SCALABLE LATTICE BOLTZMANN LEAPS TO EXASCALE

Jayesh Badwaik (Jülich Supercompting Center) <j.badwaik@fz-juelich.de>
Ondřej Vysocký (IT4Innovations) <ondrej.vysocky@vsb.cz>
SCALABLE Consortium

Introduction

Lattice Boltzmann methods (LBM) are trustworthy alternatives to conventional CFD, showing roughly an order of magnitude performance advantage than Navier-Stokes approaches in comparable scenarios. The SCALABLE project brings together industrial and academic partners to create a new framework for an industrial LBM-based computational fluid dynamics (CFD) solver which can achieve high performance, scalability, and energy efficiency.

In the context of EuroHPC, LBM is especially well suited to exploit advanced supercomputer architectures through vectorization, accelerators, and massive parallelization. To achieve its goals, SCALABLE aims to transfer leading edge performance technology between waLBerla and LaBS, thus breaking the silos between the worlds of scientific computing and physical flow modelling.

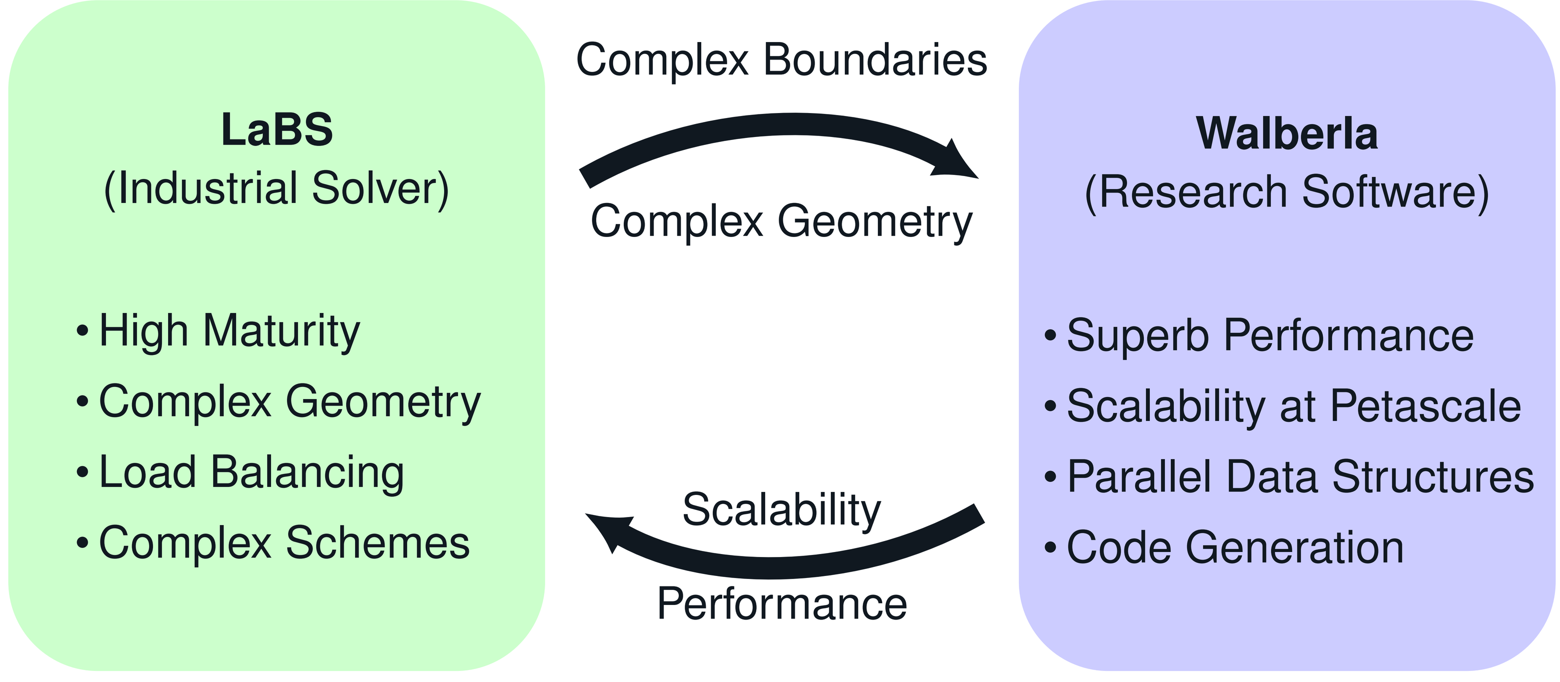

Fig. 1: Exchange of Technologies Between LaBS And waLBerla

Test Cases

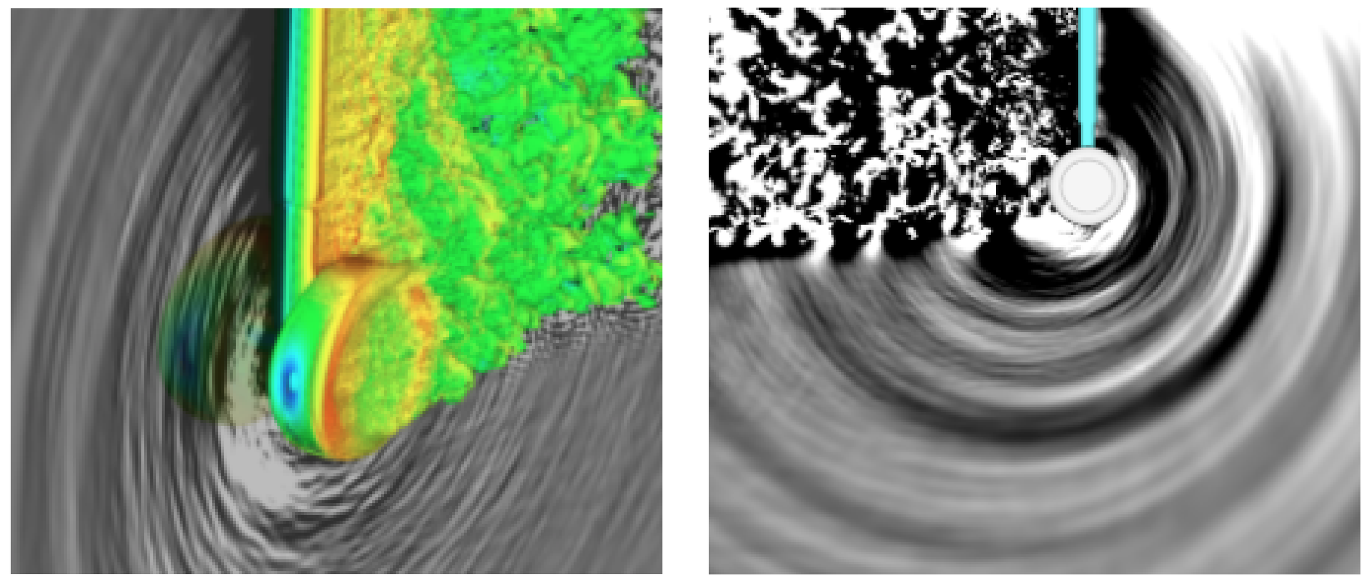

Fig. 2: Lagoon Test Case

SCALABLE has chosen a suite of test cases reflecting the interest of the industrial partners. One of the test cases in the suite is the Lagoon landing gear.

During the landing phase of an aircraft, the landing gear is an important source of noise and produces a non-trivial turbulent flow. The simple geometry and an existing large experimental database make the use case an attractive candidate to test and benchmark the performance of the new framework.

Energy Efficiency Optimization

Energy consumption of the non-accelerated execution can be reduced by exploiting the dynamic behavior of the different phases of applications. In SCALABLE, the MERIC run time system, developed as a part of H2020 READEX project, is used to scale CPU core and uncore frequencies to fit the executed phase of the applications.

