**EGI-InSPIRE**

TOWARDS A CMMST VRC virtual team project report

Proposal for the setup of the CMMST VRC in EGI that has been prepared in the related EGI-InSPIRE VT project

[**https://wiki.egi.eu/wiki/Towards\_a\_CMMST\_VRC**](https://wiki.egi.eu/wiki/Towards_a_CMMST_VRC)

|  |  |
| --- | --- |
| Date: | 18/06/2013 |
| Document Status: | **Draft** |
| Dissemination Level: | **Internal** |
| Document Link: | **https://documents.egi.eu/document/xxxx** |

|  |
| --- |
| **Abstract**The present document formulates a proposal for the setup of the CMMST Virtual Research Community (VRC) in EGI. The proposal has been elaborated as a deliverable of the EGI-InSPIRE Virtual Team (VT) project ‘Towards a Chemistry, Molecular & Materials Science and Technology (CMMST) VRC’ promoted by the members of the Virtual Organization (VO) COMPCHEM. The proposal is tailored to the specific nature of the CMMST community that is a typical aggregate of several independent academic research laboratories coupled by common scientific interests (not necessarily bound to a common experimental infrastructure) and aims at establishing a highly collaborative distributed environment.The proposal relies on the convergence of COMPCHEM with other EGI VOs (like GAUSSIAN and CHEM.VO.IBERGRID.EU), University Departmental structures (like Chemistry of the Wroclaw Technology University (PL), Chemistry of the Texas Technology University (US), Chemistry of the University of Perugia (IT), Chemistry of the University of the Basque Country in Vitoria (ES), Physical Chemistry of the University of Barcelona (ES)) and Research Centres (like FORTH (GR), Research Centre for Natural Sciences of the Academy of Science (HU), Grid Computing Competence Center (GC3 - CH)……..). It is also grounded on the support of some National Grid Infrastructures (like IGI, PL-Grid and IBERGRID) and of some resource providers (like CINECA, CYFRONET, CESCA, CSCS, …..). The proposal involves also the European Chemistry Thematic Network Association (ECTNA, http://ectn-assoc.cpe.fr/) and the Division of Computational Chemistry (DCC, http://www.euchems.eu/divisions/computational-chemistry.html) of EUCHEMS. The VRC intends to provide its members with the advantages that the virtual community status implies. Namely: meet user requirements for an efficient access and use of high throughput and high performance computing resources, aimed at composing higher level complexity applications through the sharing of hardware and applications and the integration of a quality evaluation of the work done for the community. Such proposal, detailed in the present document, enumerates the technical and non technical aspects related to the building of a virtual community and analyses technology, structure and organization evolution (at national and European level) of distributed computing (Grid and Cloud) infrastructures enabling the set up of the CMMST VRC. For this purpose, the document analyses both how to exploit the tools offered by EGI to the end of building a distributed computing community and how to engage the members of the community in the development of innovative instruments. Such instruments will enable the selection of the resources (from personal systems to supercomputers) and services (from number crunching to massive data handling on heterogeneous platforms). They will make it possible, as well, to reward the work done for the community through a credit acquisition/redemption system. This is meant to activate a sustainable quality-based enhancement of grid economy for communities that could be taken as a model for other communities making the use of distributed computing resources a real change in the way science is carried out and in the way research activities impact social life. |

Copyright notice

Copyright © Members of the EGI-InSPIRE Collaboration, 2010-2014. See [www.egi.eu](http://www.egi.eu) for details of the EGI-InSPIRE project and the collaboration. EGI-InSPIRE (“European Grid Initiative: Integrated Sustainable Pan-European Infrastructure for Researchers in Europe”) is a project co-funded by the European Commission as an Integrated Infrastructure Initiative within the 7th Framework Programme. EGI-InSPIRE began in May 2010 and will run for 4 years. This work is licensed under the Creative Commons Attribution-Noncommercial 3.0 License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc/3.0/> or send a letter to Creative Commons, 171 Second Street, Suite 300, San Francisco, California, 94105, and USA. The work must be attributed by attaching the following reference to the copied elements: “Copyright © Members of the EGI-InSPIRE Collaboration, 2010-2014. See [www.egi.eu](http://www.egi.eu) for details of the EGI-InSPIRE project and the collaboration”. Using this document in a way and/or for purposes not foreseen in the license, requires the prior written permission of the copyright holders. The information contained in this document represents the views of the copyright holders as of the date such views are published.

1. Document Log

|  |  |  |  |
| --- | --- | --- | --- |
| **Version** | **Date** | **Comment** | **Author/Partner** |
| 0.0 | 18/06/2013 | Draft0 | A. Costantini |
| 0.1 | 19/06/2013 | Revision | A. Costantini |
| 1.0 | 26/06/2013 | Draft1 | A. Laganà |
| 1.1 | 30/06/2013 | Revision | A. Costantini |
| 1.2 | 01/07/2013 | Revision | D. Cesini |
| 1.3 | 05/07/2013 | Revision | M. Ceotto |
| 1.4 | 12/07/2013 | Revision | A. Costantini |
| 2.0 | 30/07/2013 | Draft2 | A. Costantini |
| 2.1 | 02/08/2013 | Revision | A. Costantini |
| 3.0 | 08/08/2013 | Draft3 | A. Laganà |
| 3.1 | 19/08/2013 | Revision | A. Costantini |
| 3.2 | 27/08/2013 | Revision | A. Laganà |
| 3.3 | 27/08/2013 | Revision | A. Costantini |
| 3.4 | 27/08/2013 | Revision | A. Costantini |
| 3.5 | 29/08/2013 | Revision | S. Farantos |

1. ACKNOWLEDGMENT

None

1. Application area

This document is a public report produced by the members of the “Towards a CMMST VT“ EGI Virtual Team project, run under the EGI-InSPIRE NA2 virtual team framework. Further information is available at <https://wiki.egi.eu/wiki/Towards_a_CMMST_VRC>.

1. Terminology

A complete project glossary is provided at the following page: <http://www.egi.eu/about/glossary/>.

TABLE OF CONTENTS

1 Introduction 5

2 MilEstones’ outcomes 5

2.1 M1 outcomes 5

2.2 M2-M5 outcomes 6

3 TOOLS OFFERED BY EGI TO THE CMMST COMMUNITY 6

4 the CMMST-VRC proposal 6

4.1 Current scenario of computing resources allocation and usage 6

4.2 The synergistic Grid model for the CMMST VRC 7

4.3 The adoption of a credit system in the CMMST VRC 8

4.4 Coordination and Management 10

4.4.1 Management bodies for the VRC 10

4.4.2 TASK A1: User support and training 10

4.4.3 TASK A2: Services and operations 11

4.4.4 TASK A3: Requirement gathering 11

4.4.5 TASK A4: Communication and Dissemination 11

4.4.6 Sustainability 12

5 Technologies and services offered by the CMMST-VRC through EGI 12

5.1 Provided Computational Chemistry applications 12

5.2 Provided Grid tools and services 15

6 ONGOING RESEARCH AND DEVELOPMENT 17

6.1 Further applications 17

6.2 Further tools and services 18

6.3 Bridging HTC and HPC 18

6.4 Data formats 18

7 Conclusions 19

Appendix A – EGI/EUDAT/PRACE ACTIVITY 21

Appendix B – The ChEmistry MOLECULAR AND MATERIALS SCIENCES AND TECHNOLOGIES VT membership 22

Annex 23

References 24

# Introduction

Virtual Research Communities (VRCs) are groups of like-minded individuals and researchers’ clusters organised by discipline or computational model. A VRC can establish a support relationship, formalised through a Memorandum of Understanding (MoU), with the European Grid Infrastructure collaboration (EGI, http://www.egi.eu/). EGI VRCs (http://www.egi.eu/community/vrcs/) typically have an established presence in their field and represent well-defined scientific research communities. Multi-national scientific communities can draw many benefits from having a VRC partnership with EGI. For example, they can benefit from the resources and support that are available within the National Grid Initiatives (the main stakeholders of EGI.eu), they can benefit from the workshops and forums organised by EGI, they can receive support on resolving specific technical issues with EGI services, and they become involved in the user-focussed evolution of EGI’s production infrastructure.

