OpenFOAM Conference 2021

A Massive Simultaneous Cloud Computing Platform for OpenFOAM

G. Boiger[a], D. Sharman[a,b], B. Siyahhan[a], V. Lienhard[a], M . Boldrini[a], D. Drew[b]

a) ZHAW Zurich University of Applied Sciences, ICP Institute of Computational Physics; Wildbachstrasse 21, 8401 Winterthur, Switzerland; Prof. Dr.; +41 (0)58 934 77 93; gernot.boiger@zhaw.ch

b) Kaleidosim Technologies AG, Zurlindenstrasse 80, 8003 Zurich, Switzerland

1 Introduction

Today the field of numerical simulation in is faced with increasing demands for data-intensive investigations. On the one hand Engineering tasks call for parameter-studies, sensitivity analysis and optimization runs of ever-increasing size and magnitude. In addition the field of Artificial Intelligence (AI) with its notorious hunger for data, urges to provide ever more extensive, numerically derived learning-, testing- and validation input for training e.g. Artificial Neural Networks (ANN).
On the other hand the current 'age of cloud computing' has set the stage such that nowadays any user of simulation software has access to potentially limitless hardware resources.

In the light of these challenges and opportunities, Zurich University of Applied Sciences (ZHAW) and Kaleidosim Technologies AG (Kaleidosim) have developed a publically available [2] Massive Simultaneous Cloud Computing (MSCC) platform for OpenFOAM. The platform is specifically tailored to yield vast amounts of simulation data in minimal Wall Clock Time (WCT).

Spanning approximately nine-man-years of development effort the platform now features:

- An instructive web-browser-based user interface (*Web Interface*) (see [1] and [2]);
- An *Application Programming Interface* (API) (see [3]);
- A *Self-Compile* option enabling users to run self-composed OpenFOAM applications directly in the cloud;
- The *Massive Simultaneous Cloud Computing* (MSCC) feature which allows the orchestration of up to 500 cloud-based OpenFOAM simulation runs simultaneously (see [1] and [4]-[6]);
- The option to run *Paraview in Batch Mode* such that (semi-) automated cloud-based post-processing can be performed;
- The *Katana File Downloader* (KFD) allowing the selective download of specific output data (see [2]);

2 Massive Simultaneous Cloud Computing (MSCC) Work Flow

In essence the MSCC workflow is about the simultaneous orchestration of any number of OpenFOAM simulation runs on any number of cloud-based computers within a public cloud

(e.g.: Google Cloud). For now the actual platform has been tested and released to accommodate a maximum number of 500 simultaneous cloud simulation runs.
The MSCC-workflow as provided within the UI of the platform [2], encompasses the following steps (see also Fig.1):

- **Step 1:** Since the web-platform is explicitly not meant to be an OpenFOAM GUI, users proceed to prepare their OpenFOAM *Base Case(s)* locally with whatever pre-processor they see fit (see Fig.1 *Set Base Case*).

- **Step 2:** Users can copy source code of self-composed OpenFOAM applications and/or Python post-processing applications [7] (e.g.: for running *Paraview in Batch-Mode*) into the OpenFoam *Base-Case* folder.

- **Step 3:** Users zip their OpenFOAM *Base Case(s)* and transfer (see Fig.1: *Transfer Cases*) to the cloud via the *Web Interface* (see Fig.1: *Web Interface*).

- **Step 4:** Within the *Web Interface* users:
    - select any OpenFOAM version;
    - choose either an OpenFOAM *Standard Solver* [8] or
    - compile self-composed, to-be-uploaded OpenFOAM applications;
    - select cloud-based hardware with respect to the number of virtual CPU cores (vCPUs) and required RAM;
    - choose either the automatically created *Run Script* or
    - compose any *Bash Script* featuring e.g. OpenFOAM terminal command function(s);

- **Step 5:** Users then apply the *MSCC Run Creator* (se Fig.1: *MSCC Run Creator, [1]*) to set parameter ranges in order to multiply the *Base Case* such as to perform e.g. a parameter-study with up to 500 individual simulation cases.

- **Step 6:** Users initiate *MSCC* (see Fig.1: *Massive Simultaneous Computing*) where the cloud platform simultaneously:
    - reserves,
    - sets-up and
    - initializes one virtual machine per prepared simulation case yielding chosen hardware features and chosen OpenFOAM software version;

  Subsequently the platform runs:
    - pre-processing,
    - processing and
    - post-processing, as dictated by the *Run Script*.

