Mem. S.A.It. Vol. 80, 571

Memorie della

GRID-Launcher v.1.0

N. Deniskina[1], M. Brescia[2], S. Cavuoti[1], G. d'Angelo[1], O. Laurino[1], G. Longo[1,2]

[1] Department of Physical Sciences University Federico II, Naples, via Cinthia 6 80131 Naples, Italy, e-mail: natalia.deniskina@unina.it
[2] INAF - Osservatorio Astronomico di Capodimonte, via Moiariello 16, 80131, Napoli, Italy

Abstract. GRID-launcher-1.0 was built within the VO-Tech framework, as a software interface between the UK-ASTROGRID and a generic GRID infrastructures in order to allow any ASTROGRID user to launch on the GRID computing intensive tasks from the ASTROGRID Workbench or Desktop. Even though of general application, so far the Grid-Launcher has been tested on a few selected softwares (VONeural-MLP, VONeural-SVM, Sextractor and SWARP) and on the SCOPE-GRID.

Key words. distributed computing; virtual observatory

1. Introduction

The main goal of the International Virtual Observatory URL.2 (2000) infrastructure is to provide the community at large with an easy and user friendly access to astronomical data, software tools, and computing facilities. While the first two parts of the task, namely the federation and fusion of heterogeneous data archives and the implementation of flexible data reduction and data analysis tools have been widely addressed and, at least in their fundamental aspects, solved, the possibility to access large distributed computing facilities to perform computing intensive tasks has not yet been satisfactorily answered. The main reasons for this delay being mainly two: the need to match the GRID security requirements and the lack of homogeneity in the definitions of the storage space. The first issue can be easily explained: most users of a specific Virtual Organization (VO) do not possess the personal certificates which are requested to access the GRID or, even when they do have a personal certificate, the computing GRID which they need does not recognize their own certification authority. The second issue arises instead from the fact that the data usually reside locally or in a remote storage space which is not seen by the selected GRID as a storage element (SE).

The lack of an user friendly access to the GRID is a main obstacles against the use of computing intensive data mining methods and tools such as, for instance, Support Vector Machines Cavuoti (2008) or Probabilistic Principal Surfaces D'Abrusco et al. (2008) on massive data sets (MDS). Therefore, in the framework of the VONeural project, which aims at implementing a package of data mining routines capable to work in a distributed computing environment and on large data sets of high dimensionality, we have implemented

Send offprint requests to: N. Deniskina and tested a general purpose interface between the UK-ASTROGRID URL.1 (2000) and the GRID which solves part of the above quoted problems. We wish to stress that the followed approach is quite general and that GRID-Launcher can be easily adapted to other Virtual Organizations and to any GRID. For testing we used the GRID-SCOPE URL.3 (2000) which is part of the recently funded GRID infrastructure for SOuthern Italy.

2. GRID-launcher

As already mentioned, an ASTROGRID user is recognized by the UAS (User Authentication Service) of ASTROGRID but, often does not possess a personal GRID certificate. As it is the case for most virtual organizations, the ASTROGRID and GRID authentication procedures are incompatible due to a number of factors: a) different CAs (certification authority); b) the lack of compatibility between the ASTROGRID and GRID certificates and PKI policy. The EGEE-II boards are currently discussing how to allow the use of specific services by means of application certificates (or "robot" certificates) but, at the time GRID-Launcher 1.0 was developed this option had not yet be introduced and the first version discussed here needs therefore to be regarded as a test version of a more flexible tool (GRID-Launcher 2.0) which will be made available as soon as the new standards will be introduced.

In other words, the version 1.0 relies on a personal GRID certificate (signed by the INFN-GRID CA) which is recognized by the SCOPE-GRID used for testing. In the figures 1-3, we give a schematic representation of the way GRID-Launcher works: (Fig.1)

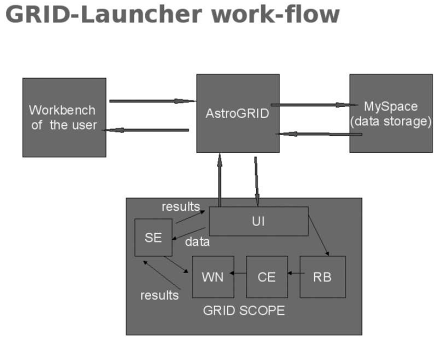

Fig. 1. Workflow. UI: user interface; RB: resource broker; SE: storage element; CE: computing element; WN: working node.

