IAF EARTH OBSERVATION SYMPOSIUM (B1)
Earth Observation Data System Development and Management (4)

Author: Dr. Paolo Mazzucchelli
ARESYS, Italy, paolo.mazzucchelli@aresys.it

Mr. Marco Codazzi
ARESYS, Italy, marco.codazzi@aresys.it
Dr. Fabio Gerace
ARESYS, Italy, fabio.gerace@aresys.it
Mr. Riccardo Piantanida
ARESYS, Italy, riccardo.piantanida@aresys.it
Dr. Luca Ferraro
CINECA, Italy, l.ferraro@cineca.it
Mr. Eric Monjoux
ESA - European Space Agency, Italy, eric.monjoux@esa.int
Ms. Betlem Rosich Tell
European Space Agency (ESA/ESRIN), Italy, betlem.rosich@esa.int

HIGH PERFORMANCE COMPUTING FOR EARTH OBSERVATION

Abstract

The availability of freely accessible remote sensing datasets (as provided within ESA Copernicus Missions) has paved the road to the definition of improved processing techniques that call for increased computing power. Moreover, the future Copernicus missions will also result in a new dimension of generated data volumes and hundreds of Earth Observation (EO) satellites are expected to be launched as large constellations over the next decade, with the satellites operating jointly and providing data for a plethora of new applications. Thus, the increased dimensionality of available data (due to both the higher number of EO satellites and to the continuously improving spatial resolution and coverage), along with the complexity and diversity of the required processing algorithms demand for novel EO processing capabilities: the ability to take advantage of massive parallel computing resources allowing to implement highly scalable applications can bridge the gap between the increased processing requirements and the new envisioned data exploitation opportunities. The European Community has also identified the availability of computing power as an outmost important competitive advantage, developing the European High Performance Computing Joint Undertaking (EuroHPC JU): a joint initiative between the EU, European countries and private partners to develop a World Class Supercomputing Ecosystem in Europe for the exascale and post-exascale era (https://eurohpc-ju.europa.eu/). In fact, the availability of (low cost) on-demand computing power has already contributed to change the scenario of data processing in the last few years: the power of a system has increased by an astonishing factor of 10^6, as average in the last 35 years, while computer clusters are the de-facto standard available machines on the market (www.top500.org). Their peak performances can be exploited only by parallel applications, that can take advantage of both multiple hosts and accelerators (i.e., GPUs). A unified processing framework for EO data processing on HPC premises is defined to effortlessly exploit parallel computing resources and allow to:

- build-up a modern and future-proof approach capable of managing a generally defined processing workflow,

- achieve optimal operational computational performances taking advantage of parallel hardware capabilities.

Such a system can be easily tailored to make the most of the different available processing models to select the most effective processing strategy with respect to any defined metric. Results on an HPC bulk reprocessing test for Sentinel-1, along with the lessons learned through the years in the Energy sector (the early adopters of HPC) prove the capabilites of HPC as a powerful asset for EO novel number-crunching applications.