

# GWpilot: a personal pilot system

*Thursday, 20 September 2012 11:30 (30 minutes)*

## Description of the work

By using pilot-job systems, Grid execution performance can be improved. This is so because it is possible to somehow reserve the available resources by means of submitting regular Grid jobs (the pilots) that lately allocate the application tasks for doing the real calculation. Nevertheless, the current implemented systems usually lack some features such as user-sharing and easy-to-installing capabilities or standard interfaces that could prevent their deployment.

GWpilot is a new developed general purpose framework system that overcomes the limitations or drawbacks derived from the aforementioned approaches. In addition to the common functionalities already implemented in other systems to overcome remote queues, correctly fit tasks to pilots or discard bad resources, it can also abandon a passive role in order to effectively coordinate the pilot-task matchmaking with advanced scheduling techniques such as pre-allocation, reservation, data-allocation awareness, etc. This is done by a more flexible, extendable and tuneable methodology based on embedding the pilot system in the GridWay metascheduler.

As consequence, the framework is suitable in both user and VO levels since it is compatible with standards such as DRMAA, OGSA-BES or JSDL and it permits a fine troubleshooting, this is, any user belonging to any scientific discipline or association can use it. Therefore, GWpilot can be used in any kind of HTC calculation since legacy applications can be easily adapted to these interfaces. Moreover, it enables a multi-level scheduling ecosystem that allows application-oriented systems (workflow managers) or self-schedulers to profit from the customized characterisation of resources supplied by pilots.

As use cases, several official EGI applications are employed to demonstrate the suitability of GWpilot as well as the performance gain obtained.

## Link for further information

<http://www.ciemat.es/portal.do?IDR=343&TR=C>

<http://www.gridway.org/>

## Wider impact of this work

GWpilot system is a framework specially devoted to single users and institutions who are interested in performing massive High Throughput Computing calculations. It can be used in principle with any Grid application and/or VO, so it is ideal for users who in principle do not count on ad-hoc systems. In this way, it is expected a high impact in the scientific community due to it offers standardised interfaces that allow a straightforward execution of previously gridified applications. It also provides an easy and standalone installation that facilitates its quickly deployment and use.

Last, in addition to the common advanced scheduling features already present in other pilot systems, GWpilot accomplishes the performance requirements of users, institutions and even VO levels.

## Printable Summary

Grid overheads can be overcome by pilot-job systems. Huge collaborations rely on ad-hoc ones that fit their need, but a researcher working on his own cannot rely on a general purpose application. This work is devoted to this kind of standalone users, either inexperienced or highly skilled ones. Some frameworks offer pilot-job advantages to conventional users, but they usually lack some features such as user-sharing, ease of installation or they suffer from standard Grid interfaces that could prevent their deployment. GWpilot is a newly developed general purpose framework based on GridWay that offers functionalities already implemented in other pilot systems to overcome remote queues, correctly fit tasks to pilots or discard bad resources, but also

coordinates the pilot-task matchmaking with advanced scheduling techniques. The system offers CLI tools and Grid standard interfaces that make it suitable for running legacy applications or being coupled to other systems such as workflow managers

**Primary author:** Mr RUBIO-MONTERO, Antonio J. (CIEMAT)

**Co-authors:** HUEDO, Eduardo (UCM); MAYO-GARCIA, Rafael (CIEMAT)

**Presenters:** Mr RUBIO-MONTERO, Antonio J. (CIEMAT); MAYO-GARCIA, Rafael (CIEMAT)

**Session Classification:** Resource Infrastructure Services

**Track Classification:** Resource Infrastructure services (Peter Solagna: track leader)