For this reason, the COMPCHEM (https://www3.compchem.unipg.it/compchem/) Virtual Organization (VO) with the help of the GAUSSIAN and CHEM.VO.IBERGRID.EU VOs, belonging to the Chemistry, Molecular & Materials Science and Technology (CMMST) community, by recognising the advantages that membership as a VRC within EGI will bring has assembled the EGI Virtual Team (VT) project [1] that has taken the first step towards the setup of the CMMST VRC. The VRC status will help the CMMST community to satisfy the requirements of its members concerning the access and use the computing resources that are federated in EGI. The VT will take care of documenting:

* the structure that the proposed VRC should have to represent the CMMST community in EGI;
* the technologies, resources and services which already exist within EGI and could be used to satisfy the requirements of the CMMST VRC;
* the tools that need to be developed or brought into EGI and then integrated with the production infrastructure so the VRC members can efficiently manage and use resources from EGI;
* the applications that need to be imported in order to assemble higher level of complexity CMMST simulations.

# MilEstones’ outcomes

The objectives addressed by the project [1] are:

* develop a plan aimed at assembling a VRC out of the already existing CMMST oriented EGI Virtual Organizations (VOs);
* exploit related applications, tools as well other resources and services that NGIs and EGI projects provide;
* identify tools, services and resources that the VRC needs to develop or bring into EGI in order to operate as a sustainable entity for the CMMST scientific community;
* elaborate a proposal for the technical, organisational and funding aspects.

## M1 outcomes

M1 outcomes were focused onto the preparation of the Investigation Planning Document (IDP) containing a list of technical and non-technical topics (like tools used, access to computing resources, structure of the VRC, attraction and training of new users, etc…) investigated by the VT. The IDP presents an analysis of the consistency of the available application patrimony and an indication of how attracting more users to CMMST (in addition to the already existing ones of COMPCHEM and GAUSSIAN VOs) and in particular the members of the Division of Computational Chemistry (DCC) of EUCHEMS and of the European Chemistry Thematic Network Association (ECTNA). Each topic has been allocated either to an individual or to a team leader from the VT.

## M2-M5 outcomes

M2-M5 outcomes are listed in the present document that provides a preliminary analysis of the means and the actions to be undertaken in order to create an infrastructure of integrated local sites and services connected in a European network whose goal is to offer to the users a stable access to distributed heterogeneous computing platforms. The proposed strategy is founded on the following pillars:

1. selection among the available resources of those suitable for the applications considered;
2. replacement of the in-use model based on individual acquisition of computing resources via “grant award” with a “community acquisition and sharing” of both hardware and software applications;
3. utilization of the resources regulated by an appropriate credit system (credits are awarded in return of activities spent on behalf of the community, work for projects as well as against financial support).

and has gained support of the following University Departmental structures (like Chemistry of the Wroclaw Technology University (PL), Chemistry of the Texas Technology University (US), Chemistry of the University of Perugia (IT), Chemistry of the University of the Basque Country in Vitoria (ES), Physical Chemistry of the University of Barcelona (ES)) and Research Centres (like FORTH (GR), Research Centre for Natural Sciences of the Academy of Science (HU), Grid Computing Competence Center (GC3 - CH)……..). It is also grounded on the support of some National Grid Infrastructures (like IGI, PL-Grid and IBERGRID) and of some resource providers (like CINECA, CYFRONET, CESCA, CSCS, …..).

# TOOLS OFFERED BY EGI TO THE CMMST COMMUNITY

--- GERGELY ---

# the CMMST-VRC proposal

## Current scenario of computing resources allocation and usage

The evolution of computing technologies has led both to the growth of remotely accessible high-performance platforms (High Performance Computing, HPC) and, on the other hand to the growth of local (departments, research groups, institutions, etc.) clustered computational platforms and networks. These two types of platforms adopt two different models of resources allocation:

- the first (grant model) assigns, as already mentioned, the resources as a result of an ex ante evaluation of a proposal following a specific call for user projects;

- the second (opportunistic model) meets the requirements of registered users as they come without a specific selection of the merits for their request.

The grant model is in general adopted by large scale supercomputing facilities (as well as by the medium-large computer centres) in which specialized staff is devoted to the maintenance of HW and SW. The opportunistic model, instead, is in general adopted either by research groups or departments to run the machines and update the software of their computational members (with the help of some permanent or temporary technical staff).

In large scale computing facilities (as are the computer centres networked in PRACE, http://www.prace-ri.eu/) the users can rely on computational resources, software and skills that evolve according to the current technological development and to time assignments (grants) born out of a centralized vision of computing needs and strategies. In such model the user develops a rather passive attitude (with respect to the hardware and to the evolution of technology, networking and policy) and is forced to either adapt his/her application to the new computing paradigms or to adopt an already adapted one.

In the case of local computing facilities, instead, the users (which are either responsible or co-responsible for managing the available resources and related environments) can customize the platform and tune it to fit their needs and those of the services they wish to use/offer. In this case, however, hardware, technology, networking and policy evolution is much slower and the user group takes care of updating (usually rather partially) skills and computational environment (with the risk of underutilizing the computing resources and applications due to a lack of expertise).

## The synergistic Grid model for the CMMST VRC

Although widely adopted both models are alone inadequate to meet the present requirements of the scientific communities. As an alternative a synergistic innovative model [2] of allocating resources was developed within a series of European projects and collaborations (DATATAG (http://datatag.web.cern.ch/datatag/), EGEE-I-II-III (http://www.egee.eu), WLCG (http://wlcg.web.cern.ch/), etc.) mainly under the pressure of the research needs of High Energy Physics. Such synergistic model, often called High Throughput Computing (HTC) is based on the aggregation of a large number of geographically dispersed CPUs and users connected over the public network and is managed through the use of appropriate middleware and tools, has made it possible to execute distributed programs on a large number (over hundreds of thousands) of processors reaching very high performances for applications composed by decoupled or loosely coupled tasks.

After the first impulse given by the High Energy Physics community the model was extended to other disciplines. After all, for academic research groups (not belonging to a single Institution or tightly bound to the management of a shared experimental infrastructure) the synergistic model is the best approach to access the necessary volume of computing resources. Moreover, the synergistic model encompasses much more than the simple task of aggregating a large amount of computing resources and users for massive distributed computing. It can target, in fact, the more radical objective of changing the model of using and managing computing resources by allowing:

- an on demand allocation of the available computing time among registered users

- a synergy among different types of programs and expertise (including experiment and monitoring)

- a service oriented organization with a (Grid economy) stimulated proactive participation of the users

- a coupled distributed and parallel (HTC-HPC) computing methodology.