Figure 3 and Table 4 shows the effect of employing MERIC on LaBS, executing the Lagoon test case. Major energy savings can be obtained over the full blade with negligible performance loss.

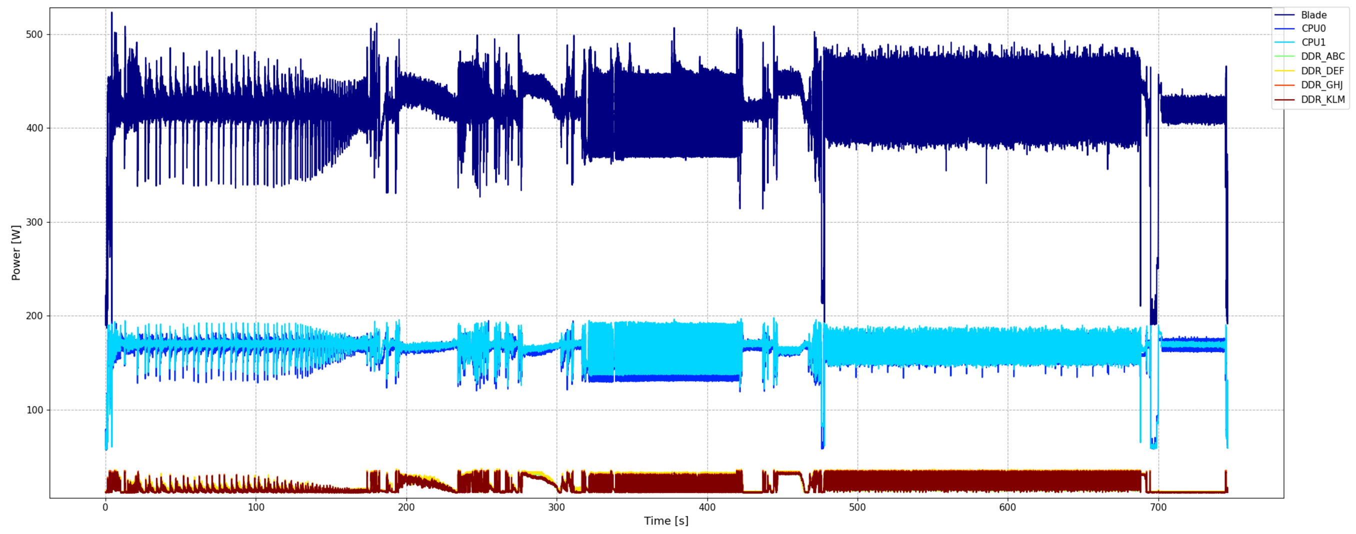

Fig. 3: Effect of Dynamic Tuning on LaBS for Lagoon Test Case. Used hardware: Intel Xeon Cascade Lake 6420

	Default	Static Tuning	Dynamic Tuning Constant Runtime	Dynamic Tuning Energy Optimized
Runtime / s	1797.9	1942.73	1807.13	1871.14
Energy / kJ	3102.3	1942.73	2726.7	2496.71
Efficiency / MLups/W	0.054	0.059	0.056	0.056
Runtime Extension / %	-	8.1	0.5	4.1
Energy Saving / %	-	15.1	12.1	19.5

Fig. 4: Numerical Data for Figure 3

waLBerla has been analyzed in both accelerated and non-accelerated configurations. Figure 5 shows waLBerla running an accelerated version of the Lagoon test case. The use case was optimized by specifying a static GPU streaming multiprocessor (SM) frequency. Scaling of the CPU core frequency does not bring any additional savings in this case since the GPU is the main consumer of energy.