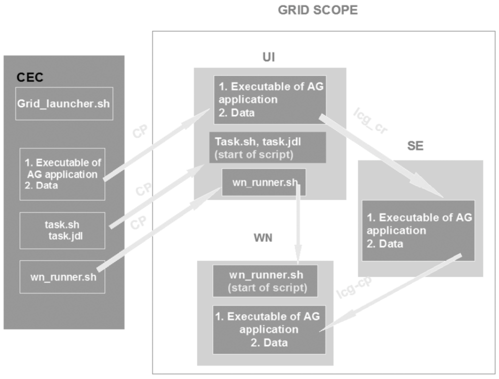

Fig. 2. The workflow from ASTROGRID to the WN of the GRID. The task.sh, wn_runner.sh and task.jdl - are automatically generated by the GRID launcher.

1. Grid-launcher:
   - takes the user input from the User Interface of ASTROGRID (Workbench);
   - collects all files, tabs and programs needed;
   - generates automatically three scripts:task.sh, task.jdl and wn_runner.sh to be executed on the GRID;
   - wraps them in an archive and sends it to the Scope-GRID UI. (The authentication on the GRID is done by public keys exchange).
2. The GRID User Interface (UI) receives the data and the JDL program, unpacks them, copies the data to the SE, copies wn_runner.sh to the WN's, starts task.sh and task.jdl. (Fig. 2)
3. wn_runner.sh starts on the WNs, takes the data from SE, starts the executable of the application, and puts the results on the SE. The GRID generates automatically two output files task.err and task.out and sends them to the UI using the Output SandBox.

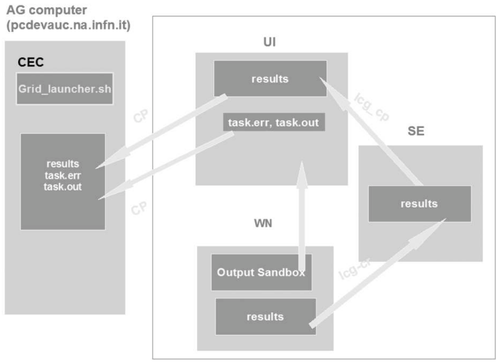

Fig. 3. The workflow from the WN to ASTROGRID. The task.err and task.out files are automatically generated by the GRID.

4. GRID-launcher periodically checks the status of job and when it ends, it moves the results from the UI to the ASTROGRID machine. (Fig. 3)
5. GRID-launcher receives the data archive, unpacks it and puts the results into the storage of ASTROGRID (VO-Space).

SO far, GRID-Launcher has been implemented and tested on:

– VO-Neural-MLP: a Multi Layer Perceptron Neural Network used for supervised clustering and classification Skordovski (2008); Brescia et al. (2008);
– VO-Neural-SVM: Support Vector Machines for supervised clustering Cavuoti (2008); Brescia et al. (2008);
– Sextractor: for the extraction of object-catalogues from astronomical images Bertin & Arnouts, (1996),
– SWarp: an application that resamples and coadds FITS images using any arbitrary astrometric projection defined in the WCS standardURL.4 (2000).

3. Future developments

In the version 2.0, we plan:

– to insert a ”switch” which will check whether the user has a GRID certificate (and in this case the task will start on the GRID of the user), or has not (in this case the program will be launched as a service on the SCOPE-GRID);
– to render the program capable to launch any program (and not only applications registered in the ASTROGRID) matching some trivial specifications;
– to offer the possibility to use GRID-Launcher as a subtask inside the ASTROGRID workflows and, finally,
– to unify the VO Space with the SE of the GRID.

Additional information about ”GRID-launcher” can be found at URL.5 (2000).

References

Bertin E. & Arnouts S., 1996, A& AS, 117, 393
Brescia M., et al., 2008, these proceedings.
S. Cavuoti, Laurea Thesis, 2008, University Federico II in Napoli
D'Abrusco R., Longo G., Walton N., 2008, astro-ph/0805.0156v1
B. Skordovski, Laurea Thesis, 2008, University Federico II in Napoli
URL.1: `http://www.astrogrid.uk/`
URL.2: `http://www.ivoa.org/`
URL.3: `http://scope.unina.it/`
URL.4: SWarp User manual , E. Bertin at: `http://terapix.iap.fr/rubrique.php?id_rubrique=49`
URL.5: `http://voneural.na.infn.it/`