After all, within the EGI-Inspire European project, significant effort has been spent to make the Grid model undergo a significant increase in participating communities, networked resources, software and skills (otherwise confined in the local sites) availability as well as inter-disciplinary collaboration enhancement. Moreover, in order to meet the peculiarities of the different research areas and strengthen their role the users of the Grid have been gathered in VOs (http://www.egi.eu/community/vos/) which are now encouraged to gather together into VRCs (http://www.egi.eu/community/vrcs). For this reason the specific commitment of the COMPCHEM VO in the CMMST VRC project is to join its forces with those of other actors of the community (especially those of similar VOs like GAUSSIAN and CHEM.VO.IBERGRID.EU and to encourage other VOs to join as well). In particular through DCC this opportunity will be offered to the researchers of ECTNA (http://ectn-assoc.cpe.fr/) and of the Division of the Computational Chemistry (http://www.euchems.eu/divisions/computational-chemistry.html) of EUCHEMS. In order to better achieve this, quarterly teleconferences for planning and management will be held and one or two face-to-face meetings a year will be held during main Grid events like conferences or forums. Such meetings will also focus on technical problems of the VRC including operations, services, support to user, communication and dissemination.

In particular the Grid model to be implemented for CMMST should realize the dream of several computational scientists (also of other disciplines) of being able to carry out simulations of realistic systems and reproduce their observable properties starting from first principles (the so called virtual experiment as discussed in Ref. [3]). This is, indeed, what the Grid Empowered Molecular Simulator (GEMS) schema has been designed for [4]. In this way, by making use of appropriate grid tools (portals, workflows, frameworks, etc.) and blocks of applications and data, VRC scientists should be able to compose such simulations a la carte from components and applications made available to the members of the community by other members (or even from external providers) and run them by having flat access to the most suited computing resources (within the limits of availability).

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), sustaining collaborative projects, pursuing special innovative targets, etc.) and will be assigned to the members of the community in return for the gained credits.

## The adoption of a credit system in the CMMST VRC

A key aspect of the synergistic nature of the proposed CMMST VRC is, in fact, the possibility of monitoring the activities of the members to the end of evaluating the work they do on behalf of the community and the services they are offered by the other members of the community. In other words all the activities and services offered to the community are to be identified, recognised and rewarded.

The solid pillars on which such endeavour relies are

1. the commitment of the members of the community to offer the products of their work (new algorithms and applications, new validated data, etc.) as a service to the other members
2. the exploitation of the functionalities embodied in two tools being developed in COMCPHEM.

Such tools are aimed at evaluating the quality of both services (QoS) and users (QoU) by means of a framework performing active and passive filtering (GriF[5]) on Grid data and at awarding credits [6] based on the evaluation of monitored activities.

Because of this, the VRC can offer to its members the clear advantage of allowing them to carry out their computational campaigns (especially when the calculations are so complex to be unfeasible on the ground of only their own competences and their own limited platforms) in return for their commitment to carry out community related duties. Obviously, one can also choose to join the community at an entry level (passive user, first lower layer as discussed in ref. [7]) member and either run the codes made available for demo purpose or his/her own software for personal use. The reason for having established this entry layer (usually meant to be temporary) is to offer an opportunity to the newcomers to check their real willingness to operate on a DCI platform. Already at this level, in fact, it becomes soon apparent that it is necessary to bear certain competences to exploit the advantages of the Grid. As a result, at this point one has choice of acquiring such competences and enter the second layer (active user) able to restructure the code so as to take advantage of distributing computing and usefully relate their codes to the ones of other users.

Obviously, one can still remain passive user after the expiration date of the trial period though he/she would have to contribute to the VRC sustainability using different means like for example being a paying user. On the contrary, the involvement reached by being active user may call the VRC members to an upper layer of involvement. As a matter of fact, opening the code implemented on the Grid to a shared usage by the other members of the VRC is the next membership level (passive software provider, third layer) that implies the validation of a stable version of a code, the assemblage of all the necessary GUIs (Grid User Interfaces) and the adoption of proper (even if only de facto) standard data formats for use by other researchers possibly in conjunction with other codes. When software providers guarantee also additional basic services (like software upgrade, maintenance as well as user support (active software provider)) they become active software providers (fourth layer).

The commitment to confer to the Grid additional hardware (especially for those suites of codes which need special devices) after a negotiation with the Operations coordinator, on behalf of the VRC Management Committee (MC), about the relevance of such a commitment to the strategic choices is also welcome. This is, indeed, the fifth layer (passive resource provider). It applies also to hardware providers the possibility of providing basic services and user support. In this case they become active resource provider (sixth layer).

Obviously, the going through all these layers by conferring both software and hardware together with related services will take place gradually along a certain period of time (and not all the members of the VRC will be willing or admitted to do so). On the contrary some members or group of users or resource providers would like to take this as a mission (for example devote to development work on behalf of the VRC also unshared resources) and reach the status of “VRC stakeholder” (seventh layer).

The belonging to a given layer is periodically revised on the ground of credits acquired and multiple layer status for different products and services may coexist within a given member. For example, the status of stakeholder does not exclude that of paying customer or paid supplier for certain items. This becomes particularly appropriate when the VRC wins a bid or becomes a funded project holder.

Crucial to such dynamical geometry of the VRC is the adoption of “terms of exchange credits” (toecs or simply credits [8]) through which all activities of the members can be rewarded. The mechanism through which credits are assigned and redeemed are established and regulated by the governing body of the VRC. In general credits are assigned to reward the work that the VRC declares as useful to its sustainability and can be then redeemed through a privileged allocation of the resources of the VRC, discounts in the use of commercial tools and/or other services. They may be even redeemed, in some clearly regulated cases, as cash to be used for research activities or hiring people to work on them.

## Coordination and Management

The formal acceptance of the Grid model and the establishing of VRC in EGI is the signature of a Memorandum of Understanding (MoU) between EGI and the VRC. The CMMST community has already expressed its wish to be formally part of EGI through an expression of interest formulated by the COMPCHEM VO. Yet, due to the fact that only the signature of a MOU provides a means of documenting the relationship that EGI has with the community as well as the activities and objectives of such a collaborative relationship (a MoU, is used when both parties do not want to pursue a contract that is legally binding while it still clarifies the relationships, responsibilities and communication channels between two or more parties that may share services, clients, and resources) we give below a sketch of the coordination and management structure that in the first instance will be provided by the members of the participating VOs for the time period needed for the first consolidation of the VRC.

### Management bodies for the VRC

The Management body of the VRC is the Board of Stakeholder (BoS) that is made by 5 representatives of the stakeholder VOs (one representative for each fraction of 20% of the total number of members of the VRC belonging to the VO), 3 representatives (one each) for the three main stakeholder resource providers and 2 representatives of the most contributing (in terms of nationality of the members) stakeholder National Grid Infrastructures.

In this spirit the VRC management body is here defined:

- VRC leader: is the person representing the VRC and is the MoU contact point and is elected by the BoS. Such position will be taken for the first three years by the COMPCHEM coordinator Antonio Laganà).

