Virt&I-Comm.5.2014.17

AN EGI-XSEDE PROJECT FOR DEVELOPPING A HIGH THROUGHPUT HIGH PERFORMANCE UNIVERSAL MOLECULAR SIMULATOR AS A CMMST COMMUNITY SERVICE

Antonio Laganà, Department of Chemistry, Biology and Biotechnologies
University of Perugia

EXECUTIVE SUMMARY

The CMMST (Chemistry, Molecular, Materials Science and Technologies) VRC (Virtual Research Community) of EGI (European Grid Infrastructure) intends to establish a collaborative/competitive work model for designing and implementing a joint EGI-XEDE Synergistic High-throughput&high-performance Universal Molecular Simulator (SHUMS). SHUMS is meant to be offered as a service to the community members to the end of developing higher level of complexity multi-scale scientific and technological innovative applications. To this end a set of requests for "startup allocations" of compute resources are submitted to XSEDE as preparatory to a subsequent development of the final product.

PREAMBLE

The collaboration established between XSEDE and the European Grid Infrastructure (EGI) has evolved in recent times into focusing on the issue of supporting VRCs in their effort in developing workflows, portals, tools and user support initiatives. Through the participation to 2014 EGI Community Forum (Helsinki) of May and to XSEDE 14 (Atlanta) of July, in fact, important steps necessary to foster a proactive user participation have been envisaged. Among them efficient authentication and authorization procedures, robust competence centres support, procedures for code selection and performance evaluation, user requirements adapted quality of service parameters, scheduling, data transfer, frameworks and cloud computing techniques.

THE PROJECT

In this spirit the CMMST (Chemistry, Molecular, Materials Science and Technologies) VRC (Virtual Research Community) of EGI has proposed to XSEDE the implementation of a collaborative/competitive work model in a joint endeavour with researchers of the same area in US. The proposal will aim at assembling a Synergistic High-throughput&high-performance Universal Molecular Simulator (SHUMS) as a generalized Computational Engine for a vast range of higher complexity multi-scale applications (entailing accurate ab initio treatments, various dynamical methods and detailed statistical approaches) and its components on a networked platform of XSEDE and EGI computers. The project will rely on EGI-XSEDE bridged computations and on the assistance of related support teams, tools and Competence Centres as already agreed during the dedicated BOF held at XSEDE 14 and Workshop held at CF14.

THE SYNERGISTIC MODEL

The networked platform will consist of HTC, HPC machines and specialized HW interoperating through the middleware adopted by XEDE and EGI, accessible via specialized portals and gateways, whose packages, programs and data are chained in shared workflows. De facto data format standards and stable versions of common packages will be made available by the members of the community and third bodies. The most appropriate compute elements will be selected using the functionalities of GRIF (a grid framework allowing the ranking of provided resources on the ground of the provided Quality of Service (QoS) according to user requirements), the contributions to the community activities made by active users will be rewarded through the award of credits based on QoS and Quality of User (QoU) parameters.

THE PROPOSED EGI-XSEDE COMPUTE TIME ALLOCATIONS

Basic EGI compute time allocations are opportunistic. On top of them, the CMMST VRC can provide its users with preferential allocations based on internal criteria (eg the already mentioned quality parameters).
XSEDE allocations, instead, are bound to requests linked to a US Principal Investigator (PI) and a EU based coPI. For the purpose of the project tentative requests (preliminarly discussed in the already mentioned EGI and XSEDE events) are:

- Full ab initio electronic structure and quantum nuclei dynamics treatment of few body chemical reactions as virtual crossed beams experiments;
- High level ab initio calculations for electronic structure for large molecules structural properties;
- Accurate algebraic formulations of large molecules ab initio potential energy surfaces;

ISSN: 2279-8773

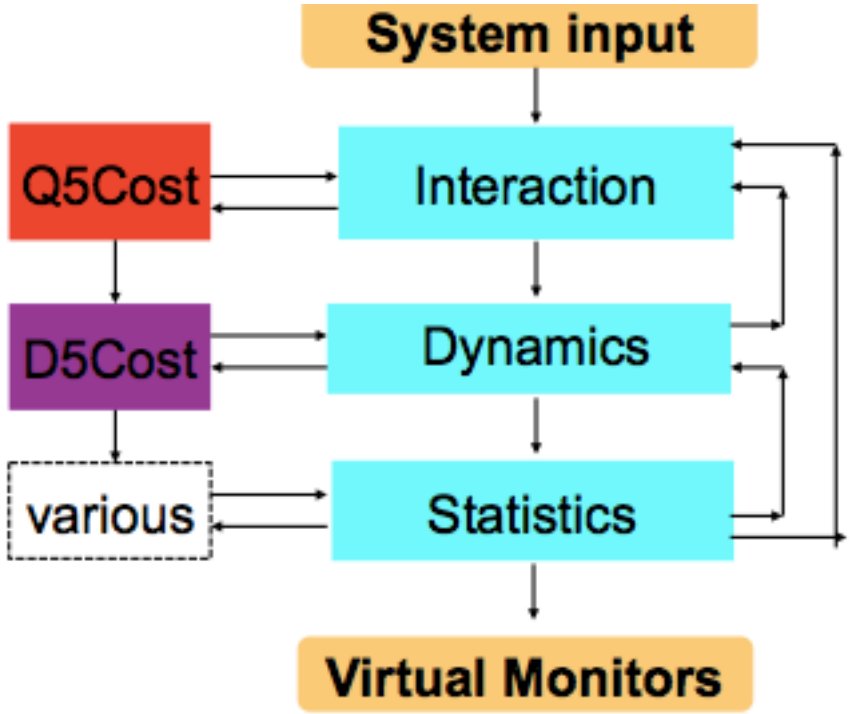
- Direct classical and semiclassical dynamics for large system;
For these applications startup allocation requests will be made on some XSEDE platforms (up to 200,000 CPU-hours (or Service Units, SUs)).
Parallel to the above mentioned computational activities, additional work will be necessary on the network, middleware, data handling, workflows, etc. Support on this will be provided by EGI and XSEDE. However specific work is proposed to be carried out by specialize teams
Quality of service Monitoring using GriF functionalities – (C. Manuali, University of Perugia, IT)
Adaptation of workflows and portals to CMMST needs (P. Kacsuk, MZTAKI, Budapest, HU)
Large data handling and data formats for molecular sciences (G. Graciani, University of Barcelona, ES)

THE PARTNER SUPPORT TO THE SIMULATOR

Parallel to the above mentioned computational activities, additional work will be necessary on the network, middleware, data handling, workflows, etc. Support on this will be provided by EGI and XSEDE.
User support of EGI (G. Sipos and NGI specific structures), XSED (M. Rynge, S. Marlu, S. Senelievici)
ECSS Direct coordination S. Senelievici
Possible support of the CMMST Competence Centre
Moreover, additional specific work will be carried out by the following specialized teams
- Quality of service Monitoring using GriF functionalities – (C. Manuali, University of Perugia, IT)
- Adaptation of workflows and portals to CMMST needs (P. Kacksuc, MZTAKI, Budapest, HU)
- Large data handling and data formats for molecular sciences (G. Graciani, Univesity of Barcelona, ES)XR

THE ROADMAP

From now on, frequent contacts with the CMMST members will be activated in order to define the details of the proposals.
By the fall the first batch of proposals for preparatory allocations will be submitted and technical problems with the different installations, networking, libraries and middleware will be tackled first.
The work on the related project is expected to last about one year.
Following the completion of the preparatory work, where appropriate, requests for regular allocations will be submitted.



ISSN: 2279-8773