  As long as the total amount of simulation cases stays ≤500, the total Wall Clock Time (WCT) (see Fig.1: T*) of the entire operation will stay within a maximum of 150% of WCT required to execute one single simulation yielding a *Net-Speed-up-Factor* (NSF) of up to 100.

- **Step 7:** The platform returns simulation results from individual virtual machines to a cloud database such as to ensure availability via the *Web Interface* (see Fig.1: *Return Results*).

- **Step 8:** Users can apply the *Katana Downloader* feature (see Fig.1: *Katana Downloader*) to selectively download any portion of data (e.g.: log-files, field data, JPEGs from Paraview post-processing, etc.) from any sub-group of computed cases.

- **Step 9:** Users apply any desired additional local post-processing procedure(s) (e.g.: Paraview-based, Python-based, OpenFOAM-based), (see Fig.1: *Post-Process Locally*).

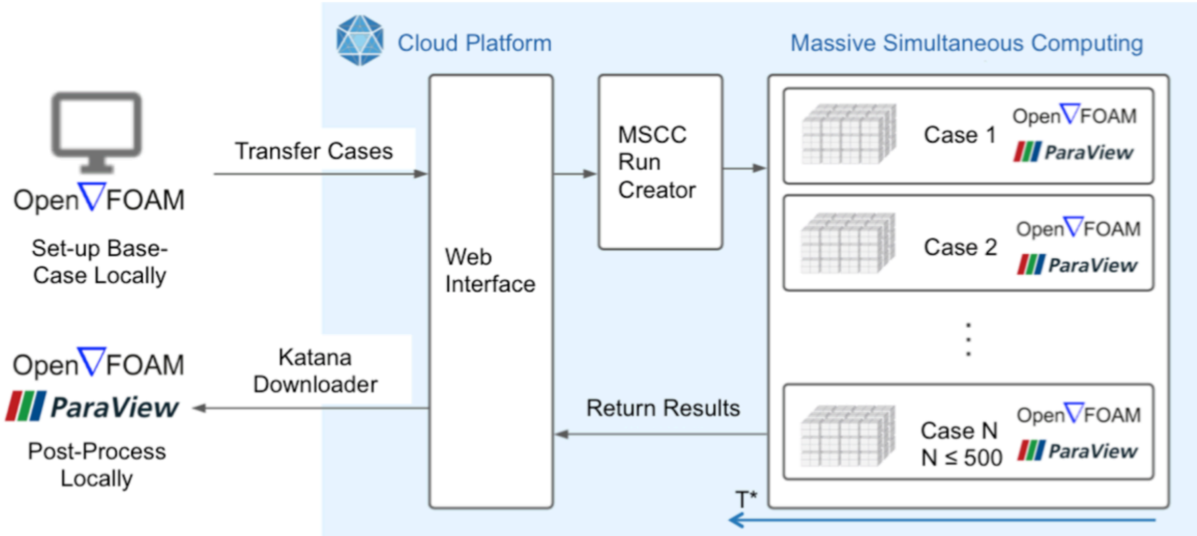

Fig.1: Scheme of MSCC workflow within novel cloud platform. T signifies the Wall Clock Time per individual simulation run (see also [2]).*

3 Application Programming Interface

Users retain the option to conduct the entire MSCC workflow as depicted within Fig.1 via an Application Programming Interface (API), accessible via [3]. Thus the necessity to access the *Web Interface* can be omitted. In this context user-defined scripting can allow for further elevated levels of cloud-machine-orchestration.

4 Fields of Application: Parameter Studies, Optimization and Artificial Neural Network Training

While the most obvious application of the platform lies within the swift execution of large *Parameter Studies* (see [1] and [4]-[6]), the API functionality can significantly extend the field of application. More specifically API scripting can be used to arrange sequences of MSCC runs in loops such as to enable e.g. *Optimization Algorithms* of a novel generation. Where common optimization procedures in CFD (see [9]) would rely on relatively many sequential iteration steps of relatively low order (see Fig.2: *Sequential Optimization*), the presented platform now allows the application of relatively few sequential iteration steps of relatively high order (see Fig.2: *MSCC Optimization*). While the overall amount of required computation would thus not necessarily decrease, global optima could be identified with higher confidence and the total WCT of optimization algorithms could potentially be reduced quite dramatically.