- VRC deputy leader: is the person elected by the BoS acting on the behalf of the VRC leader and appointed by him/her. Such position will be taken for the first three years by …

- VRC activity coordinators: are the persons appointed by the BoS representing the activity of the VRC in strong coordination with EGI and the participant VOs and acting as experts in the assigned roles. In the first instance they will be agreed by the VOs participating to the CMMST VRC

- User support and training Coordinator: is the person taking care of managing user support and training activities (A1). Such position will be taken for the first three years by …

- Operations Coordinator: is the person taking care of managing the services and operation activities (A2). Such position will be taken for the first three years by …

- Technical Coordinator: is the person taking care of managing requirements gathering activities (A3). Such position will be taken for the first three years by …

- Communication and Dissemination Coordinator: is the person taking care of managing communication and dissemination activities (A4). Such position will be taken for the first three years by …

### TASK A1: User support and training

VRC representatives: Name Surname – VRC Role – Affiliation

 ………………. - ……………. - ………

 Antonio Laganà – VRC leader – UNIPG

- provide information about estimates on the size of the VRC and possibly its by country;

-provide dedicated user support activities for code building, adapting and porting on multi platform computing models and storage;

-provide tutorials aimed at enhancing the use of services made available to the community;

-provide examples, training material, details of specialist applications, documentation and presentations to be made accessible to members of the scientific community.

### TASK A2: Services and operations

VRC representatives: Name Surname – VRC Role – Affiliation

 ………………. - ……………. - ………

 Antonio Laganà – VRC leader – UNIPG

-provide robust, well designed, user centric services to scientific user communities;

-provide local and global operational services as needed to support international users and operations;

-drive the adoption of standards within tools and applications production;

-provide compliance with the operation interfaces needed to ensure seamless and interoperable access to networked (Gird and Cloud) resources;

-utilize a credit system to encourage CMMST users to cooperate in developing higher level of complexity applications;

produce VRC-specific availability statistics make use of the EGI availability computation system.

### TASK A3: Requirement gathering

VRC representatives: Name Surname – VRC Role – Affiliation

 ………………. - ……………. - ………

 Antonio Laganà – VRC Leader – UNIPG

-define and monitor Service Level Agreement (SLA) for third-level support on incidents and requests;

-participate to EGI security police team to contribute to the development to the security police fabric of the infrastructure;

-participate in the Operations Management Board to contribute to the EGI operations agenda;

-subscribe to a mandatory set of EGI policies, procedures and Operation Level Agreements (OLA)s.

-design a roadmap for the harmonization of the participating VO and research groups structure

### TASK A4: Communication and Dissemination

VRC representatives: Name Surname – VRC Role – Affiliation

 ………………. - ……………. - ………

 Antonio Laganà – VRC Leader – UNIPG

-define communication strategies aimed at attracting more CMMST users into a common endeavour offering the possibility of assembling higher level of complexity applications and services;

-establishing contact points for communication channels and publications;

-disseminate results of the collaboration 1) by planning joint sessions at EGI.eu and CMMST events; 2) by informing EGI.eu and each VRC members of any scientific/academic publications published/issued related to the VRC activities.

### Sustainability

The members of the VRC are requested to be proactive in providing either their own work or attract financial resources specifically for the development of the VRC. As to contributing to the VRC by providing work this may occur under the form of participation to the management of the Grid, to the development of tools and services, etc. As to attracting financial resources VRC members should elaborate joint applications for funding, research projects and even develop within the VRC commercial services. However, the most important contribution to the sustainability that is requested to the stakeholders is a high dynamism in research and in the transfer of its outcomes into innovation and developments (R&D). This means that, ideally, all members of the VRC should excell in basic and applied research and be ready to provide work to be rewarded in terms of credit.

With this assumptions, the VRC management body collaborates with EGI and in order to assure the VRC sustainability and in particular by tackling the following joint activities:

-contribute to enable the vision of providing European scientists and international collaborations for sustainable distributed computing services to support their work;

-exchange ideas and collaborate on the definition of sustainability models;

-collaborate in business relationships development;

-develop a community economy also via a design and implementation of a credit system.

# Technologies and services offered by the CMMST-VRC through EGI

In this section a prospect of technical and non technical aspects related to the applications and services provided by the VRC to the CMMST users are given. Applications and services are the common patrimony of the VRC and are also the ground on which the VRC supports its members and grows in quality and becomes able to develop higher complexity activities.

In particular, attention has been focused to the software packages ported into the Grid environment and currently in use by the CMMST members and to the capabilities of the existing EGI and Grid-related tools.

## Provided Computational Chemistry applications

Leveraging on the Molecular and Materials sciences research competences of the Chemistry, Molecular & Materials Science and Technology (CMMST) members, the technology of distributed computing has been exploited to the end of tackling multi-scale problems starting from a rigorous treatment of atomic and molecular structures and processes with the aim of enabling real-like simulations of a wide range of problems. These real-like problems concern research, innovation and development in various fields ranging from chemical engineering to biochemistry, chemometrics, omic-sciences, forensic chemistry, medicinal chemistry, food chemistry, energy production and storage, new materials, space technologies, etc. To this end, in a joint endeavour with other international Research Laboratories, various programs devoted to

* A) the ab initio calculation of the electronic structure of molecular systems and the assemblage of ab initio based potential energy surfaces,
* B) the integration of quantum and/or classical equations of motion plus the handling of

the additional statistical and higher scale treatments necessary to work out the value of measurable quantities

have been developed and/or gathered together to form a Grid Empowered Molecular Simulator (GEMS) [4] devoted to the assemblage of realistic innovation after being ported on the computing Grid.

A set of in-house developed and third body programs (a full list of chemistry and solid state physics software is given in ref. [9]) has been implemented on the computing Grid to be used as a cooperative computational engine for real world high-level of complexity applications (aimed at supporting research, innovation and development in various fields ranging from pure science to innovative technologies) of strong economic and social impact through a service oriented approach. Such cooperative endeavour is based on

1. the combined expertise and effort of experts in the field of the molecular science
2. the adoption (with appropriate adaptations) of high level ICT instruments.

A list of the CMMST applications already (or being) ported on the GEMS Grid environment and receiving dedicated user support is shown in Table 1. The following applications are registered in the EGI AppDB (see Section 5.2 for details) and related contact points are provided. The VRC is committed to take care of its usability, evolution and maintenance.

A) AB INITIO ELECTRONIC STRUCTURE SOFTWARE

Several ab initio electronic structure packages (Hartree-Fock (HF), post Hartree-Fock, Density Functional Theory (DFT)), either open source or commercial, based on the Born Oppenheimer approximation, as those given in ref. [9], are suitable for insertion into the INTERACTION block of GEMS. Most of them are large packages and often contain several separate programs developed over many years.

The packages ported to the Grid as part of the INTERACTION block are briefly described in the followings and information on literature reference and VRC Reference Members (RM) are given:

*GAMESS-US* [10] RM(K Baldridge, S. Maffioletti)

GAMESS-US is an ab initio electronic structure molecular quantum chemistry package that calculates potential energy values for moderately large molecular systems using direct and in parallel techniques on appropriate hardware.

