Workshop: Strategies for Data Science and Data Management



Contribution ID: 52 Type: Poster

Parallel, Distributed, Adaptive Simulation Data Management

A key to defining the organizational structure of simulation data is the computational mesh. It encodes where in space, and possibly in time, simulation data points are located. This is essential for the inner functioning of the simulation on the one hand and for in-situ processing, storing, and post-processing the data on the other. The primary danger is losing the parallel, highly optimized access to the data when moving it from simulation memory to intermediate exchange formats and/or mass storage.

We illustrate perspectives that reach beyond pure simulation at the example of p4est. The p4est software library implements fast algorithms for large-scale distributed adaptive mesh refinement and data location and serves as the parallel mesh backbone for various third-party simulation codes. It employs a distributed forest-of-octrees data structure and has been demonstrated to scale to 3e6 MPI processes and .5e12 mesh elements. p4est lends itself as a generic tool for partition-independent management of spatial data. This means that the specific division pattern of data among the many parallel processing units shall influence neither the results of the computation nor the definition of any data exchange format. This capacity is a precondition for flexible, reproducible re-processing.

On this poster we present several aspects. First of all, we illustrate the concept and practical role of parallel partitioning in general. Second, we present a partition-independent I/O mechanism for simulation data. It employs the MPI I/O standard and at the same time guarantees write- and read-equivalent files over an arbitrary partition and even without MPI support. The stored data continues to be amenable to our highly scalable algorithms used for spatial search and simulation. Third, we provide an example of a p4est-based partition-independent simulation. The octree-based design enables efficient remote search functionalities, for example to locate physical measurement points (such as floating tide gauges) and to integrate over rays or curves (such as lines of sight in satellite imaging).

Primary authors: Mr GRIESBACH, Tim (University of Bonn Institute for Numerical Simulation); Mr BRANDT, Hannes (University of Bonn Institute for Numerical Simulation); Prof. BURSTEDDE, Carsten (University of Bonn Institute for Numerical Simulation)

Presenter: Mr GRIESBACH, Tim (University of Bonn Institute for Numerical Simulation)

Session Classification: Posters