Flexible description of earth observation data processing

Big Data is a challenge in many fields of research, one important domain being those dealing with Earth Observation (EO) data. The main issues, in this case, are the volume and velocity of data acquired by an increasing number of aerial and space-borne sensors, but as well the complexity of data mining process. Thus, the sheer volume and acquisition rates may threaten to overwhelm many organizational storage capabilities, leading to situations in which data value is overtaken by storage costs, which will, in turn, lead to loss of data. Data can create value only when it is used, and the data protection has to be oriented toward allowing innovation that sometimes depends on creative people, which achieve unexpected valuable results through a flexible and adaptive manner. The users need to describe and experiment themselves different complex algorithms through analytics in order to valorize data. The analytics uses descriptive and predictive models to gain valuable knowledge and information from data analysis. The BigEarth platform offers solutions for increasing the efficiency of the data processing efforts by combining two approaches. The first one experiments the high-performance, cloud-based computing solution in order to shorten the overall execution time. The second approach provides a highly flexible description of the data processing tasks in order to improve the modularity of the design and promote reusability. As a means of specifying the structure of the workflow, the BigEarth platform uses a specially designed description language called WorDeL (Workflow Description Language). This presentation aims to highlight and exemplify some of the main features of the language, and demonstrate their usage in defining Earth Data processing tasks. The WorDeL language is based on the flexible description of processing tasks as workflows, composed of basic processing operators. With this approach, the language offers an intuitive way of representing processing tasks, without requiring programming expertise from its users. It also allows its users to employ and integrate existing functionality into their design, thereby reducing the complexity and development effort of newly defined processing workflows. The WorDeL language supports the transparent adaptive parallelization of the processing tasks over high performance computation architectures, such as cloud based solutions.

Biography

Dorian Gorgan is Professor in Computer Science Department of the Technical University of Cluj-Napoca and PhD supervisor in Computers and Information Technology, and coordinator of the Computer Graphics and Interactive System Laboratory. His fields of interest involve parallel and distributed processing over HPC infrastructures such as Grid, Cloud, Multicore, and cluster, development of platforms and applications for spatial data processing and visualization, interdisciplinary research in the domains of Earth Sciences and Earth Observations. He has been involved as scientific coordinator and WP leader in national and international research projects such as BIGEARTH, PECSA, envirolGRIDS, IASON, SEE-GRID-SCI, GISHEO, mEducator, iTRACE, MedioGrid, COMPLEXHPC, and KEYSTONE. He has been member of scientific and reviewing committees of many ISI journals and international conferences, and gave more than 300 papers and presentations in journals and prestigious conferences in the domains of Computer Science and Earth Observation.

dorian.gorgan@cs.utcluj.ro

Notes: