MAX PLANCK INSTITUTE
FOR MULTIDISCIPLINARY SCIENCES

Contact:
ckutzne@mpinat.mpg.de

Scaling of the GROMACS Molecular Dynamics Code to 65k CPU Cores on an HPC Cluster

C Kutzner,[1] V Miletić,[2] K Palacio Rodríguez,[3] M Rampp,[2] F Leidner,[1] G Hummer,[3] BL de Groot,[1] and H Grubmüller[1]

[1]MPI for Multidisciplinary Sciences, Göttingen [2]Max Planck Computing and Data Facility, Garching [3]MPI of Biophysics, Frankfurt/Main

Abstract

We benchmarked the performance of the GROMACS 2024 molecular dynamics (MD) code[1] on a modern high performance computing (HPC) cluster with AMD CPUs on up to 65,536 cores. We used 5 different MD systems, ranging in size from about 82 thousand to 204 million atoms, and evaluated their performance using two different MPI libraries, Intel-MPI and Open-MPI.

We achieved peak performances of 687 ns/d for the 82k atom system, 116 ns/d for the 53M atom system, and about 35 ns/d for the largest 204M atom system.
The largest system showed near-perfect scaling up to 65,536 cores, maintaining a parallel efficiency above 0.9 even at the highest level of parallelization. Energy efficiency for a given number of nodes was generally equal to or slightly better than parallel efficiency.

These results demonstrate that highly optimized software running on a state-of-the-art HPC cluster provides sufficient computing power to simulate biomolecular systems at the mesoscale of viruses and organelles, and potentially small cells in the near future.

Introduction

While most protein simulations involve between 10,000 and several hundred thousand atoms, recent advances in computer hardware and software enable the study of increasingly larger systems,[2,3] with the current world record exceeding one billion atoms.[4] These developments suggest that it may soon be possible to simulate an entire biological cell, which would comprise 500 million to six billion particles.[5]
When preparing for large-scale simulations on expensive high-end computing resources, it is generally advisable to compare the achieved performance with reference benchmarks. Reference performance numbers also help to estimate the achievable simulation throughput and aid in the planning of new simulation campaigns.

Methods

The Viper HPC Cluster

- AMD-based CPU cluster[6] operated at the Max Planck Computing and Data Facility (MPCDF) since July 2024
- 768 nodes with two EPYC 9554 “Genoa” CPUs @ 3.1 GHz base frequency, with a total of 128 physical cores per node, and 512 GB RAM
- NVIDIA/Mellanox NDR200 Infiniband network (200 Gb/s) with a non-blocking fat-tree topology

Software

- On Viper, we built GROMACS 2024.2 in mixed precision with AVX_512 SIMD instructions and OpenMP support, using GCC 14.1.0, and linking against FFTW 3.3.10 for the FFTs required by PME
- We built two versions, one with Intel-MPI 2021.11 and one with Open-MPI 4.1.6
- The MEM benchark result on a Viper GPU node used Paul Bauer’s HIP branch of GROMACS 2022 (https://github.com/ROCm/Gromacs, which is planned to be incorporated in the 2026 release of GROMACS), and VkFFT[9]
- On JEDI, GROMACS 2024.3 was used with OpenMPI 4.1.6 and CUDA 12. NVIDIA’s cuFFTMp library was used to distribute the 3D FFTs across multiple GPUs

Benchmark Setup

- We relied on the well-performing empirical defaults for parameters GROMACS can control, such as the number of separate PME ranks and whether or not to enable dynamic load balancing
- We always used all 128 physical cores of each node exactly, in the following combinations of MPI ranks x OpenMP threads: 1 x 128, 2 x 64, 4 x 32, and 8 x 16. We report the best-performing one
- Benchmarks ran for 0.2–0.5 hours of wallclock time. We excluded the first half of each run from the performance measurements, as load balancing mechanisms take some time to reach their optimum
- Power consumption was measured using Eviden’s “Smart Energy Manager Suite” (SEMS)

