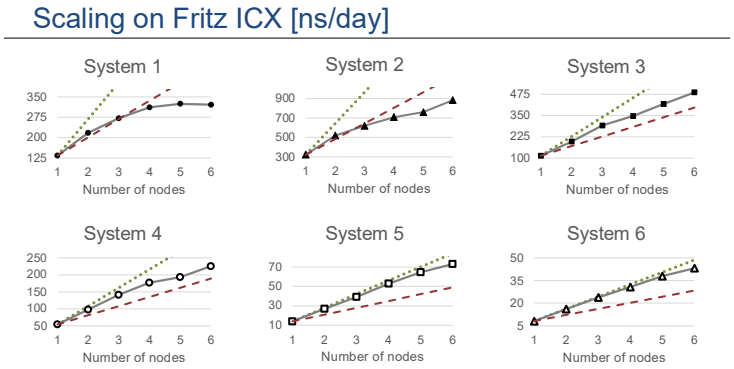
Improving MD performance on HPC clusters through in-depth hardware knowledge and advanced program usage

Anna Kahler, Thomas Zeiser, Gerhard Wellein
NHR@FAU, Friedrich-Alexander-Universität Erlangen-Nürnberg

Most modern MD simulation programs run out of the box on HPC clusters and yield reasonable performance results. To shed some light on the backgrounds of optimized performance, we present three case studies from user support.



First, an REMD simulation with GROMACS reached 124 ns/day for 26 replicas on 12 dual-socket Intel Ice Lake. We were able to port this simulation to eight NVIDIA A40 GPUs while retaining a performance of 120.6 ns/day, yielding a reduction in hardware costs by 2/3. Since the number of replicas is not a multiple of eight, porting required assignment of PP- and PME-tasks to the GPUs by hand.

The second case is about calling the GROMACS runtime correctly to obtain a performance gain, especially when dealing with a large simulation system of 2,600,000 atoms and a multiple GPU-setup. Starting from a performance of 11.8 ns/day on eight NVIDIA A40 GPUs, we nearly quadrupled performance to 20 ns/day on four A40 GPUs. Thus, about twice the performance on half of the resources by setting environment variables for improved GPU communication and adjusting runtime parameters.

Our third proof of in-depth hardware knowledge is represented by our third case where ORCA underperformed on our high throughput cluster: A single numerical calculation of molecular frequencies took 76.4 hours to finish. Multiple setups on various CPU architectures followed by detailed examinations sped up this simulation to 11 hours on the same node; the statically linked OpenBLAS library falsely detected the underlying hardware.