*GAUSSIAN* [11] RM(GAUSSIAN VO – M. Sterzel) An ab initio electronic structure molecular GAUSSIAN is a quantum chemistry package to calculate potential energy values for moderately large molecular ems performed using direct techniques, or in parallel on appropriate hardware.

*NB-MCTDH* [12] RM(D. Skouteris)

NB-MCTDH is a multiconfigurational time-dependent Hartree (MCTDH) program for calculating bound states of a generalized N-Body system (including Non Born Oppenheimer cases)

B) MOLECULAR MOTIONS AND OBSERVABLES

The specificity of CMMST is that of adding to the just mentioned popular Ab initio electronic structure packages several other ones designed or available for the other 2 blocks of GEMS (DYNAMICS and OBSERVABLES). The codes listed in the followings are all tailored to deal either with quantum or semiclassical and classical equations of molecular motions as part of the DYNAMICS block of GEMS. Most of the codes, however, include also routines for the calculation of quantities related to the OBSERVABLES block of GEMS (that could be also taken as separate programs).

The packages ported to the Grid as part of these two blocks are briefly described in the followings and information on literature reference and VRC Reference Members (RM) are given:

B1) QUANTUM PACKAGES

*ABC* [13] RM (D. Skouteris, L. Pacifici, E. Garcia, S. Rampino)

ABC is a time independent atom diatom quantum reactive scattering program using a coupled-channel hyperspherical coordinate method to solve the Schrodinger equation for the motion of the three nuclei (A, B, and C) on a single Born-Oppenheimer potential energy surface.

*FLUSS-MCTDH* [14,15] RM(F. Huarte)

FLUSS-MCTDH is a pair o programs carrying out a multiconfigurational time-dependent Hartree (MCTDH) calculation of thermally averaged quantum dynamics properties of multidimensional systems based on a modified Lanczos iterative diagonalization of the thermal flux operator.

*RWAVEPR* [16] RM(D. Skouteris, L. Pacifici)

RWAVEPR is a time dependent atom diatom quantum reactive scattering program using Jacobi coordinates to integrate rigorously the three-dimensional time-dependent Schroedinger equation by propagating wave packets.

*SC-IVR* [17] RM(M. Ceotto)

SC-IVR is a Semi-classical (SC) initial value representation (IVR) program based on the outcome of a classical trajectory code used to calculate the thermal rate coefficients for the gas-phase reactions.

*DIFFREALWAVE* [18] RM (E. Garcia)

DIFFREALWAVE is a parallel real wavepacket code for the quantum mechanical calculation of reactive state-to-state differential cross sections in atom plus diatom collisions.

B2) SEMICLASSICAL AND CLASSICAL PACKAGES

In the case of classical and semiclassical dynamics, contrary to what occurs for the quantum codes, use is made of some general purpose packages. In particular, the most used packages by the CMMST community are:

*VENUS96* [19] RM(W. Hase)

VENUS96 is a program developed and maintained by W.L.Hase (QCPE-671). It calculates the trajectory for two reactants (atoms or molecules) by integrating the Hamilton equation in cartesian coordinates. VENUS96 is linked to the Semi-Classical (SC) Initial Value Representation (IVR) program SC-IVR [13] in order to perform semiclassical calculations of the thermal rate coefficient for gasphase reactions.

*DL\_POLY* [20] *RM* (M. Albertì)

DL\_POLY is the most used code in the CMMST VRC for the integration of the classical equation of motion of Molecular dynamics calculations. It is a general purpose package of subroutines, programs and data designed to facilitate Molecular Dynamics simulations. DL POLY is continually developed at Daresbury Laboratory by W. Smith and I.T. Todorov under the auspices of the British EPSRC and NERC in support of CCP5. It can be used to simulate a wide variety of molecular systems including simple liquids, ionic liquids and solids, small polar and non-polar molecular systems, bio- and synthetic polymers, ionic polymers and glasses solutions, simple metals and alloys.

*GROMACS*[21] RM (IGI)

GROMACS is a versatile package to perform molecular dynamics by integrating the Newtonian equations of motion for systems with hundreds to millions of particles. It has been used in a large number of case studies and a complete workflow aimed at exploit the interoperability between a local cluster platform (HPC capable) and a Grid platform (mainly HTC capable) has been developed. In the workflow, the coupling among the various jobs is taken care using a link (a semaphore) that de\_nes the dependency job chain.

*NAMD* [22] (IGI)

NAMD is a parallel molecular dynamics code designed for high-performance simulation of large biomolecular systems and it has been used to study the behaviour of a lipidic bi-layer in a water box. Ported on the Grid environment by using OpenMPI parallel libraries, a Direct Aciclic Graph (DAG) has been implemented to run the code in a semi-automatic way and facilitate the used on carrying out his/her calculations.

## Provided Grid tools and services

Thanks to the activities of CMMST members, several CMMST applications have been ported in the Grid environment where dedicated user support activities are provided. However, the use of different applications and technologies often requires advanced skills not always available to the user. For such reason a set of tools have been developed both to govern complex ensembles of data, models and programs of an increasing number of applications and to offer a unified user friendly way of composing related tools. Among them are those provided by EGI and EGI members in connection with initiatives and projects aimed at supporting the various communities operating on the European Grid for their work. In particular, in the followings we list those which have been adopted to be offered to the community members with the obvious proviso that they are in a further development stage:

*AppDB* [23] RM (EGI)

The EGI Applications Database (AppDB, (http*://*appdb.egi.eu/)) is a central service that stores and provides to the public, information about: tailor-made software tools for scientists and developers to use, the programmers and the scientists who developed them, and the publications derived from the registered software items. All software filled in the AppDB is ready to be used on the European Grid Infrastructure.

*gUSE WS-PGRADE* [24] RM(MTA-SZTAKI)

Is an open source DCI gateway framework that enables users the convenient and easy access to Grid and Cloud infrastructures and supports the development, visualization, configuration and submission of distributed applications executed on the computational resources. gUSE WS-PGRADE has the structure of a workflow enabling application developers to define their investigations by means of a graphical environment and to generate out of the user-specified description the Grid scripts and commands allowing the execution of the various computational tasks on the distributed computing platforms. For the above mentioned reasons gUSE WS-PGRADE has been used as a workflow engine in many works by developing a set of reusable application-specific workflows;

*GC3PIE* [25] *RM* (GC3 - Maffioletti)

Is a library of Python classes for running large job campaigns (high throughput) on diverse batch-oriented execution environments (such ARC) providing facilities for implementing command-line driver scripts in the form of Python object classes whose behavior can be customized by overriding specified object methods. The adoption of the framework GC3Pie in many computational campaigns allowed the users to define event-related dependencies between different applications and execute them simultaneously on a large-scale distributed computing infrastructure. Thanks to its the programmatic approach to workflows, the entire execution schema is assembled at runtime and steps can be added and removed dynamically as the program progresses, adapting to the outcome of individual computations;

*GriF* [26] RM (COMPCHEM VO - Manuali)

Is a Service Oriented Architecture (SOA) Collaborative Framework designed to facilitate the use of the DCIs by non specialists aimed at optimizing the selection of different computing elements for running single and parameter study applications. It consists on a set of java modules aimed at submit and monitor jobs in different computing platforms such as Grid or HPC. The framework has been equipped with a Quality of User (QoU) - Quality of Service (QoS) evaluation module called GCRES based on the monitoring of the user activities. Both GriF and GCRES have been developed within the activities of the CMMST community;

