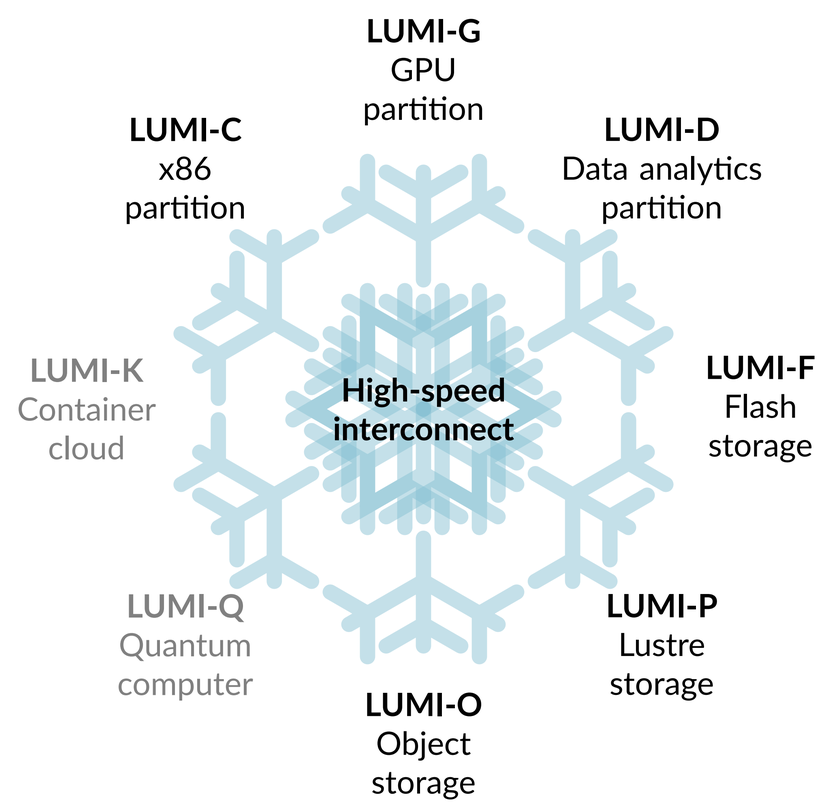

Figure 2. Hardware partitions of the LUMI supercomputer. Greyed out partitions are not yet available.

System Specifications

The measured LINPACK performance of LUMI is 0.38 Eflop/s [1]. The primary computing power of the system comes from its GPU partition, LUMI-G, featuring AMD Instinct MI250X GPUs. LUMI-G is augmented by a smaller CPU partition (LUMI-C) with 64-core AMD EPYC "Milan" CPUs as well as a data analytics and visualization partition (LUMI-D) featuring large memory nodes with fast local disks and Nvidia A40 GPUs. The total amount of memory and storage space available are 2 PiB and 118 PiB, respectively.

Partition	Nodes	GPUs per node	CPUs per node	Memory per node	Storage
LUMI-G	2978	4 AMD MI250X	1 AMD EPYC	512 GiB	
LUMI-C	2048	–	2 AMD EPYC	256–1024 GiB	
LUMI-D	16	8 Nvidia A40	2 AMD EPYC	2048–4096 GiB	312 TiB
LUMI-P	–	–	–	–	80 PiB
LUMI-F	–	–	–	–	8 PiB
LUMI-O	–	–	–	–	30 PiB

Table 1. Specifications of the hardware partitions of LUMI.

Benchmarking GROMACS on LUMI-G

GROMACS [2] is a free and open-source software suite for high-performance molecular dynamics (MD). While GROMACS has had excellent support for Nvidia GPUs for a long time, support for AMD GPUs has only recently matured following developments in the AdaptiveCPP SYCL implementation [3] that GROMACS uses to enable GPU offloading to AMD hardware.

