FEMZIP-OpenFOAM: Data Compression tailored for transient OpenFOAM cases

Stefan P. Müller, Pinaki S. Banerjee

SIDACT GmbH, Grantham Allee 2-8, 53757 Sankt Augustin, Germany, +49 228 534 80430,
stefan.mueller@sidact.com, pinaki.banerjee@sidact.com

OpenFOAM has a large user base across all application fields in Computational Fluid Dynamics (CFD). Its modular and object oriented structure has many advantages like expandability, but also leads to a huge number of files as well as a high redundancy in them. Especially in the case of transient models together with decomposition into several hundreds of processors, a single simulation can easily reach 1 Million files together with an overall size of several 100GBs. The challenge is to handle these large amounts of data in a time and storage saving manner. Difficulties regarding limitations on bandwidth and disk space can be approached by an application of lossy data compression.

In order to keep the number of files as well as the data size small, we created an on-the-fly compression in parallel to an ongoing transient OpenFOAM simulation, which is unique in the CAE context. The simultaneous compression process saves disk space on the cluster's scratch and improves the transfer times from clusters to servers. The applied compression tool FEMZIP-OpenFOAM is tailored for the compression of OpenFOAM simulation results and thus achieves compression factors above 10 on binary OpenFOAM data even when the data was initially compressed with gzip. Reductions in data volumes and number of files lead to a significant decrease in the requirements on storage, bandwidth and stability of filesystems. Moreover, the application on stationary OpenFOAM cases is also beneficial resulting in compression factors of about 8 to 10.

To improve the usability of compressed datasets, our compressed output format is designed similar to the OpenFOAM format structure. This results into an easy integration into individual post-processing workflows and on-the-fly access of data subsets without the requirements of original data reconstruction. The compression techniques are designed in a way such that reading uncached, compressed OpenFOAM cases from a local HDD along with decompression and merging of distributed data is faster than reading the corresponding uncached, original cases into a postprocessor.