Gaia is ESA’s ambitious space astrometry mission the main objective of which is to astrometrically and spectrophotometrically map 10^6 celestial objects (mostly in our Galaxy) with unprecedented accuracy. The satellite will downlink close to 10^9 bytes of raw telemetry data over 5 years. To achieve its required accuracy of 10^{-5} as of micro-second astrometry, intricate processing of this data is required. In addition to the main astrometric mission Gaia will host a Radial Velocity instrument, two low-resolution dispersers for multi-colour photometry and two Star Mappers.

The basic principle of the astrometric data reduction is to determine all the variables of the system (source positions, attitude, calibration, photometry, spectroscopy). In order to achieve an optimal match with the observations, the data reduction will be performed in several cycles; in each cycle new data from the satellite will be processed and the starting solution will be improved.

The Gaia Data Processing and Analysis Consortium (DPAC) has been formed recently and will answer ESA’s announcement of opportunity for the data processing. All of Europe’s expertise in astrometry will be needed in order to reduce Gaia data to the level of accuracy needed to meet the mission goals.

ABSTRACT

The data processing requirements for Gaia are amongst the most challenging of any scientific endeavour to date. Due to the immense volume of data that will be collected, for 10^9 CCDs, it will be a major challenge even by the standards of computational prowess in the next decade to process, maintain and extract the scientific results necessary to build a unique space map of our Galaxy, the Milky Way. Gaia is in some sense the astronomical equivalent of the world’s largest terrestrial project: a engineering undertaking launched by ESA.

The Gaia data processing chain consists of several stages: Initial and intermediate data reduction, Main Data Base, Spectral Processing and Data Structure, Stochastic computing. The work is supported by a set of Data Processing Centres (DPCs) responsible for the data reduction. These activities are supported by a set of Data Processing Centres (DPCs) responsible for the data processing. The DPCs are tasked with the production of the processing systems. The CUs will be scaled and exercised using up to 10^6 million sources and more complex data.