Virt&I-Comm.5.2014.14

Requirements of the Chemistry, Molecular & Materials Sciences and Technologies community for evaluating the quality of a service

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

The present is a list of requirements that the Chemistry, Molecular & Materials Sciences and Technologies (CMMST) has developed to the end of implementing a synergistic community model [1] allowing selection of resources and a sharing of software and data. The synergistic computing model, in fact, meets the dreams of the most recent generations of computational scientists because it allows to run highly complex applications not only by having flat access to heterogeneous platforms but also by exploiting the advanced features of some workflow structured computing procedures offering an a la carte combination of the applications made available to the members of the community by the European Grid Infrastructure (EGI), CMMST and other providers (including external ones).

In other words, the proposed synergistic model encompasses much more than the simple feature of aggregating a large amount of computing resources and users for massive distributed computing that is the key mission of EGI. It targets, in fact, the more radical objective of changing the model of using and managing research oriented computing resources by fostering:

- a synergy among complementary types of programs (or sections of programs) and expertise (including the monitoring of experiments and the access to data bases) [2];
- a coupled distributed and parallel (HTC and HPC) use of computational platforms [3];
- a service oriented organization with a rewarded active participation of the users in the different activities of the community [4].

This enhances the possibility of building higher complexity applications sustainability by stimulating the use of tools facilitating collaborative activities and the introduction of a metrics suited for ranking the interoperability of the offered software. Within the synergistic model users can, in fact, provide and maintain stable versions of their own programs and applications software properly structured in terms of (input and output) data formats so as to be chained in efficient simulators like the Grid Empowered Molecular one (GEMS) [5].

It is, also, important here to point out that the synergistic model grounds also the rewarding of the resource providers on the ground of quality evaluation of the service provided. Within a synergistic model, in fact, in addition to finding computing resources one may wish to select at run time the most appropriate one and to access as well High Performance Computers from a High Throughput one.

In a synergistic model it is also equally important to evaluate the users' contribution to the community activities. This calls the users to develop tools facilitating collaborative activities and the offering of stable versions of their programs and applications and to adopt as well standard formats for data. software properly structured in terms of (input and output) data formats so as to be chained in GEMS and offered as services (in return for credits as we shall detail later) to other users as a kind of shared library of applications.

All this enhances the community sustainability by calling both for the development of tools facilitating collaborative activities and for the introduction of a metrics suited for ranking services.

ISSN: 2279-8773

Within the synergistic model users can, in fact, provide and maintain stable versions of their own programs and applications software properly structured in terms of (input and output) data formats so as to be chained in GEMS and offered as services (in return for credits as we shall detail later) to other users as a kind of shared library of applications.

To this end COMPCHEM is going to join its efforts with those of other partners either belonging to the CMMST community (like GAUSSIAN, MoSGrid, Scalalife and CHEM.VO.IBERGRID.EU) or external to it like the Division of Computational Chemistry of EUCHEMS (http://www.euchems.eu/divisions/computational-chemistry.html) and the VEC standing committee of the ECTN (http://ectn-assoc.cpe.fr/) Association.

This grounds also the VRC purpose of competing for acquiring, as a general policy, computing resources from resource providers for the community (to be added to the ones available to the VRC members from other sources). These resources will be used by the community for supporting fundamental activities (like basic research and algorithms development but also production runs), sustaining collaborative projects, pursuing special innovative targets, etc. whose assignments to the members of the community will be regulated by the credit system. For this reason the following requirements have been set:

For each job or subjob:

1. Username owning the job
2. Grid service name (wrapper)
3. Grid site name used
4. Job Type (e.g. Single, Parameter Study or Workflow)
5. Submission date
6. Ending date
7. Program name (if custom)
8. Input provided (if any)
9. Results address (e.g. Storage Element)
10. Computing Element address (e.g. queue name used for run)
11. Computing Element Exit status (e.g. 0 -> done with success; 1 -> CE error)
12. Cpu time
13. Wall time
14. Memory consumed
15. Additional job info

ISSN: 2279-8773

For each Grid Service (wrapper):

16. Name

17. Description

18. Grid site name location

19. Maintainer / Manager

20. Number of Functions (e.g. # of Web Services composing the Grid Service)

21. Number of Errors

22. Elapsed (averaged) time for satisfying (initiating) a call

23. Elapsed (averaged) TTR (Time-to-Repair) for repairing an error

24. QoS (calculated by us)

25. Cost (assigned by us)

26. Additional service info

For each User:

27. Number of runs corresponding to those applications already made available by the Framework;

28. Number of runs corresponding to new programs provided by the user (if any);

29. Number of results generated from 27)

30. Number of results generated from 28)

31. Number of results accessed from 27)

32. Number of results accessed from 28)

33. Average cpu time elapsed for runs generating results

34. Average cpu time elapsed for runs having results accessed

35. Average wall time elapsed for runs generating results

36. Average wall time elapsed for runs having results accessed

37. Average memory consumed for runs generating results

38. Average memory consumed for runs having results accessed

ISSN: 2279-8773

39. Number of positive feedbacks

40. Number of neutral feedbacks

41. Number of negative feedbacks

42. QoU (calculated by us)

43. Credits (assigned by us)

44. Additional user info

For each Computing Element (CE):

45. Queue name

46. Ranking evaluation (position)

47. Last used (date) or new

48. Performance

49. Average Latency

50. Additional ranking info (e.g. Round robin strategy flag)

[1] A. Laganà, A. Costantini, A virtual research community for synergistic computing in Chemistry, Molecular and Materials Science and Technologies, Virt&l-comm 3 (2013); A. Laganà, A. Riganelli, O. Gervasi, On the structuring of the computational chemistry virtual organization COMPCHEM, Lecture Notes in Computer Science (ISSN:0302-9743) 3980, 665-674 (2006)
[2] A. Laganà, E. Garcia, A. Paladini, P. Casavecchia, N. Balucani, The last mile of molecular reaction dynamics virtual experiments: the case of the OH (N=1-10) + CO (j=0-3) → H + CO_2 reaction, Faraday Discussion of Chem. Soc. 157, 415 - 436 (2012)
[3] C. Manuali, S. Rampino, A. Laganà, GRIF: A Grid Framework for a Web Service Approach to Reactive Scattering, Comp. Phys. Comm. 181, 1179-1185 (2010); C. Manuali, A. Laganà, GRIF: A New Collaborative Framework for a Web Service Approach to Grid Empowered Calculations, Future Generation of Computer Systems, 27(3), 315-318 (2011)
[4] C. Manuali, A. Laganà, A Grid Credit System Empowering Virtual Research Communities Sustainability, Lecture Notes Computer Science 6784, 397-411 (2011)
[5] A. Costantini, O. Gervasi, C. Manuali, N. Faginas Lago, S. Rampino, A. Laganà, COMPCHEM: progress towards GEMS a Grid Empowered Molecular Simulator and beyond, Journal of Grid Computing 8(4), 571-586 (2010)

ISSN: 2279-8773