GROMACS Scaling on Viper

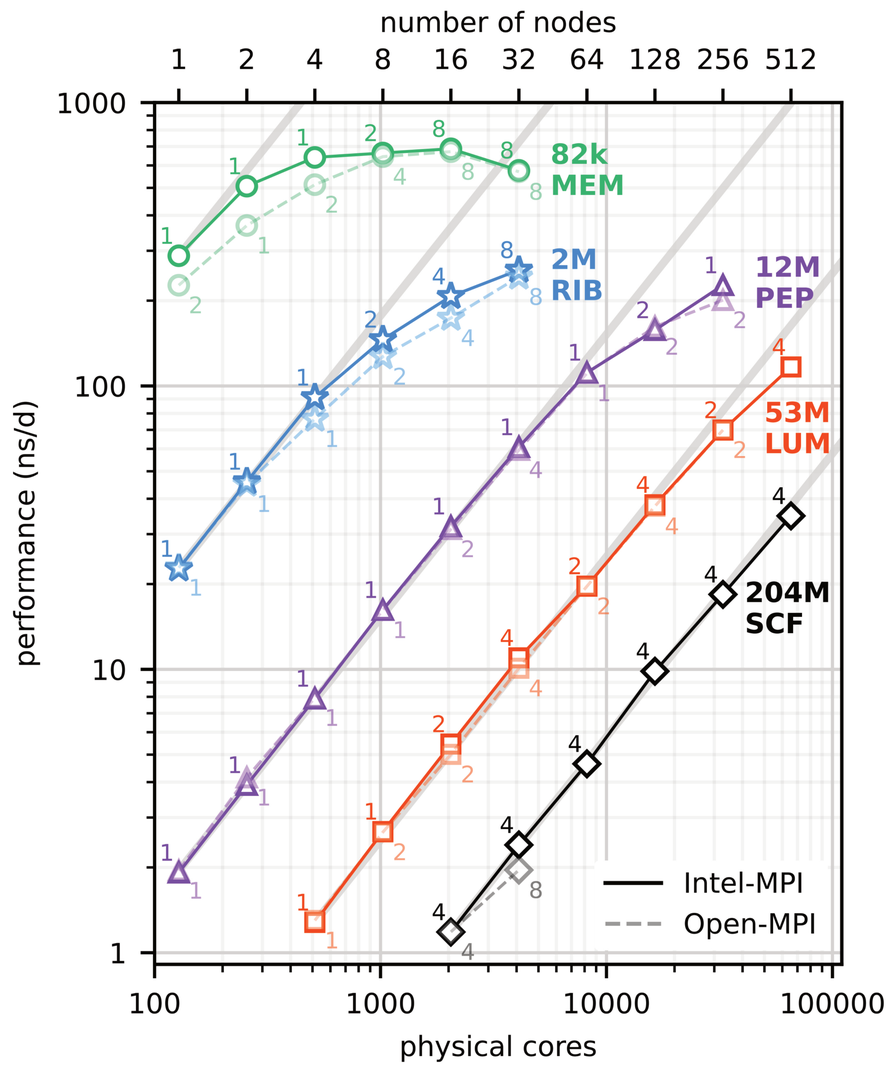

Figure 1. Strong scaling of GROMACS on up to 65,536 CPU cores.

Performances of GROMACS 2024 on up to 512 Viper nodes with Intel-MPI (solid) and Open-MPI (dashed, semi-transparent) for five MD systems (colors) ranging from 82k to 204M atoms in size. Numbers beside symbols refer to the optimal number of OpenMP threads per MPI rank. Diagonal gray lines indicate perfect scaling.

Benchmark Systems

https://www.mpinat.mpg.de/grubmueller/bench

MEM 82 k atoms, Aquaporin tetramer in lipid membrane + ions + water, Δt = 2 fs

RIB 2.1 M atoms, solvated Ribosome, Δt = 4 fs

PEP 12.5 M atoms, Peptides in water, Δt = 2 fs

LUM 53 M atoms, luminal ring of human nuclear pore complex (hNPC), Δt = 2 fs

SCF 204 M atoms, hNPC scaffold (inner ring, nuclear ring, cytoplasmic ring, luminal ring), Δt = 2 fs

all use PME

Comparison to other HPC systems

To put Viper performance in context, we compare with three other HPC platforms (Fig. 3):

1. **AWS hpc6a:** Cluster of hpc6a.48xlarge instances in the Amazon Web Services (AWS) cloud based on AWS ParallelCluster.[7] Compute nodes with two AMD EPYC “Genoa” CPUs @ 2.95GHz base frequency, with a total of 96 physical cores per node, and 384 GB RAM, elastic fabric adapter network (100 Gb/s)
2. **JEDI:** Cluster of GH200 nodes with four NVIDIA GH200 Grace-Hopper superchips each, operated at JSC Jülich. Four ARM Neoverse-V2 CPUs @ 3.1 GHz base frequency with 120 GB RAM each plus four H100 GPUs with 96 GB HBM3 memory each. Nodes have 288 physical cores in total, 4x NDR200 Infiniband network (4x 200 Gb/s)
3. **Viper/GPU:** In contrast to the Viper CPU nodes with 128 cores, a Viper GPU nodes has 48 AMD 9554 “Genoa” cores @ 3.1 GHz base frequency, plus two AMD Instinct MI300A APUs with 128 GB HBM3 memory each, NVIDIA/Mellanox NDR200 Infiniband network (400 Gb/s) with a non-blocking fat-tree topology