*IGI Portal* [27] *RM* ( IGI)

Is a science gateway developed within the activities of the Italian Grid Initiative (IGI), which is operating the Italian Grid infrastructure. Already used by other communities \cite{ansys}, the web-based portal provides to the final user several important services such as job submission, workflow definition and data management for both Grid and Cloud environments. Moreover, it implements a robust Authentication and Authorization (AA) mechanisms to provide a secure computational environment and, at the same time, hiding the complexity of the X.509 digital certificates on which most of the DCI resources rely;

*InSilico Lab* [28] RM (CYFRONET)

Is an application portal designed to support in silico experiments by easily running computational chemistry software on grids. Unlike manual job submission or grid portals, InSilicoLab enables to run computations on grids without technical knowledge of how to operate it;

*MyGAMESS* [29] RM (GC3 - Maffioletti)

Is a Framework for Integration of Chemistry applications with the Swiss Grid Portal aimed at providing dynamic access to applications services. Users will profit for example by a considerable reduction of time-consuming data format conversions and/or other code restructuring, once the capabilities of different applications are integrated;

*G-LOREP* [30] RM (COMPCHEM – Tasso; Varella)

(Grid LOs Repository) is a manager of a federation of distributed repositories of research based learning objects. It relies on a central database and allows registered users to create and/or download learning contents for research based education activities. G-LOREP makes also the content automatically shareable among the federation servers and enables the creation of dependency management software attachment as well as fault tolerant submission of the simulations from any federated site.

# ONGOING RESEARCH AND DEVELOPMENT

In the process of implementing the proposed synergistic model, the CMMST community is carrying out within various projects a significant extent of research, design, implementation and validation work. In this process the pillar activity of CMMST is, indeed, the search for new methods and technologies. In this respect both the field of Molecular and Materials Sciences Technologies and the field of related computational aspects are subject of extended research and production of new ideas. Therefore new approaches to the description of atomic and molecular properties together with their gas and condensed phase will be investigated and considered for development of algorithms and programs to be structured as services to the community and to the Society at large.

## Further applications

Other applications planned to be ported in the Grid environment but not yet integrated in the AppDB are:

*CRYSTAL* [31](COMPCHEM, F. Filomia)

CRYSTAL is a commercial quantum chemistry ab initio program written by V.R. Saunders, R. Dovesi, C. Roetti, R. Orlando, C.M. Zicovich-Wilson, N.M. Harrison, K. Doll, B. Civalleri, I.J. Bush, Ph. DArco, and M. Lunell within a collaboration of the Theoretical Chemistry Group of the University of Torino (IT) and the Computational Materials Science Group at the Daresbury Laboratory (UK) with other researchers. CRYSTAL is recognized within the computational chemistry community as a powerful tool for carrying out solid state simulations useful for scientific and technological applications;

*QUANTUM ESPRESSO* [32] RM (S. Cozzini)

QUANTM ESPRESSO is an integrated suite of Open-Source computer codes for electronic-structure calculations and materials modeling at the nanoscale based on density-functional theory, plane waves, and pseudopotentials;

*APH3D* [33] RM (G. Parker)

APH3D is a time independent atom diatom quantum reactive scattering program using a coupled-channel hyperspherical coordinate method to solve the Schrodinger equation for the motion of the three nuclei (A, B, and C) on a single Born-Oppenheimer potential energy surface;

*GFIT3C and GFIT4C* [34] RM (L. Pacifici)

*GFIT3C and* GFIT4C are routines devoted to the fit of the Potential Energy surface of three and four body molecular systems using the formulation of M. Paniagua based on Many Body expansions and polynomials of exponentials time internuclear distances;

*POMULT* [35, 36] RM (S. Farantos)

ROMULT is a fortran program supporting the spectroscopy studies of large molecules by locating periodic orbits and equilibrium points in Hamiltonian systems based on 2-point boundary value solvers which use multiple shooting algorithms;

*FORWCONV* [37] RM (N. Balucani)

FORWCONV is a routine devoted to the forward convolution of the Newton Diagrams of a distribution of Newton diagrams to simulate the composition of a gas phase molecular beam.

## Further tools and services

Among other technologies suited for being used by the CMMST community we list here some applications whose usability is currently under investigation by selected computational chemistry experts belonging to the community with the support of the related technology developers:

*KEPLER* [38]

KEPLER is an open source, scientific workflow application. Using Kepler graphical interfaces and components, scientists can create and execute scientific workflows which allow accessing scientific data and executing complex analyses on them. Kepler has been recently adopted in a recent work [] where a prototypical workflow has been designed and tested for the calculation of the time independent quantum probabilities of the H+H2 benchmark reaction. For this purpose a set of quantum mechanical codes, both belonging to the set of the aforementioned computational applications, have been used;

*SHIVA SIMULATION PLATFORM* [39]

SHIVA is a project providing a multi-systems workflow execution platform and an interoperability solution. The platform currently supports seven workflow systems: Askalon, Kepler, LONI pipeline, MOTEUR, Taverna, Triana and WS-PGRADE.

## Bridging HTC and HPC

As already mentioned, in order to effectively bridge HPC and HTC machines to the end of overcoming those highly unsatisfactory situations in which neither HPC nor HTC are completely fit to meet the requests of complex CMMST applications at present use is made of a SSH procedure. Such situation is going to be overcome by further developing the resource selection function of GriF. Moreover, also on the resource providers side (and not only on the user one) there are good reasons for coordinating the use of HPC and HTC e-infrastructures to the end of interoperating large computational applications. This in fact allows an optimization of the usage of both HTC and HPC computing resources because it is not infrequent the case in which a user utilizes HPC platforms not as such but as a bunch of loosely coupled processors underutilizing their fast dedicated network. At the same time HTC users may utilize massively distributed HTC platforms to solve tightly coupled computational tasks ending up by wasting a large amount of time in transferring data on the net. A coordination of the two types of platforms to interoperate via a single workflow (or workflow of workflows) and properly manage the various components on the most appropriate hardware, would instead allow a clever composition of complex applications optimizing the use of the various computing resources and providing the users with the best level of performance.

## Data formats

An already mentioned advantage of the synergistic distributed model consists in the possibility for the user to choose the platform better fitted for retrieving/storing programs and data for his/her applications. This provides the CMMST members with the possibility of better combining different pieces of software when assembling complex realistic applications using GEMS [4] (or any other collaborative simulator) and undertake more ambitious research projects. Such possibility of building workflows of shared programs has, for example, stimulated the setting of proper (de facto) standards of data in quantum chemistry and quantum dynamics [40]. This has prompted, as already mentioned, the development of tools (like the framework GriF) enabling the redirecting of computer applications to run on the best suited sites (including HPC machines) and enhancing cooperative compute capabilities by opening the perspective of combining different complementary know how into single (higher level of complexity) realistic applications and applying as a community for more ambitious research grants. A fundamental necessary implication of all that is the further development of standards for the representation of data (sometimes huge sets) of different types within the CMMST area. Moreover the CMMST community wishes to collaborate in some other EGI activities as, for example, the EUDAT and PRACE pilot ones [41] aimed at bringing together infrastructures and user communities by focusing their interest in the data sharing facilities (see Appendix A).