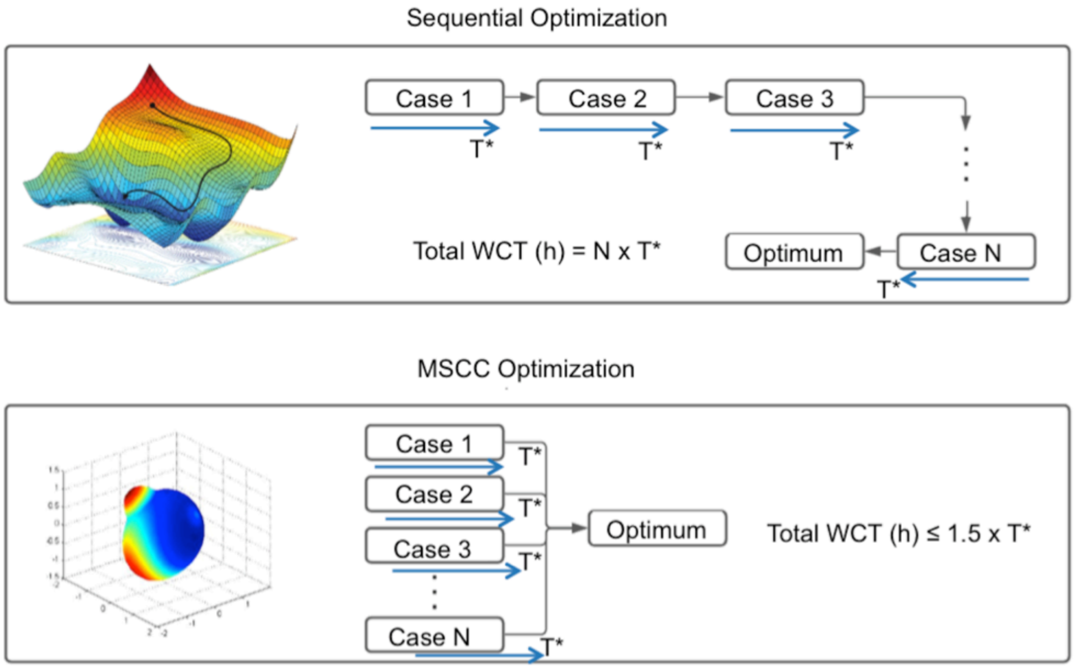

Fig. 2: Basic illustration of comparison of Sequential- vs. MSCC-based Optimization schemes. T signifies the Wall Clock Time per individual simulation run.*

API based scripting can also turn the OpenFOAM cloud platform into a comprehensive tool for producing potentially vast amounts of simulation data for the training, testing and validation of Artificial Neural Networks (ANNs).
Fig. 3 depicts the basic idea. Thereby the web-based *MSCC Run Creator* would be used to conduct a multitude of simultaneous OpenFOAM simulation runs whose outcome would provide testing-, validation- and training data for ANNs. As shown in Fig.3, it would make sense to add a database, which would store process specific data (e.g.: boundary conditions BCs) as well as weights of trained networks. Thus trained ANNs could potentially take over specific functionalities of the original numerical solver and could be integrated into accelerated later version of the same.