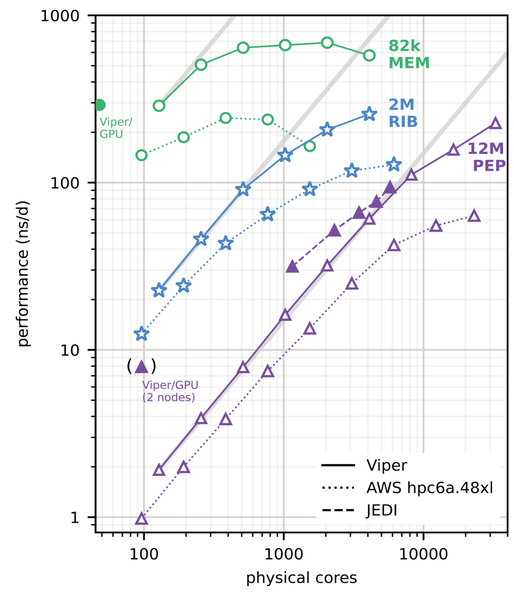

Figure 3. How does Viper-CPU compare to other modern HPC systems?

GROMACS performance for 3 benchmark systems (colors) on the Viper CPU nodes (solid lines), on a Cloud-based cluster of AWS hpc6a instances (dotted), and on JEDI (dashed). Open (filled) symbols denote CPU (GPU) nodes.

Computational Efficiency

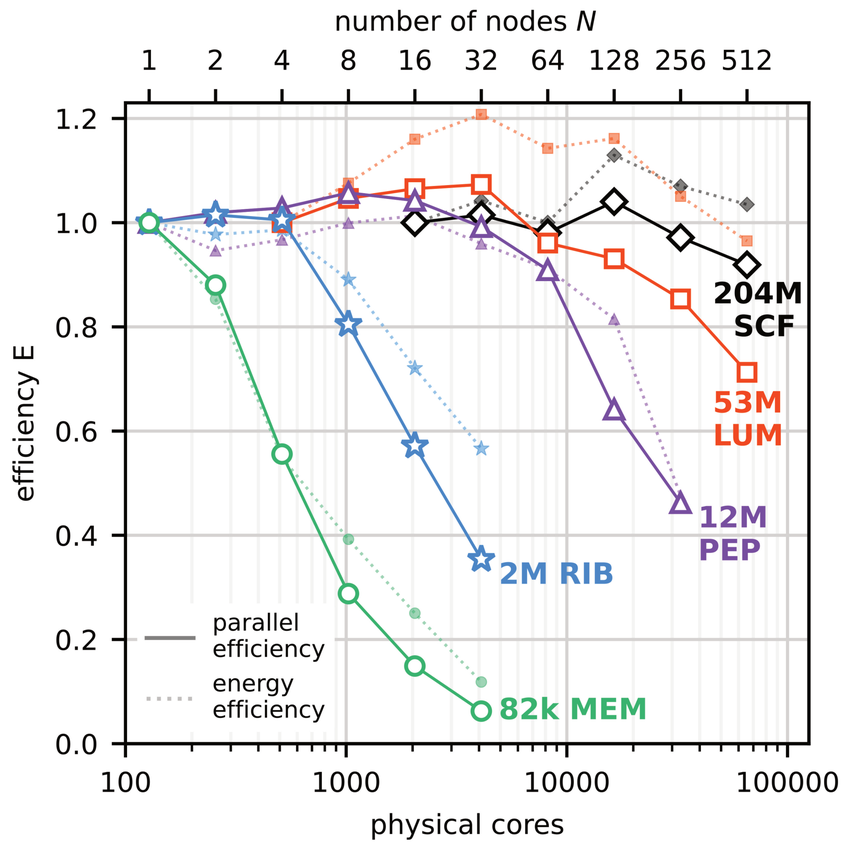

Figure 2. Parallel efficiency of GROMACS on up to 65,536 CPU cores.

For the five benchmarked systems (colored), two efficiency metrics are shown for the Intel-MPI benchmarks: (a) the parallel efficiency E_{par} (solid lines), and (b) the normalized energy efficiency E_{en} (dotted lines).

