Extending gLite VOs with volunteer and institutional BOINC-based desktop grids to execute parameter sweep applications

Peter Kacsuk
MTA SZTAKI

Current Grid systems can be divided into two main categories: service grids (SG) and desktop grids (DG). Service grids (like gLite-based grids) are typically organized from managed clusters and provide a 24/7 service for a large number of users who can submit their applications into the grid. The SG middleware is quite complex and hence relatively few managed clusters take the responsibility of providing grid services. As a result the number of CPUs in a Virtual Organization (VO) of SGs is moderate typically in the range of 1.000-10.000.

BOINC-based volunteer desktop grids are collecting large number of volunteer desktop machines to exploit their spare cycles. These desktops have no SLA requirement, their client middleware code is extremely simple and hence typical number of volunteer desktops in BOINC DGs is in the range of 10.000-1.000.000. However, their drawback is that they can execute only some very limited number of pre-registered applications, typically compute-intensive bag-of-task applications. Interestingly, the vast majority of applications executed in gLite VOs are parameter sweep applications that also belong to the class of bag-of-task applications and hence BOINC-based desktop grids could be used to support these VOs.

Further on BOINC can also be used to build institutional DGs where the PCs of a university or a research institute are connected into a local DG system. Building such an institutional DG system is very economical since it does not require purchasing managed clusters or supercomputers, it simply utilizes the existing PCs and the size of such a DG system could easily be in the range of 1.000-10.000 machines that is comparable with the typical VO size.

A natural idea would be to extend the existing VOs with volunteer and institutional DG systems. In this way the size of the VOs could be significantly enhanced which is particularly important in the case of parameter sweep applications where large number of jobs run within a single application. This was the goal of the EDGeS project that was running between January 2008 and March 2010. EDGeS has developed the 3G Bridge (Generic Grid-Grid Bridge) technology that enables the extension of gLite VOs with BOINC and XtremWeb DGs. The concept of 3G Bridge is so generic that it was successfully applied later in the EELA-2 project to extend gLite VOs with OurGrid P2P desktop grid. Based on the 3G Bridge technology EDGeS has created a production infrastructure where any gLite VO can be extended with volunteer and institutional DG systems. EDGeS has also ported 12 gLite applications from various scientific areas to the EDGeS infrastructure and created an application repository from where gLite users can access these applications.

Built on the success of EDGeS a follow-up project, called as EDGI (European Desktop Grid Initiative), was launched in June 2010. The objectives of EDGI include the extension of the EDGeS infrastructure with ARC and Unicore service grid support and to enable the execution of even data-intensive applications. Further on to provide QoS for the integrated infrastructure DG systems of the EDGI infrastructure will be extended with Cloud resources when required.

All these experiences of EDGeS and the new achievements of EDGI will be explained in detail in the talk. The main focus of the talk is the explanation how the gLite VOs can be extended with volunteer and institutional DG systems and how parameter sweep applications can be ported to such extended gLite VOs. Finally, the talk will show the way how end-users can access and exploit the extended infrastructure.