IAF EARTH OBSERVATION SYMPOSIUM (B1)
Interactive Presentations - IAF EARTH OBSERVATION SYMPOSIUM (IP)

Author: Dr. Fabian Hoppe
Deutsches Zentrum für Luft- und Raumfahrt e.V. (DLR), Germany

Mr. Wadim Koslow
Deutsches Zentrum für Luft- und Raumfahrt e.V. (DLR), Germany
Dr. Kathrin Rack
Deutsches Zentrum für Luft- und Raumfahrt e.V. (DLR), Germany
Dr. Alexander Rüttgers
DLR (German Aerospace Center), Germany

EXPLORING AND PROCESSING LARGE DATA SETS IN EARTH OBSERVATION ON HPC-SYSTEMS WITH HEAT

Abstract

Handling and analyzing massive data sets is particularly important in the field of earth observation. Nevertheless, this can be challenging, especially for researchers and developers without a background in high-performance computing (HPC). The Python library Heat ("Helmholtz Analytics Toolkit"), jointly developed by DLR, Research Center Jülich, and Karlsruhe Institute for Technology, aims at supporting such researchers and developers by providing general-purpose, memory-distributed and hardware-accelerated array manipulation, data analytics, and machine learning algorithms in Python, targeting the usage by non-experts in HPC.

The pillars of its development are...

- ...to enable memory distribution of n-dimensional arrays,

- to adopt PyTorch as process-local compute engine (hence supporting GPU-acceleration),

- to provide memory-distributed (i.e., multi-node, multi-GPU) array operations and algorithms, optimizing asynchronous MPI-communication (based on mpi4py) under the hood, and

- to wrap functionalities in NumPy- or scikit-learn-like API to achieve porting of existing applications with minimal changes and to enable the usage by non-experts in HPC.

In this talk we will show-case how Heat can help to facilitate exploration and processing of large amounts of data in earth observation on HPC-systems. This includes, e.g., typical pre-/post-processing tasks such as dimensionality reduction, clustering, n-dim. FFT etc. In particular, we will discuss results of and experiences with Heats application in a current DLR-project on anomaly detection in an earth observation context.