Such activity is expected to provide a set of tools and recommendations that the CMMST community can use to integrate its CMMST workflows.

# Conclusions

In the present document the advantages that a VRC status would offer to the members of the CMMST community in terms of access and use of the computing resources federated in EGI are listed. The document depicts the present scenario of computational resources and their usage by analysing the characteristics of the grant and the opportunistic models as opposed to the synergistic one proposed for the CMMST community. The document enumerates also the technical and non technical aspects of such cooperative model based on the selection of the resources (from personal systems to supercomputers) and services (from number crunching to massive data handling on heterogeneous platforms) as well as on an advanced usage of QoS and QoU and on an introduction of a community economy based on a credit. On this ground, the CMMST VRC (as well as other communities which use distributed computing resources) will be able not only to share hardware and software that (thanks to the expertise offered the various groups) can be integrated to address complex problems and computing simulations of high social impact.

Table 1 – A list of packages and frameworks in-use by the CMMST Community over EGI Grid

|  |  |  |
| --- | --- | --- |
| **Application** | Description | **License** |
| ABC | Solve the Schrodinger equation for triatomic systems using the time independent quantum method | Academic |
| FLUSS-MCTDH | MultiConfigurational Time Dependent Hartree method Lanczos iterative diagonalization | Academic |
| NB-MCTDH | N-Body MultiConfigurational Time Dependent Hartree method  | Academic |
| VENUS96 | Quasi-classical dynamics of reactive collisions |  |
| DL\_POLY | Classical Molecular Dynamics |  |
| NAMD | Classical Molecular Dynamics | Academic |
| GAMESS-US | General Atomic and Molecular Electronic Structure Package | Academic |
| RWawePR | Time Dependent Method to Solve the quantum reactive Scattering equations for triatomic systems  | Academic |
| GROMACS | GROningen MAchine for Chemical Simulations | Academic |
| SCIVR | Semiclassical initial value representation method | Academic |
| **Framework** | Description |  |
| GriF | Grid Framework enabling efficient and user-friendly scientific massive calculations | Free |
| Gcres | Quality of Users (QoU), Quality of Services (QoS) evaluation Framework | Free |
| MyGAMESS | Front-end script for submitting multiple GAMESS-US jobs | Free |
| InSilicoLab | Application portal designed to support in silico experiments | Free |
| IGI Portal | science gateway providing a workflow environment and a Cloud access | Free |
|  |  |  |
| GC3PIE | Framework providing building block to build dynamic workflows | Free |

# Appendix A – EGI/EUDAT/PRACE ACTIVITY

The goal of this EGI activity is to bring together European infrastructures (EGI, EUDAT and PRACE) and user communities to discuss the requirements in sharing and using their data between different environments, which can be general infrastructures or domain specific facilities.

The main objectives of this activity are here highlighted

* identify common data access and transfer tools and protocols
* demonstrate real benefit for the involved communities
* identify technology and/or organisational gaps and suggest measures for improvements

On behalf of such objectives, pilot activities were set up as a result of the EGI/EUDAT/PRACE workshop on data management [40] and address the user communities' requirements by defining and setting up test environments using existing technologies like GridFTP, UNICORE, Globus Online, iRODS, EMI data management services, etc.

Several pilot activities were defined to address specific but generic use cases of usage of different resource infrastructures:

* Pilot 1. & 2: Data sharing and uniform data access across e-infrastructures and community centres
* Pilot 3. ScalaLife and PRACE
* Pilot 4. Monitoring and resource discovery

# Appendix B – The ChEmistry MOLECULAR AND MATERIALS SCIENCES AND TECHNOLOGIES VT membership

(list all those involved in the team and doing the work)

|  |  |  |  |
| --- | --- | --- | --- |
| Country Code: | Name: | e-mail: | Role: |
| IT | Antonio Lagana (UNIPG)  | lagana05@gmail.com | VT Leader |
| IT | Daniele Cesini (INFN)  | daniele.cesini@cnaf.infn.it |  |
| IT | Alessandro Costantini (INFN)  | alex.costantini@gmail.com |  |
| IT | Carlo Manuali (UNIPG) | carlo.manuali@gmail.com |  |
| IT | Michele Ceotto (UNIMI) | michele.ceotto@unimi.it |  |
| GR | Stavros C. Farantos, IESL/FORTH and Chemistry UoC  | farantos@iesl.forth.gr |  |
| GR | Stamatis Stamatiadis, IESL/FORTH and Materials UoC  | stamatis@materials.uoc.gr, |  |
| GR | Manos Giatromanolakis, IESL/FORTH  | gmanos@iesl.forth.gr |  |
| GR | Giannis Remediakis, Materials UoC  | remed@materials.uoc.gr |  |
| HU | George Lendvay (MTA TTK Research Center for Natural Sciences)  | lendvay@chemres.hu |  |
| HU | Anna Vikar (MTA TTK Research Center for Natural Sciences)  | vikaranna@hotmail.com |  |
| HU | Peter Szabo (MTA TTK Research Center for Natural Sciences)  |  |  |
| HU | Tibor Nagy (MTA TTK Research Center for Natural Sciences)  | nt2001@freemail.hu |  |
| HU | Akos Bencsura (MTA TTK Research Center for Natural Sciences) | bencsura@gmail.com |  |
| NL | Remco Havenith (RUG) | r.w.a.havenith@rug.nl |  |
| ES | Fermin Huarte Larrañaga (IQTCUB)  | fermin.huarte@gmail.com |  |
| ES | Amaia Saracibar (University of the Basque Country)  | amaia.saracibar@ehu.es |  |
| ES | Ernesto Garcia (University of the Basque Country)  | e.garcia@ehu.es |  |
| FR | Marco Verdicchio (CNRS)  | m.verdicchio@gmail.com |  |
| CR | Petr Hanousek (CESNET) | petr.hanousek@cesnet.cz |  |
| UK | Peter Oliver (STFC) |  |  |
| USA | Alan Sill (CERN)  | alan.sill@ttu.edu |  |
| USA | Bill Hase (TTU) | Bill.Hase@ttu.edu |  |
| CH | Sergio Maffioletti (UZH)  | sergio.maffioletti@gc3.uzh.ch |  |
| PL | Mariusz Sterzel (CYFRONET) | m.sterzel@cyf-kr.edu.pl |  |
| EGI | Elena Tamuliene (VU) | jelena.tamuliene@tfai.vu.lt |  |
| EGI | Gergely Sipos (UCST)  | gergely.sipos@egi.eu |  |
| EGI | Richard McLennan (UCST)  | richard.mclennan@egi.eu |  |
| EGI | Karolis Eigelis (UCST)  | karolis.eigelis@egi.eu |  |

# Annex

None

# References

[1] <https://wiki.egi.eu/w/images/5/5f/VT_CMMST_Proposal_v1.pdf>

[2] A. Laganà, A. Costantini, A virtual research community for synergistic computing in Chemistry, Molecular and Materials Science and Technologies, Virt&l-comm, 3 (2013)

[3] 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 + CO2 reaction, Faraday Discussion of Chem. Soc. 157, 415 - 436 (2012)

[4] 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)

[5] C. Manuali, S. Rampino, A. Lagana’, GRIF: A Grid Framework for a Web Service Approach to Reactive Scattering, Comp. Phys. Comm. 181, 1179-1185 (2010); C. Manuali, A. Lagana’ GRIF: A New Collaborative Framework for a Web Service Approach to Grid Empowered Calculations Future Generation of Computer Systems, 27(3), 315-318 (2011)

[6] C. Manuali, A. Laganà, A Grid Credit System Empowering Virtual Research Communities Sustainability, Lecture Notes Computer Science, 6784, 397-411 (2011)

[7] 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).