Efficiency estimates

Parallel efficiency E_{par} measures how efficient parallel computation is relative to the serial case,

$$E_{\mathrm{par}} = \frac{P(N)}{N \cdot P(1)}$$

Energy efficiency E_{en}, the performance on N nodes divided by the average power consumption pow(N) of the involved nodes, measures how energy efficient parallel computation is compared to the serial case,

$$E_{\mathrm{en}} = \frac{P(N)}{\mathrm{pow}(N)}$$

Summary

- We achieved **peak performances** (Fig. 1) of
  687 ns/d for the 82k atom system (MEM),
  256 ns/d for the 2M atom system (RIB), and
  226 ns/d for the 12M atom system (PEP).[9]
  These are the highest performances ever reported for these specific benchmark systems, including on GPU-accelerated nodes
- At the highest parallelization of 65,536 cores, the two largest systems still show high **parallel efficiencies** (Fig. 2) of 0.71 (LUM) and 0.92 (SCF), yielding remarkable simulation speeds of
  116 ns/d for the 53M atom system (LUM), and
  35 ns/d for the 204M atom system (SCF)
- Normalized **energy efficiency** generally follows the same trend as parallel efficiency (Fig. 2). At higher node counts, where simulation performance deviates from ideal scaling, nodes become underutilized and energy efficiency becomes higher than parallel efficiency
- In the regime where ideal scaling is observed in Fig 1, power draw is about 1.2 kW per node
- Preliminary tests on Viper GPU nodes (Fig. 3, filled purple triangles) indicate that the used MI300A APUs are on par with NVIDIAS H100’s

Conclusions

Our GROMACS benchmarks of realistic biomolecular systems on the new AMD-CPU-based HPC system of the Max Planck Society, Viper, show that the combination of powerful CPUs, a fast interconnect, and well-tuned software makes it possible today to simulate extremely large MD systems with up to several hundred million atoms.[9]

For the largest system with 204 million atoms, the scaling behavior was close to ideal, with parallel efficiencies above 0.9 even at 65,536 cores, while maintaining high energy efficiency. This demonstrates that modern HPC machines now provide the level of performance needed for simulating MD systems of the size of viruses, organelles, and eventually, small cells.

Compared to other state-of-the-art HPC systems, including those with GPUs, Viper/CPU achieves excellent simulation performance.

Acknowledgments

Many thanks to Harald Braun, John Donners (Eviden), Rafael Lago (Intel), Renate Dohmen, and Mykola Petrov (MPCDF) for their assistance. This work was supported by the German Federal Ministry of Education and Research (BMBF) as part of its initiative “SCALEXA – New Methods and Technologies for Exascale Computing” (BMBF project 16ME0713), and by BioExcel CoE (www.bioexcel.eu), a project funded by the European Union contract HORIZON-EUROHPC-JU-2021-COE-01-02. K.P.R. acknowledges support from the “Hessen Horizon Marie Skłodowska-Curie-Stipendium” program. All authors acknowledge support from the Max Planck Society.

References

1. S Páll, MJ Abraham, C Kutzner, B Hess, E Lindahl: *Tackling exascale software challenges in molecular dynamics simulations with GROMACS.* EASC 2014, Stockholm, 3–27 (Eds. Markidis, S.; Laure, E.) Springer, Cham (2015)
2. E Wilson, J Vant, J Layton, R Boyd, H Lee, M Turilli, B Hernández, S Wilkinson, S Jha, C Gupta, D Sarkar, A Singharoy: *Large-Scale Molecular Dynamics Simulations of Cellular Compartments.* In Structure and Function of Membrane Proteins; Schmidt-Krey, I., Gumbart, J. C., Eds.; Springer Cham (2021) pp 335–356
3. V Gapsys, W Kopec, D Matthes, BL de Groot. *Biomolecular simulations at the exascale: From drug design to organelles and beyond.* Curr. Opin. Struct. Biol. (2024) 88, 102887
4. J Jung, C Kobayashi, K Kasahara, C Tan, A Kuroda, K Minami, S Ishiduki, T Nishiki, H Inoue, Y Ishikawa, M Feig, Y Sugita. *New parallel computing algorithm of molecular dynamics for extremely huge scale biological systems.* J. Comput. Chem. (2021) 42, 231–241.
5. JA Stevens, F Grünewald, PM van Tilburg, M König, BR Gilbert, TA Brier, ZR Thornburg, Z Luthey-Schulten, SJ Marrink. *Molecular dynamics simulation of an entire cell.* Front. Chem. (2023) 11, 1106495
6. Max Planck Computing and Data Facility, *Viper user guide.* https://docs.mpcdf. mpg.de/doc/computing/viper-user-guide.html (2024)
7. C Kutzner, C Kniep, A Cherian, L Nordstrom, H Grubmüller, BL de Groot, V Gapsys: *GROMACS in the Cloud: A Global Supercomputer to Speed Up Alchemical Drug Design,* J. Chem. Inf. Model. 62 (7), 1691–1711 (2022)
8. C Kutzner, V Miletić, K Palacio Rodríguez, M Rampp, G Hummer, BL de Groot, H Grubmüller. *Scaling of the GROMACS Molecular Dynamics Code to 65k CPU Cores on an HPC Cluster.* J. Comp. Chem. (2025) 46: e70059
9. D. Tolmachev, *VkFFT-A Performant, Cross-Platform and Open-Source GPU FFT Library,* in IEEE Access (2023) vol. 11, 12039–12058