ER-flow Application Description Template

Application Name *(will be used as workflow name in the repo)*:
Simulations of the ESA Planck mission

Application domain *(choose one existing in the repo, otherwise will be created)*:
Astronomy & Astrophysics / Cosmology

Brief description of application *(explain implemented function, inputs, outputs, usage)*:
The workflow implements the Planck LFI mission. The workflow is designed on the requirements of applications that cannot be handled by a single computing farm, both in terms of computing power and data storage.
The workflow consists of a very simple pipeline which is constituted of different software modules. The basic steps of the pipeline are described below:

- the CMB power spectrum is created with cmbfast starting from cosmological parameters;
- the CMB maps are built starting from the CMB power spectrum with synfast code being part of the HEALPix package;
- the CMB is combined with foregrounds with their own frequency dependent intensities and the final sky is convolved with the beam pattern for each of the detectors considered in the simulation;
- the map is contaminated by introducing instrumental noise which is computed and added to the "observed" sky signal, therefore the TOD (Time Ordered Data) is built.

The knowledge level increases over the time, hence new details are introduced and the whole computational chain is iterated many times, even during the operative phase of the mission.
In order to speed up calculation we can assume a perfect overlapping between samples in two consecutive scan circles of the spacecraft when it remains in the same pointing position. In this way the sky signal is always the same for all the 60 scan circles corresponding to the same pointing position and we can therefore simulate it only once. We refer to this "fast" simulation procedure as "short" run; "long" runs instead correspond to complete simulation procedures where each scan circle is kept distinguished from the other ones.

   data:
      input data format: tar file containing a set of configuration files used by the pipeline stages and the command line attributes used by the pipeline.
      output data format: a complex set of tar files containing TODs (up to ~50 GB), frequency maps, maps of components, noise contributions
   sample data   (link): ---
   application   (link): ---
   documentation (link): ---
   publication   (link): http://www.sciencedirect.com/science/article/pii/S0167739X06000860

Execution environment
The DCI is the one operated by EGI. A new dedicated VO was set up for the Planck community to run simulations; any other VO supporting the A&A community, however, is used.
middleware: gLite   workflow system: ---

Execution characteristics
   data size (per unit, typical number of units): 22 units, one for each radiometer

	input	temporary	output
short	a few MB	as for output	0.9 GB (30 GHz), 1.2 GB (44 GHz), 1.7 GB (70 GHz)
long	a few MB	as for output	34.2 GB (30 GHz), 45.3 GB (44 GHz), 75 GB (70 GHz)

processing time (per unit):
short   12 m   (30 GHz), 13 m (44 GHz), 17 m (70 GHz)
long   389 m (30 GHz), 623 m (44 GHz), 834 m (70 GHz)
memory usage: 2 GB   disk usage: 1.3 TB

Target users

At a first instance the end users interested to run the Planck simulations are the members of the Planck collaboration itself. All cosmologists however are interested to run these simulations.

Usage scenario for workflow in the ER-FLOW *(how workflow will be reused, meta-workflow, how expected to contribute to project indicators, etc.)*.

A simulation starts from a set of values for cosmological parameters. The simulation builds an ideal sky, contaminates it and extracts new maps; a new set of parameters is obtained starting from them.

As shown above, different software components contribute to build the whole pipeline run; this typical modular structure fosters the reuse of single simulation modules to build new applications and workflows.

Contact information (author)

    name: Claudio Vuerli   e-mail: vuerli@oats.inaf.it

    name: Giuliano Castelli   e-mail: giuliano.castelli@oats.inaf.it

    name: Giuliano Taffoni   e-mail: giuliano.taffoni@iasfbo.inaf.it