 [8] A. Laganà, S. Crocchianti, N. Faginas Lago, A. Riganelli, C. Manuali, S. Schanze, From Computer Assisted to Grid Empowered Teaching and Learning Activities in Higher Chemistry Education in Innovative Methods in Teaching and Learning Chemistry in Higher Education, I. Eilks and B. Byers Eds, RSC Publishing (2009) p. 153-190 ; ISBN 978-1-84755-958-6

[9] Quantum Chemistry packages list:

http://en.wikipedia.org/wiki/List\_of\_quantum\_chemistry\_and\_solid\_state\_physics\_software

[10] GAMESS-US see http://www.msg.ameslab.gov/gamess/ M.W.Schmidt, K.K.Baldridge, J.A.Boatz, S.T.Elbert, M.S.Gordon, J.H.Jensen, S.Koseki, N.Matsunaga, K.A.Nguyen, S.Su, T.L.Windus, M.Dupuis, J.A.Montgomery J. Comput. Chem., 14, 1347 (1993).

[11] <http://www.gaussian.com/>

[12] D. Skouteris, O. Gervasi, A. Laganà, Non-Born-Oppenheimer MCTDH calculations on the confined H2+ molecular ion Chem. Phys. Letters 500 (1-3), 144-148 (2010)

[13] Skouteris, D., Castillo, J.F., Manolopulos D.E.: ABC: a quantum reactive scattering program. Comp. Phys. Comm. 133 (2000) 128–135

[14] U. MANTHE Direct calculations of reaction rates in Reaction and Molecular Dynamics A. LAGANA' and A. RIGANELLI Eds, Springer Verlag BERLIN (ISBN 3-540-41202-6) (2000) p. 130

[15] Beck, M., Jakle, A., Worth, G., Meyer, H. D.: The multiconfiguration timedependent Hartree (MCTDH) method: a highly efficient algorithm for propagating wavepackets. Phys. Rep. 324 (2000) 1–05

 [16] D. Skouteris,  L. Pacifici, A. Lagana’, Time dependent wavepacket calculations for the N(4S) + N2(1Σ+g) system on a LEPS surface: inelastic and reactive probabilities, Mol. Phys. (ISSN:0026-8976) 102(21-22), 2237-2248 (2004).

[17] M. Ceotto, S. Atahan, S. Shin, First principles semiclassical initial value representation molecular dynamics, Phys. Chem. Chem. Phys. 11, 3861-3867 (2009)

[18] Hankel, M., Smith, Sean C., Gray, S. K. and Balint-Kurti, G. G. (2008) DIFFREALWAVE: A parallel real wavepacket code for the quantum mechanical calculation of reactive state-to-state differential cross sections in atom plus diatom collisions. Computer Physics Communications, 179 8: 569-578.

[19] venus96 Chemical Dynamics Software and Simulation System (CDSSIM System) https://cdssim.chem.ttu.edu/nav/htmlpages/licensemenu.jsp. Last seen May 2013

[20] Smith W, Forester T. R. (1996): DL POLY2: a general purpose parallel molecular dynamics simulation package. Journal of Molecular Graphics, Vol. 14 (3), 136-141

[21] Hess, B., Kutzner, C., van der Spoel, D., Lindahl, E. (2008) `GROMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation', J. Chem. Theor. Comp. Vol 4, pp.435{447

[22] James C. Phillips, Rosemary Braun, Wei Wang, James Gumbart, Emad Tajkhorshid, Elizabeth Villa, Christophe Chipot, Robert D. Skeel, Laxmikant Kale, and Klaus Schulten. Scalable molecular dynamics with NAMD. Journal of Computational Chemistry, 26:1781-1802, 2005.

[23] AppDB website: http://appdb.egi.eu/. Last seen May 2013

[24] Sipos, G., Kacsuk, P.: Multi-Grid, Multi-UserWorkflows in the P-GRADE Portal. Journal of Grid Computing, 3 (2005) 221-238

[25] GC3Pie website: http://code.google.com/p/gc3pie/. Last seen May 2013

[26] 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)

[27] IGI portal website: https://portal.italiangrid.it/. Last seen May 2013

[28] In Silico Lab website: <http://insilicolab.grid.cyfronet.pl/>. Last seen May 2013

[29] MyGAMESS project: http://www.switch.ch/aaa/apps/projects/htdocs/detail/UZH.7

[30] S. Tasso, S. Pallottelli, M. Ferroni, R. Bastianini, A. Laganà, Taxonomy management in a Federation of Distributed Repositories: a chemistry use case Lecture Notes Computer Science 7333, 358-370 (2012)

[31] R. Dovesi, R. Orlando, B. Civalleri, C. Roetti, V. R. Saunders, and C. M. Zicovich-Wilson, Z. Kristallogr. 220, 571 (2005).

[32] Quamtum Espresso webpage: http://www.quantum-espresso.org/

[33] Laganà, A., Crocchianti, S., Ochoa de Aspuru, G., Gargano, R., Parker, G.A.: Parallel time independent quantum calculations of atom diatom reactivity. LNCS 1041 (1995) 361-370

[34] A. Aguado, C. Tablero, and M. Paniagua, Comput. Phys. Comm. 134, 97 (2001).

[35] S. C. Farantos, POMULT: A Program for Computing Periodic Orbits in Hamiltonian Systems Based on Multiple Shooting Algorithms, Comp. Phys. Comm., 1998, 108, 240-258

[36]  S. C. Farantos, Periodic Orbits in Biological Molecules: Phase Space Structures and Selectivity in Alanine Dipeptide, J. Chem. Phys., 126(17):175101-175107, 2007.

 [37] N. Balucani, P. Casavecchia, L. Banares, FJ Aoiz, T. Gonzalez-Lezana, P. Honvault and

J.M. Launay, J. Phys. Chem. A, 2006, 110, 817.

[38] Kepler website: https://kepler-project.org/. Last seen May 2013

[39] SHIWA project website: http://www.shiwa-worklfow.eu/project. Last seen May 2013

[40] E. Rossi, S. Evangelisti, A. Laganà, A. Monari, S. Rampino, M. Verdicchio, K. Baldridge, G.L. Bendazzoli, S. Borini, R. Cimiraglia, C. Angeli, P. Kallay, H.P. Lüthi, K. Ruud, J. Sanchez-Marin, A. Scemama, P. Szalay, A. Tajti, Code Interoperability and Standard Data Formats in Quantum Chemistry and Quantum Dynamics: the Q5/D5cost Data Model submitted to the J. Comp. Chem.

[41] EGI\_EUDAT\_PRACE pilot project:

https://wiki.egi.eu/wiki/EGI\_EUDAT\_PRACE\_collaboration