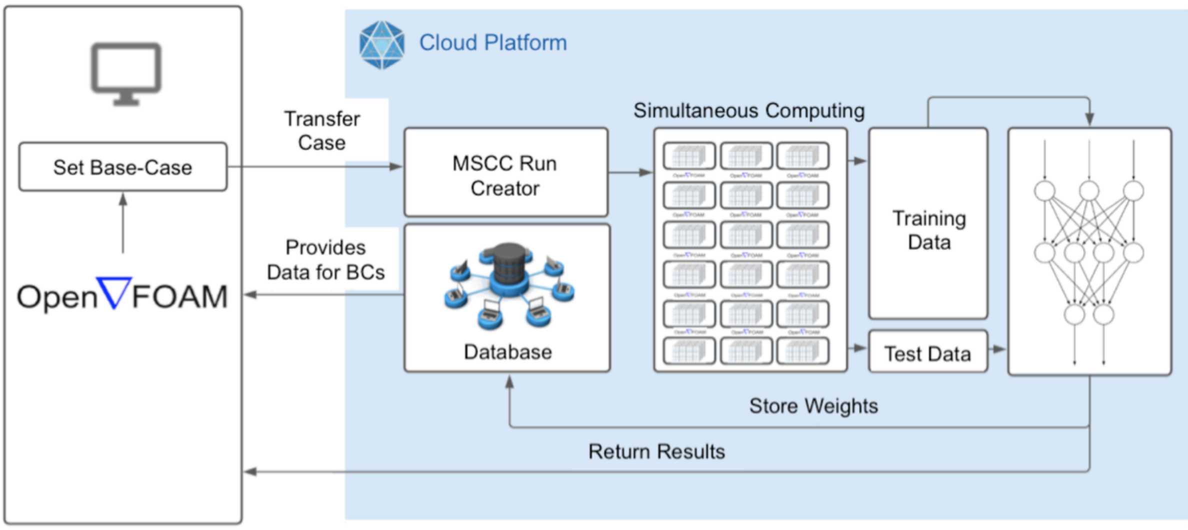

Fig. 3: Workflow scheme for using the MSCC cloud platform to simultaneously produce vast amounts of training data for ANNs.

5 Case Studies

Two case studies shall hereby be presented (see 5.1 and 5.2) in order to demonstrate the potential of the presented MSCC platform. Further application examples of the platform are documented in [4], [5] and [6].

5.1 Case Study A: Motorbike - WCT and Cost Comparison of Parameter Studies

Example A focuses on the well-known *Motorbike* tutorial case of *OpenFOAM 5.x* in the context of a steady state *simpleFoam* parameter study. The original tutorial case geometry was re-built, parameterized with respect to four geometric parameters (see $\Phi_1 - \Phi_4$ in Fig.4) and the grid was refined up to a total of approximately $1.6x10^7$ cells as compared to $2.2x10^5$ cells of the original tutorial case. In addition the onset airflow velocity vector was parameterized with respect to its magnitude Φ_5, its frontal- Φ_6 as well as its lateral- Φ_7 angle-of-attack. No other simulation parameters were modified as compared to the original OpenFOAM *Motorbike* tutorial case. Each of the seven *Degrees of Freedom* (DOF) Φ_i was then varied by three to four settings. The full bandwidth of possible combinations, amounting to a total of approximately 10'000 individual *simpleFoam* simulation cases was then run. The idea of this relatively extensive study was not to gain engineering-relevant insight but to be able to estimate cost- and time consumption in a direct comparison of:

- Running the full workflow on the modestly sized ZHAW internal cluster yielding 128 Intel Xeon 8164/8270 vCPUs with 2.0 – 3.7 GHz, parallelizing each case on 32 vCPUs.
- Running the full workflow on the MSCC cloud platform yielding up to 500 AMD Epyc Rome, 2.25-2.7 GHz machines per batch, where only machines with 32 vCPU cores were chosen for the same level of parallelization as the one used on the local cluster.

When working on the internal ZHAW cluster only 100 out of 10'000 cases were actually conducted. Batches of four cases each were uploaded, run and evaluated (i.e.: log-files containing cd and cl values were downloaded) by employees. Costs and time consumption were extrapolated. Working on the MSCC cloud platform the full parameter study of 12'000 cases was conducted where a total of 20 batches of 500 cases each were created, run and evaluated by employees in analogy to the workflow at the ZHAW cluster. Costs and time consumption were monitored.

The core-results are shown within Fig. 5 and Fig. 6 where the *Wall Clock Time* (WCT) consumption for running 100 cases as well as *Human Resource Costs* (HRC) for running all 10'000 cases are compared respectively.

The results presented in Fig. 6 are basically an extension of those shown in Fig.5. They demonstrate that the total WCT of running 10'000 *simpleFoam* $1.6x10^7$ cells cases on the ZHAW cluster (blue-dashed) would have amounted to 3356h, while running on MSCC cloud platform (red-dashed) amounted to a total of merely 58h. In order to estimate the HRC according to Fig.6, the following assumptions were made: hourly rates of 100€/h were chosen, 30 seconds working time per case for the ZHAW-cluster workflow and 240 seconds working time per case-batch for MSCC-platform workflow were assumed. On this basis the HRC of running the entire study on the ZHAW cluster (green-full) would have amounted to a total of 8317€ where the comparable number for running on the MSCC platform (purple-full) would have been a mere 133€.

