COMBINING DESKTOP AND SERVICE GRIDS TO SUPPORT E-SCIENTISTS TO RUN SIMULATIONS

G. Testyanszky[1], T. Kiss[1], T. Kukla[1], S. Winter[1], J. Kovacs[2], Z. Farkas[2], P. Kacsuk[1,2]

[1]Centre for Parallel Computing, School of Electronics and Computer Science, University of Westminster, London, United Kingdom
[2]Computer and Automation Research Institute, Hungarian Science, Budapest, Hungary
G.Z.Terstyanszky@westminster.ac.uk

The paper will describe the EDGI infrastructure which can provide resources for compute- and data-intensive applications. The FP7 European Desktop Grid Initiative (EDGI) project created a production infrastructure that integrates clouds, (built on OpenNebula and OpenStack), desktop Grids (based on BOINC and XWHEP) and service Grids (built on ARC, gLite and Unicore) into a single platform to support European Grid Initiative (EGI) user communities. The EDGI infrastructure provides access to resources through the EDGI Science Gateway which contains the EDGI Portal and the EDGI Application Repository (EDGI AR). The EDGI Portal [1] enables submitting and monitoring applications, and retrieving and displaying results. The security models of desktop and service Grids are significantly different. While service Grids trust the users and identify them by unique certificates, desktop Grid systems trust the applications. As a result, only trusted and validated applications can run on DG systems. If users of a service Grid infrastructure want to utilise desktop Grid resources, the applications should be priori validated and pre-deployed on the supporting desktop Grid. This requires uploading validated applications into a repository (EDGI AR) and make them publicly available. Users can browse and search the repository in order to find and download applications they want to execute. The infrastructure contains two bridges: SG→DG bridge and DG→Cloud bridge. The bridges manage differences between job description and job submission of desktop and service grids.

The paper will outline how to port and run docking simulations on the EDGI infrastructure. In the first two phases of the simulation the receptor and ligands molecules are selected to create a homology model and then the model is optimised by energy minimization and a consecutive short molecular dynamics (MD) simulation using AMBER. After defining the target region for the docking, sets of ligand molecules are established and together with the receptor molecule are submitted as inputs to the Autodock Vina program which is used as a docking program. It provides a measure of the docking quality expressed as a so called “docking energy”, which is the measure for the ability of the algorithm to fit the ligand into the binding pocket of the receptor. The docking and docking evaluation phases are the compute- and data intensive phases of the simulation. The Autodock Vina was ported to desktop Grid and deployed on the EDGI infrastructure: on the Carmen cluster and on the Westminster Campus Grid. The Campus Grid is a BOINC [2] based private desktop Grid. It consists of approximately 1600 PCs installed in student labs. E-scientist can run the docking simulation through the EDGI Portal. The EDGI infrastructure enables significant speed-up of the simulation. The paper will present outputs of the docking simulation.

[1] Z. Farkas and P. Kacsuk: P-GRADE Portal: A generic workflow system to support user communities, Future Generation Computer Systems, Volume 27 Issue 5, May, 2011, doi>10.1016/j.future.2010.12.001

[2] Anderson, D. P. 2004, BOINC: A System for Public-Resource Computing and Storage. 5th IEEE/ACM International Workshop on Grid Computing, November 8, 2004, Pittsburgh, USA.