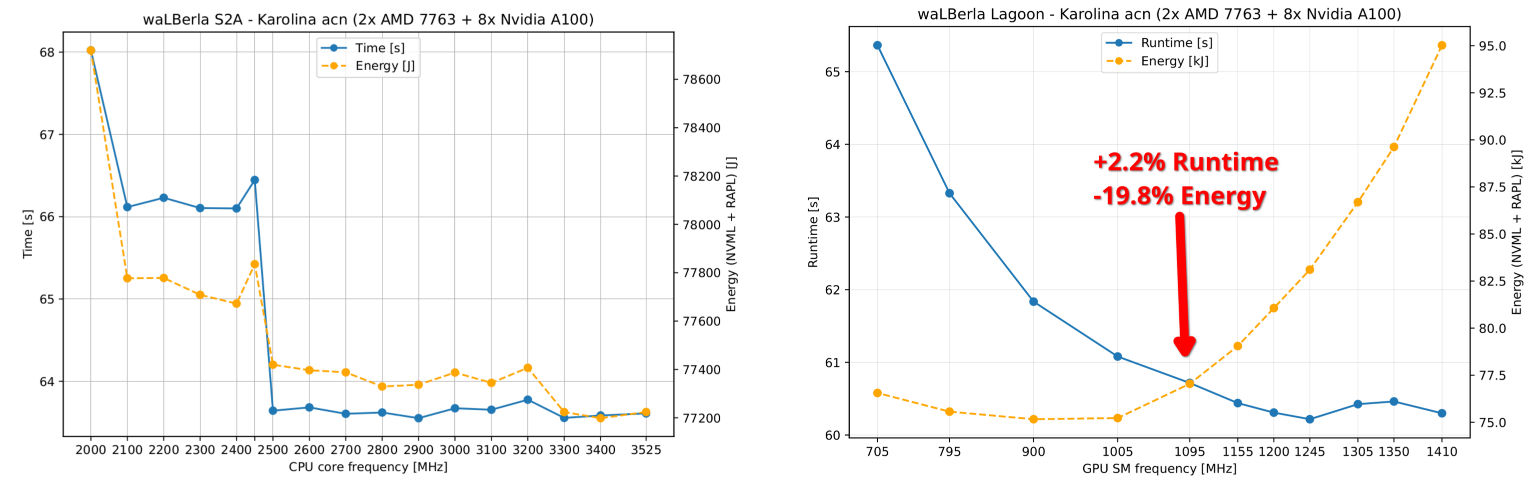

Fig. 5: Effect of Static Tuning for Walberla on GPUs

Performance Optimization

Figure 6 demonstrates strong scaling for LaBS on the Lagoon use case. The test case was performed on the CPU partition of the Karolina supercomputer at IT4I. The test case contains approximately 277 million fluid nodes and can scale acceptably up to 10240 cores, corresponding to approximately 30 thousand fluid nodes per core.