In essence the numbers demonstrate that the MSCC concept makes the conduction of parameter studies as vast as the chosen example possible in the first place. While running on a (modestly sized) local cluster would invoke working- and computing times of 1000s of hours and thus be practically unfeasible, the hereby presented concept would actually enable the completion of the same work-load within two to three days.

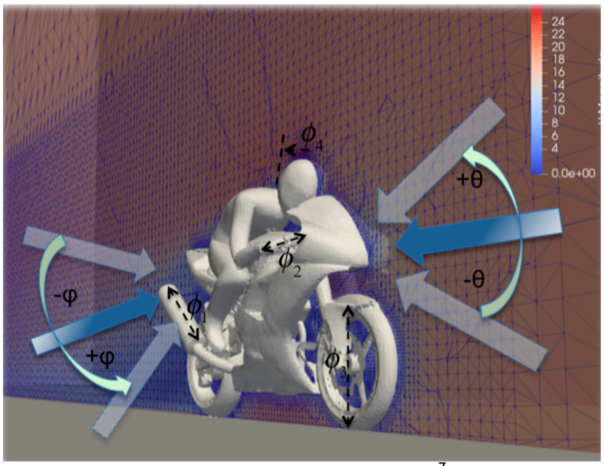

Fig. 4: Motorbike tutorial case of OpenFOAM 5.x, re-built to yield a 1.6x10^7 cells mesh instead of a 2.2x10^5 cells mesh.

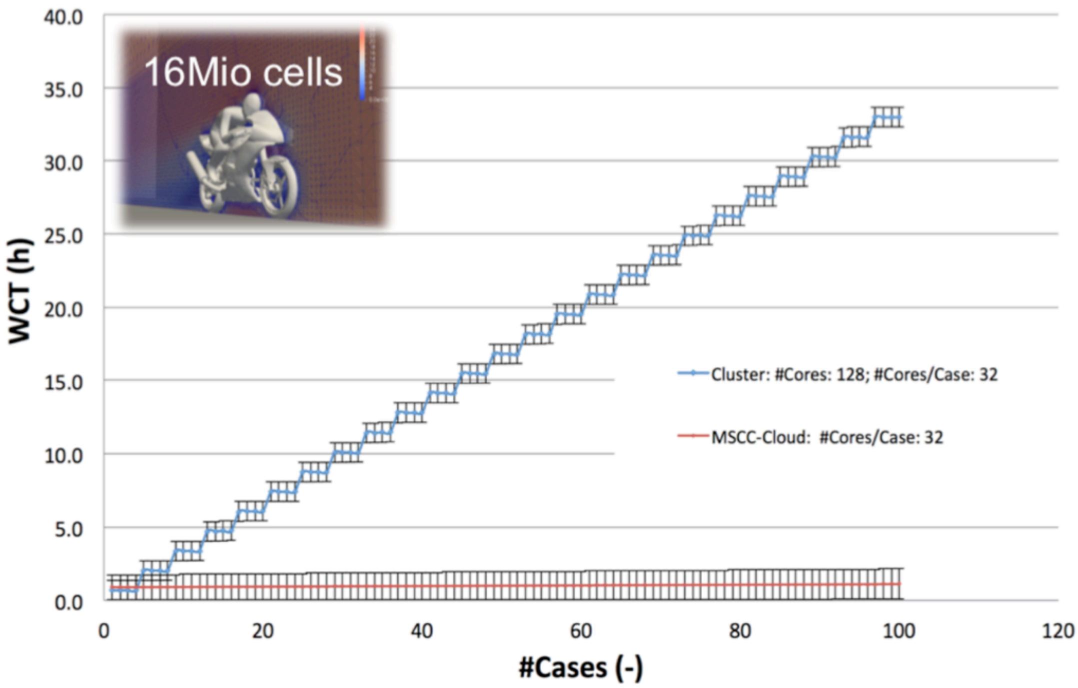

Fig. 5: Wall Clock Time (WCT) against number of cases run on ZHAW cluster (blue) and on MSCC cloud platform (red). Black candlesticks signify run-time of individual simulation cases (approximately 2h per case). On the internal cluster, batches of a maximum of four cases at a time can be run, while the cloud platform does not require any batch turn-over in this example, since maximum batch sizes of up to 500 cases can be realized.

