



Distributed Processing of Sentinel-2 Products using the BIGEARTH Platform

Victor Bacu, Teodor Stefanut, Constantin Nandra, Danut Mihon, and Dorian Gorgan
Technical University of Cluj-Napoca, Cluj-Napoca, Romania (victor.bacu@cs.utcluj.ro)

The constellation of observational satellites orbiting around Earth is constantly increasing, providing more data that need to be processed in order to extract meaningful information and knowledge from it. Sentinel-2 satellites, part of the Copernicus Earth Observation program, aim to be used in agriculture, forestry and many other land management applications. ESA's SNAP toolbox can be used to process data gathered by Sentinel-2 satellites but is limited to the resources provided by a stand-alone computer. In this paper we present a cloud based software platform that makes use of this toolbox together with other remote sensing software applications to process Sentinel-2 products.

The BIGEARTH software platform [1] offers an integrated solution for processing Earth Observation data coming from different sources (such as satellites or on-site sensors). The flow of processing is defined as a chain of tasks based on the WorDeL description language [2]. Each task could rely on a different software technology (such as Grass GIS and ESA's SNAP) in order to process the input data. One important feature of the BIGEARTH platform comes from this possibility of interconnection and integration, throughout the same flow of processing, of the various well known software technologies. All this integration is transparent from the user perspective. The proposed platform extends the SNAP capabilities by enabling specialists to easily scale the processing over distributed architectures, according to their specific needs and resources.

The software platform [3] can be used in multiple configurations. In the basic one the software platform runs as a standalone application inside a virtual machine. Obviously in this case the computational resources are limited but it will give an overview of the functionalities of the software platform, and also the possibility to define the flow of processing and later on to execute it on a more complex infrastructure. The most complex and robust configuration is based on cloud computing and allows the installation on a private or public cloud infrastructure. In this configuration, the processing resources can be dynamically allocated and the execution time can be considerably improved by the available virtual resources and the number of parallelizable sequences in the processing flow.

The presentation highlights the benefits and issues of the proposed solution by analyzing some significant experimental use cases.

Main references for further information:

- [1] BigEarth project, <http://cgis.utcluj.ro/projects/bigeearth>
- [2] Constantin Nandra, Dorian Gorgan: "Defining Earth data batch processing tasks by means of a flexible workflow description language", *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, III-4, 59-66, (2016).
- [3] Victor Bacu, Teodor Stefanut, Dorian Gorgan, "Adaptive Processing of Earth Observation Data on Cloud Infrastructures Based on Workflow Description", *Proceedings of the Intelligent Computer Communication and Processing (ICCP)*, IEEE-Press, pp.444-454, (2015).