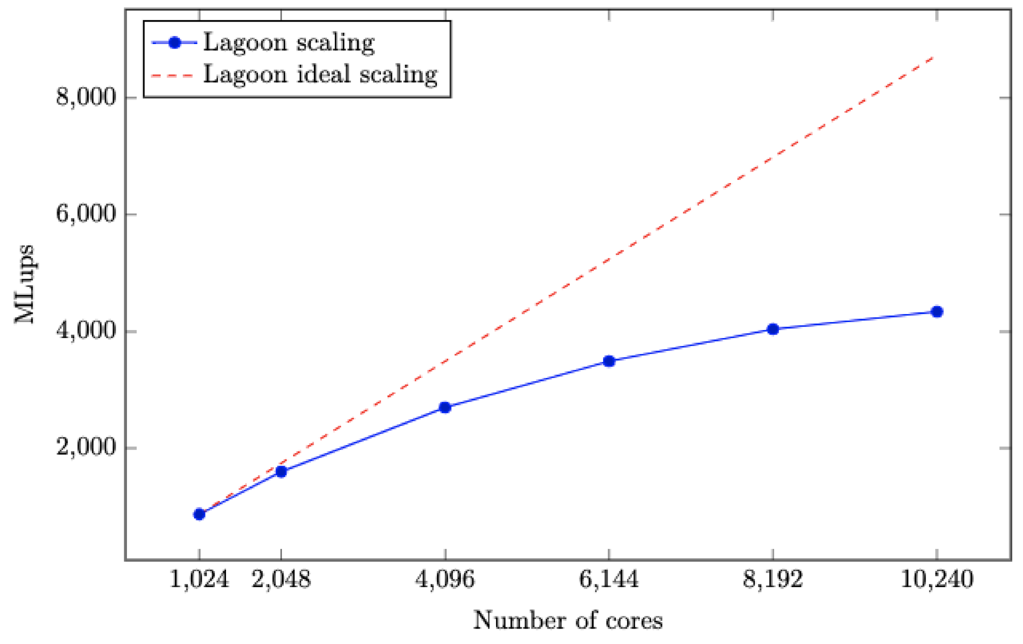

Fig. 6: Strong Scaling of LaBS for Lagoon Test

Figure 7 demonstrates strong scaling for waLBerla on the Lagoon use case. The test case was performed on the Booster module of the JUWELS supercomputer at JSC. The test case contains approximately 277 million fluid nodes and can scale acceptably up to 1024 A100 GPUs, corresponding to approximately 2 million fluid nodes per GPU.

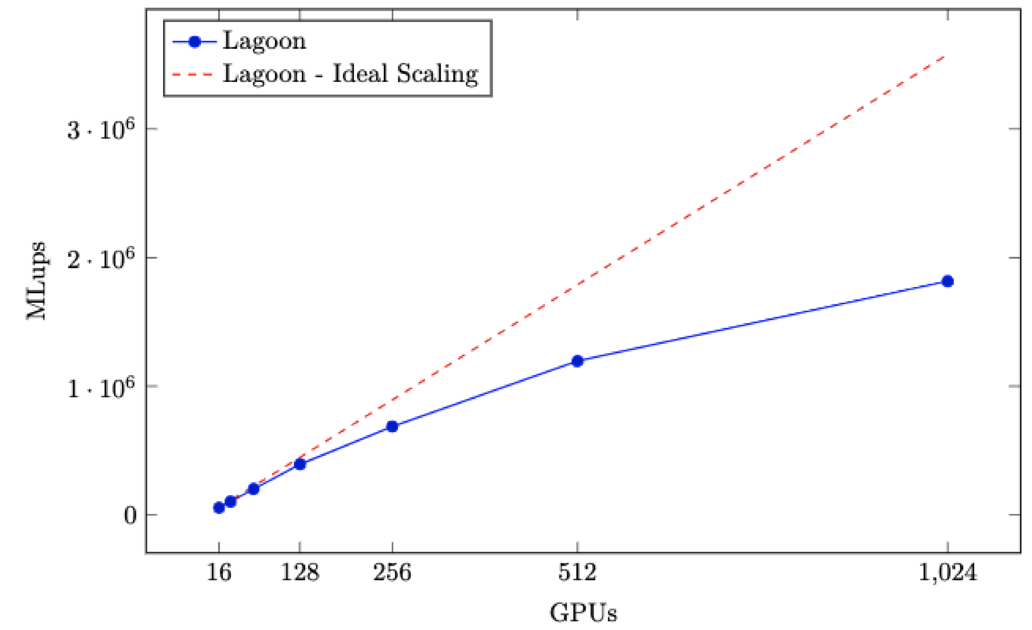

Fig. 7: Strong Scaling of waLBerla for Lagoon Test

Status and Future Plans

The SCALABLE project is in its final phase finishing at the end of 2023. In the final phase of the project, the project aims to improve upon and consolidate the results from the use cases into a single framework.

► This project has received funding from the European High-Performance Computing Joint Undertaking Joint Undertaking (JU) under grant agreement No 956000. The JU receives support from the European Union's Horizon 2020 research and innovation program and France, Germany, Czech Republic.