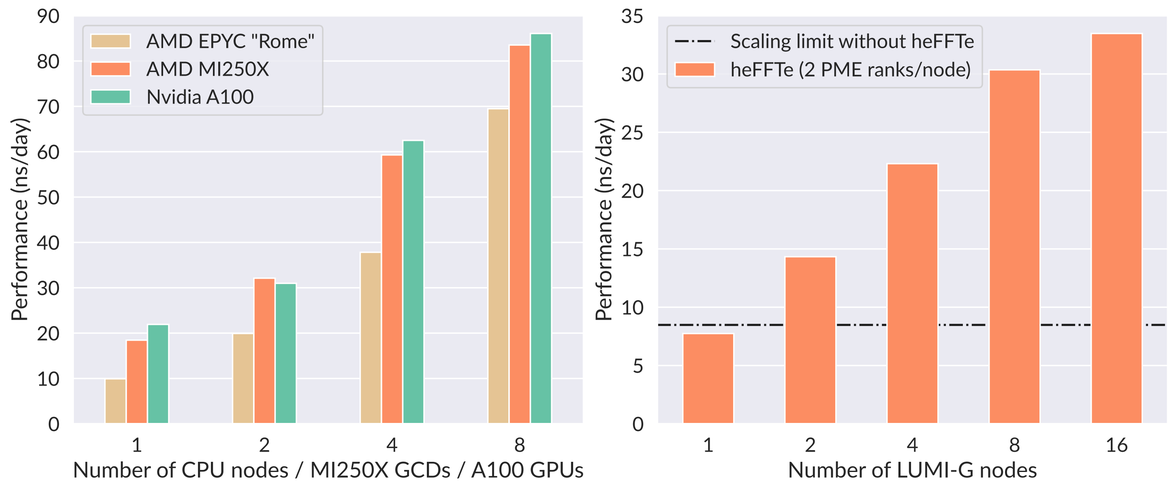

Figure 3. Scalability of GROMACS 2024.0. **Left:** Benchmarked system is a solvated satellite tobacco mosaic virus (STMV, 1067k atoms). Note that each AMD MI250X GPU is composed of two distinct graphics compute dies (GCD). **Right:** Effect of GPU PME decomposition enabled by heFFTe on the scalability of large systems, in this case a 12 million atom box of peptides in water (benchPEP-h) [4].

Benchmarking GROMACS 2024.0 on LUMI-G shows that large systems (few 100k–1M atoms) are typically able to utilize multiple AMD GPUs efficiently. The results for the STMV benchmark (Fig. 3, left) illustrates that a single MI250X GCD (half a GPU) outperforms a 128-core AMD EPYC "Rome" CPU node while being roughly as efficient as a full Nvidia A100 GPU.

The scaling of systems composed of several million atoms may be limited by single GPU PME. This bottleneck can be avoided by using the heFFTe library [5] to enable PME decomposition on AMD GPUs. The effect of PME decomposition on the scalability of the 12M atom benchPEP-h benchmark on LUMI-G is shown in Fig. 3.

Speed *vs.* Throughput

The AMD MI250X GPUs have native support for running multiple concurrent processes on a single GCD. This allows increasing the GPU utilization of small systems by sharing GCDs among several independent MD trajectories launched using *e.g.* the `-multidir` feature of GROMACS. For example, for a test system of 96k atoms, sharing one GCD among four trajectories increases the aggregate performance on two LUMI-G nodes by ∼1 µs/day compared to running just one simulation per GCD.

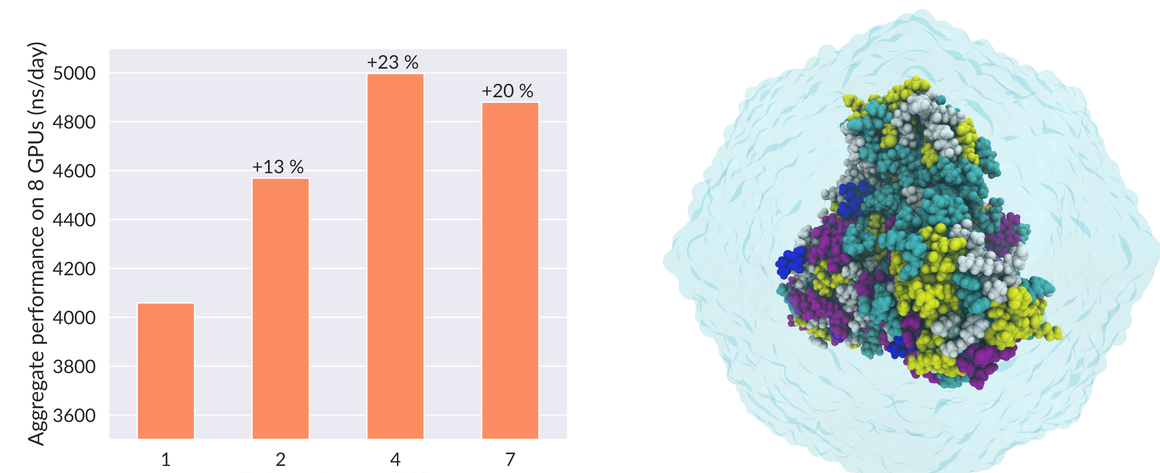

Figure 4. **Left:** Aggregate performance of GROMACS 2024.0 multi-simulations on LUMI-G. Benchmarked system is a solvated alcohol dehydrogenase enzyme (ADH, 96k atoms). The aggregate performance is calculated as the sum of the performance of each independent trajectory. **Right**: Visualization of the ADH system.

References

[1] 62nd TOP500 list. https://www.top500.org/lists/top500/2023/11/. Accessed: 2024-02-10.

[2] GROMACS. https://www.gromacs.org/. Accessed: 2024-02-10.

[3] AdaptiveCPP. https://adaptivecpp.github.io/. Accessed: 2024-02-10.

[4] A free GROMACS benchmark set. https://www.mpinat.mpg.de/grubmueller/bench. Accessed: 2024-02-10.

[5] Highly Efficient FFT for Exascale. https://icl-utk-edu.github.io/heffte. Accessed: 2024-02-10.

The acquisition and operation of the EuroHPC supercomputer is funded jointly by the EuroHPC Joint Undertaking, through the European Union's Connecting Europe Facility and the Horizon 2020 research and innovation programme, as well as the Participating States FI, BE, CH, CZ, DK, EE, IS, NO, PL, SE.