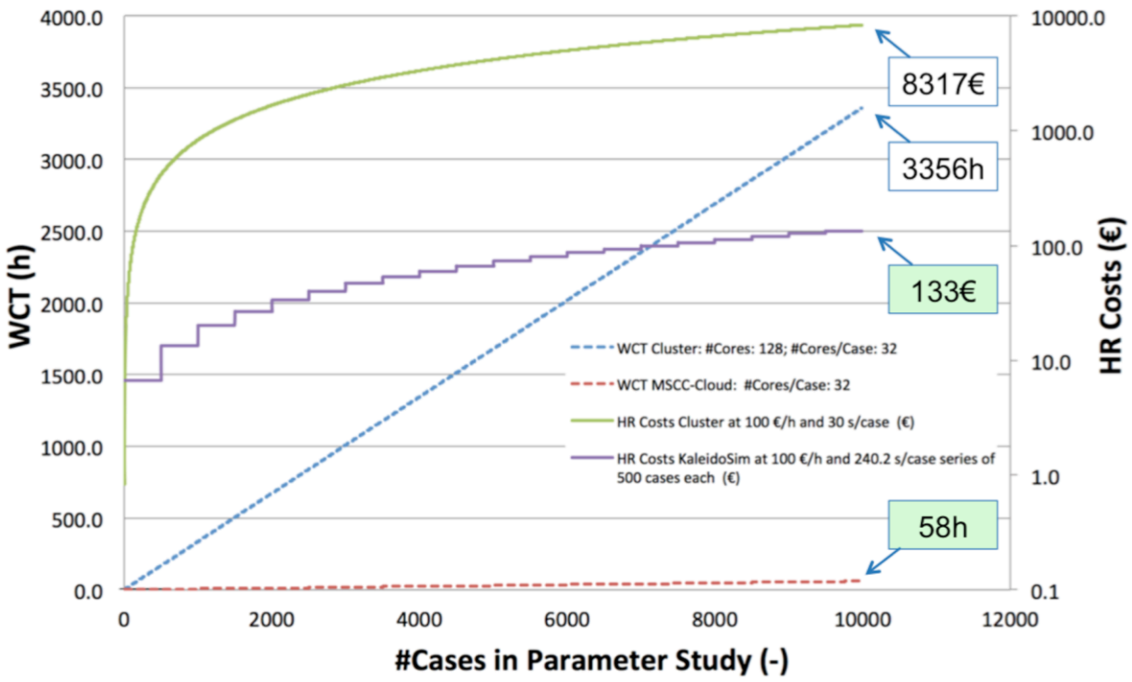

Fig. 6: Wall Clock Time (WCT) in hours (left axis) and Human Resource Costs in Euro (right axis) against number of cases run within parameter study.

5.2 Case Study B: Motorbike – 360° Angle of Attack

Example B is about the same basic simulation case as example A (*Motorbike*, *OpenFOAM 5.x*, *simpleFoam*, 2.2x10^5 cells). This time however a full 360° aerodynamic parameter study of the drag- and lift- coefficients was conducted. In this context the case was parameterized such that the angle-of-attack of the onset air velocity as well as the entire surrounding wind-channel was laterally rotated step by step. Step sizes of 3.6° and a full rotation of 360° were choose. The zero-degree *Base Case* was uploaded to the *Web Interface* and the *MSCC Run Creator* was applied such that 100 individual simulation cases were created and semi-automatically run in the cloud. A prepared *Python* post-processing utility was uploaded along with the *Base Case* and executed such that one *Paraview* snapshot of the fluid-dynamic situation was created per simulation case (see e.g. 9 out of 100 selected snap-shots in Fig. 7). Another prepared *Python* utility was applied to subsequently parse the downloaded *simpleFoam* logs, to extract and to evaluate one pair of drag- and lift coefficients per simulation case. The whole procedure yielded i) a series of 100 *Paraview*-created JPGs and a full 360° chart of the aerodynamic properties of the *Motorbiker* (for 180° result see Fig. 8). The remarkable point in this context is not so much the result itself but the fact that the whole work-flow could be conducted within a mere 20min.

