A Central Processing Facility within a Distributed Data Processing System

S. de Witte, S.M. Rispens, and R.M. van Hees

SRON Netherlands Institute for Space Research, Earth Oriented Science Division, Utrecht, Netherlands (s.de.witte@sron.nl)

In a complex scientific data processing project, where raw satellite data (Level 1) is processed to end products (Level 2), you may need specific expertise from various groups in different locations. Collaboration between these groups can lead to better results and give the opportunity to try several different scientific approaches and choose, objectively, the best result. Furthermore, such a distributed data processing system or DDPS can be used for independent validation before the end products are released.

All participating groups need common and specific data products for their processing. This involves many interfaces needing and producing different data products. Without a central storage location all groups involved have to implement their own checking routines and transformations in order to use the data products. A central processing facility, acting as a single point of interface between the DDPS and the main data provider as well as for all groups within the DDPS, can facilitate in collecting all scientific data necessary for high-level processing, transforming the Level 1 input data to a DDPS internally agreed format, checking all data products on integrity, format and validity, distributing these data products within the DDPS, monitoring the whole data distribution chain and distributing all end products to the main data provider.

A DDPS has been implemented for ESA's gravity mission, GOCE (Gravity field and steady-state Ocean Circulation Explorer). GOCE's DDPS is called the High-level Processing Facility (HPF) and is part of the GOCE Ground Segment, developed under ESA contract by the European GOCE Gravity consortium (EGG-c). The HPF is set up as a distributed facility consisting of several sub-processing centers for scientific pre-processing, orbit determination, gravity field analysis and validation. The sub-processing facilities are connected through a central node, the Central Processing Facility (CPF).

The CPF has been thoroughly tested and is ready to accept real GOCE data. The CPF concept and its unique features can be very useful for future missions that require complex scientific data processing in different locations.