ALOMA: A Parallelization Framework – Not Only for Seismic Applications

One of the challenges of Big Data is to correctly execute complex computing operations on growing volumes of data. It can only be mastered with the massive parallelization of computers and computing power.

Researchers in the High Performance Computing department have been using GPI-Space for years as a programming platform for high-performance systems. What makes GPI-Space so special is its user-friendliness: Customers do not need to acquire any special HPC knowledge, because the software system takes care of the efficient execution of the algorithms. This also applies to the specialized version ALOMA.

ALOMA detects data dependencies

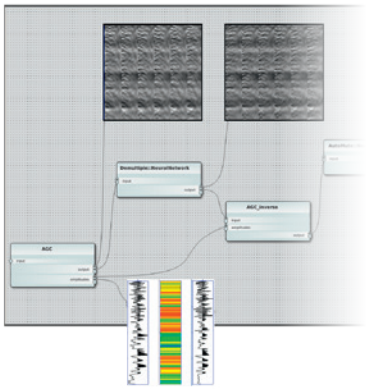
The tool is primarily used in seismics. However, it is designed as a general framework for the execution of workflows on distributed systems, because ALOMA recognizes dependencies in the data sets that are supplied as input – regardless of the data source – and answers the following questions: How is the data distributed? Which data can be processed simultaneously and independently? Where can the data be processed?

"In order to identify dependencies and answer these questions, ALOMA must have or generate appropriate information about the individual modules and their combination in the workflow. For this workflow, we use a representation as a Petri net, which the internal workflow engine can analyze accordingly," explains project manager Dr. Dirk Merten. To do this, users only have to connect outputs of algorithmic modules with the inputs of further modules in a graphical workflow editor. Information about the individual modules is provided when they are integrated into ALOMA. This information includes the number and types of the modules' input and output data, as well as their granularity. These are values specific to seismic (e. g. "Seismic Data" or "Velocity Model") and the granularities "Trace", "Gather" or "Inline".

Traditional algorithms and ML

In addition to the granularities typical for seismics, ALOMA also supports data splits from pattern recognition and machine learning. This means that modules derived from Deep Learning can also be integrated. In an exemplary workflow, a volume of input data can be corrected multiple times, stacked into one volume, and analyzed for error patterns. Processing and analysis modules for traditional algorithms and neural network inference are easily combined in one workflow. ALOMA automatically handles the parallelization of all modules within the workflow.

With ALOMA, seismic experts can integrate new traditional or machine learning-based algorithms and prototypes within a very short time, combine them with existing algorithms in workflows, and test them in parallel on realistic data sets and under everyday conditions.



Section of an exemplary workflow with ALOMA

Further information on our website www.itwm.fraunhofer.de/aloma_en

Kontakt

Dr. Dirk Merten
Department "High Performance Computing"
Telefon +49 631 31600-4616
dirk.merten@itwm.fraunhofer.de