Interactive Visualization of Large Simulation Datasets

Abstract

At the German Aerospace Center (DLR), numerical simulation is an essential tool for many application areas. The increasing compute power of supercomputers enables tremendous improvements in detail and accuracy of simulation results. At the same time, however, the volume of the resulting data files is growing rapidly. For example, a high-end aerodynamics simulation of a complete airplane today can produce up to 100 Terabytes of data. Post-processing of such large amounts of data for visualization and then analyzing the simulation results become both time consuming and demanding in terms of hardware resources. Therefore, parallelization of the post-processing pipeline is required.

We give an overview of state-of-the-art methods for the optimization of the parallel post-processing pipeline, to eventually enable interactive exploration of very large datasets. Typically the heavy post-processing work, such as converting raw datasets and feature extraction, is located at the beginning of the pipeline. We present a parallelization framework based on the visualization toolkit (VTK). It distributes the first post-processing stages to multiple processes to be executed on hardware with high-bandwidth access to the data sources. The extracted visualization objects, which are significantly smaller, are sent to the user site, where they are assembled by software that plugs directly into a virtual reality toolkit for rendering and user interaction. Geometric partitioning, e.g. using multi-block data structures, has proven to be useful for parallelization, as the resulting datasets can be easily distributed and processed in parallel. Additionally it allows for an overlap of the processing of data at the backend and visualization at the frontend. As soon as a block is processed, the resulting visualization primitives are sent to the front end. This results in a significant acceleration of the visual response.

We will close with our view of future research, including the exploitation of modern hardware architectures, such as many-core CPUs and GPUs. Co-location of computation and extraction processes within an HPC system promise a substantial reduction of the disk I/O bottleneck. Rather than first writing the simulation results to disk and then reading them back in for post-processing, the simulation and the first stage of the post-processing pipeline run together on the same HPC system and communicate by message-passing. Additionally, the use of multi-resolution data structures allows fast access to data and enables progressive data streaming. This provides a rough overview of results quickly to the engineer with the high resolution visualization emerging gradually. Finally, interactive data exploration enables computational steering since the engineer can change simulation parameters while monitoring the results during an ongoing computation.

Andreas Gerndt, Rolf Hempel, Robin Wolff
German Aerospace Center (DLR)
Simulation and Software Technology