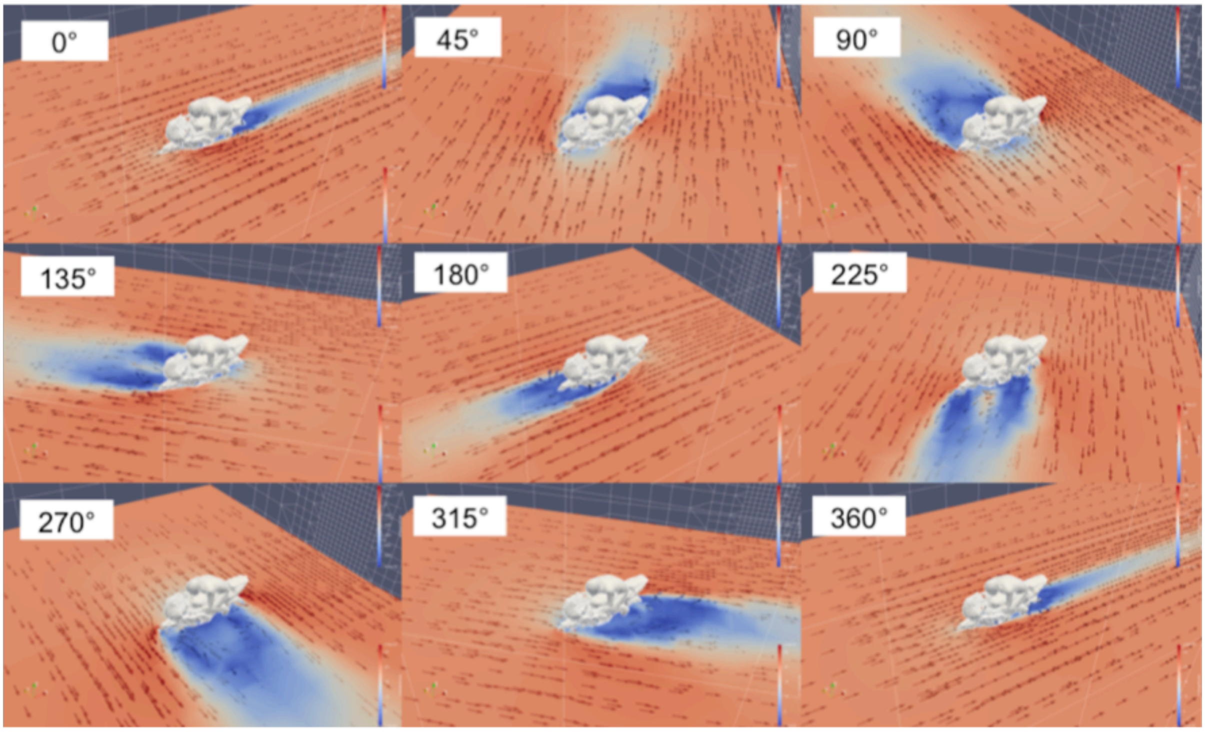

Fig.7: Nine out of 100 individual snapshots of the Motorbike simpleFoam tutorial cases computed with OpenFOAM 5.x in the MSCC cloud platform, where the angle-of-attack of the onset relative airflow velocity as well as the wind-channel were automatically rotated by a total of 360° around the Motorbike. 100 individual cases were conducted and thus evaluated simultaneously within 20min.

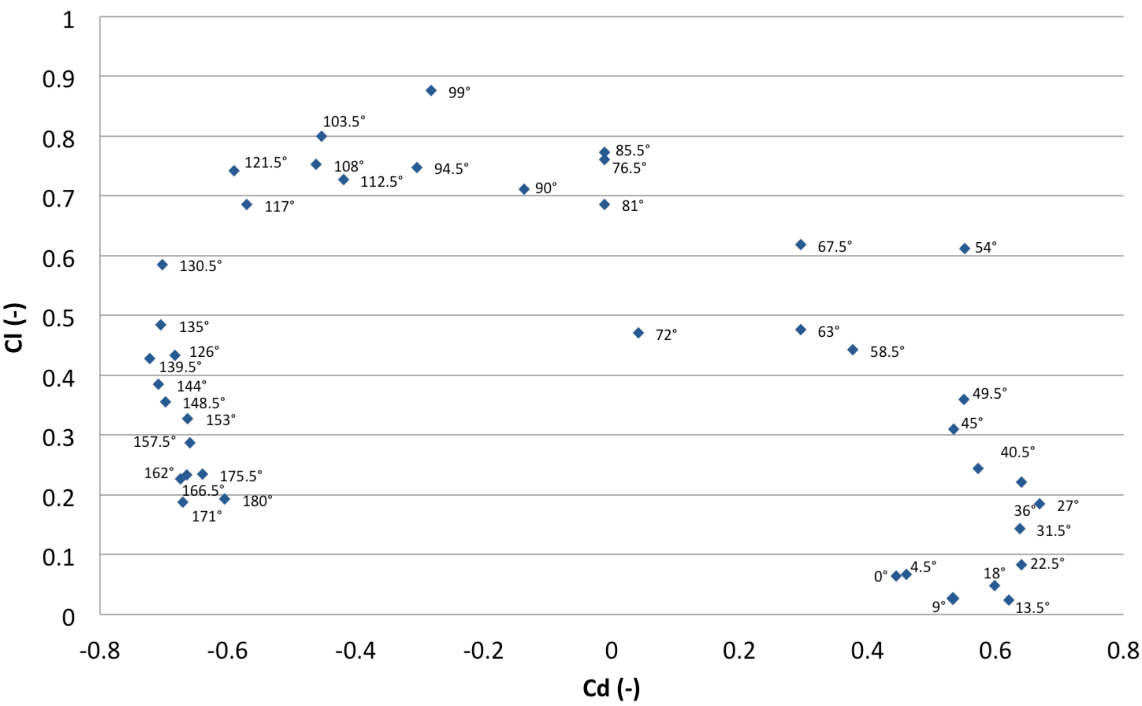

Fig. 8: Lift coefficients (c_l) versus drag coefficients (c_d) of OpenFOAM 5.x, simpleFoam tutorial case Motorbike with onset relative airflow velocity as well as entire wind channel being rotated step-by-step counter-clock wise around the Motorbike. Results derived from semi-automatically conducted 20min parameter study using MSCC platform depict 180° out of a total of 360° computed.

6 References

[1] Boiger, G.; Sharman, D.; Boldrini, M.; Everitt, V.; Dominic, D., 2020. Introducing Massive Simultaneous Cloud Computing (MSCC) for Multiphysics-Simulation Applications (MSCC). Multiphysics 2020. 15th International Conference of Multipysics, Online, 10-11 December 2020. International Society of Multiphysics. ISSN (online) 2409-1669.

[2] https://app.kaleidosim.com/

[3] https://api.kaleidosim.com/

[4] Boiger, G.; Buff, V.; Sharman, D.; Boldrini, M.; Lienhard, V.; Dominic, D., 2020. Simulation-based investigation of tar formation in after-treatment systems for biomass gasification (2020). Biomass Conversion and Biorefinery. 2020. pp. 1-18. DOI: 10.1007/s13399-020-00915-7

[5] Boiger, G.; Boldrini, M.; Lienhard, V.; Siyahhan, B.; Khawaja, H.; Moatamedi, M., 2019. Multiphysics Eulerian-Lagrangian Electrostatic Particle Spray- And Deposition Model for OpenFoam® and KaleidoSim® Cloud-Platform (2020). Int.Journal of Multiphysics. 14(1), pp. 1-15. DOI: 10.21152/1750-9548.14.1.1

[6] Boiger,G., Bercan,S.; Viktor,L.; 2020. Advancing the Validation and Application of a Eulerian-Lagrangian Multiphysics Solver for Coating Processes in Terms of Massive Simultaneous Cloud Computing. Multiphysics 2020. 15th International Conference of Multipysics, Online, 10-11 December 2020. International Society of Multiphysics. ISSN (online) 2409-1669.

[7] https://www.paraview.org/Wiki/ParaView/Python_Scripting

[8] https://www.openfoam.com/documentation/user-guide/a-reference/a.1-standard-solvers

[9] Thévenin D., Janiga G.; 2008. Optimization and computational fluid dynamics. Berlin, Heidelberg: Springer, 2008. DOI: 10.1007/978-3-540-72153-6