**EGI-InSPIRE**

TOWARDS A CMMST VRC virtual team project report

M1 outcome

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

|  |  |
| --- | --- |
| Date: | 18/03/2013 |
| Document Status: | **Draft** |
| Dissemination Level: | **Internal** |
| Document Link: | **https://documents.egi.eu/document/xxxx** |

|  |
| --- |
| **Abstract**  This planning document is the first Milestone (M1) outcome of the EGI-InSPIRE Virtual Team project ‘Towards a Chemistry, Molecular & Materials Science and Technology (CMMST) VRC’.  The goal of the VT project is the elaboration (on the ground of the present situation of the CMMST community and of the advantages that a VRC status would offer in satisfying the requirements of its members concerning the access and use of the computing resources federated in EGI) of a proposal guiding the set up of a CMMST VRC. This M1-outcome planning document  - provides a list of technical and non technical topics to be investigated;  - assigns the elaboration of such topics either to individuals or teams of the VT.  The technical topics proposed for investigation are:  - the consistency of the present patrimony of applications available to the community and the way of evolving to higher complexity applications  - the present ways of accessing compute resources and the EGI technologies, resources and services (already exist within EGI) that could be used to satisfy the requirements of the CMMST community and in particular to build workflows and workflows of workflows for the existing applications  - the possibility of a distributed execution of CMMST applications requiring both HTC and HPC on EGI and PRACE  - the use of Quality of User (QoU) and Quality of Service (QoS) parameters to build a credit system fostering collaboration  - the extension of CMMST applications to use in research based education  The non technical topics proposed for investigation are:  - the structure of the VRC and the development of a coordinated management body  - the attraction and training of more CMMST users  - the sustainability of the CMMST VRC  The various topics are assigned to the various members. |

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** |
| 1.0 | 18/03/2013 | Draft | A. Costantini |
| 2.0 | 19/03/2013 | Draft | A. Laganà |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

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 AimS 5

2.1 M1 outcomes 5

3 Applications and tools 6

3.1 Investigated Computational Chemistry applications 6

3.2 Investigated EGI tools 6

3.3 Additional tools 7

4 computational and Management aspects 7

4.1 Additional points 8

5 Conclusions 8

5.1 Conclusion 1 Errore. Il segnalibro non è definito.

Appendix Errore. Il segnalibro non è definito.

Appendix A – The “full name”VT membership 10

Annex 11

References 11

# Introduction

Virtual Research Communities (VRCs) are groups of like-minded individuals organised by discipline or computational model. A VRC can establish a support relationship, formalised through a Memorandums of Understanding (MoU), with the European Grid Infrastructure collaboration (EGI). [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.

The Chemistry, Molecular & Materials Science and Technology (CMMST) community recognises the advantages that membership as a VRC within EGI will bring. The VRC status could help the CMMST community satisfy the requirements of its members concerning the access and use of national computing resources that are federated in EGI. The proposed EGI Virtual Team (VT) project [1] will take the first step towards the setup of a CMMST VRC, by documenting:

* the structure that such a VRC should have to represent the CMMST community in EGI;
* the technologies, resources and services that already exist within EGI and could be used to satisfy the requirements of the CMMST VRC;
* the technologies that need to be developed or brought into EGI, then integrated with the production infrastructure so the VRC members can efficiently manage and use resources from EGI.

# AimS

The aims of the project will touch the following aspecs:

- develop a plan aimed at assembling a VRC out of the existing Chemistry, Molecular & Materials Science and Technology oriented EGI VOs and from the applications, tools and other resources and services that NGIs and projects of EGI 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.

- develop a proposal to establish a new CMMST VRC in EGI. Besides the technical aspects, the proposal will define the organisational and funding models for the VRC.

## M1 outcomes

In particular, in the present report, attention has been focused on the preparation of the Investigation Planning Document which contains a list of technical and non-technical topics that the project will investigate. Each topic is allocated to an individual or to a team with a leader from the VT.

# Applications and tools

Synthetic prospects of technical and non technical aspects to be investigated by the VT and to be filled and discussed at the next teleconference (March 19) are given below together with a list the specific arguments considered. In particular attention has been focused to the software packages (ported into the Grid environment) currently in use by the CMMST members and to the capabilities of the existing EGI and Grid-related tools.

## Investigated Computational Chemistry applications

In Table 1 a list of CMMST applications ported in the Grid environment and receiving dedicated user support activities are given (a full list of chemistry and solid state physics is given in <http://en.wikipedia.org/wiki/List_of_quantum_chemistry_and_solid_state_physics_software>). The ported packages are here briefly described :

*ABC*: The ABC program uses 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.   
*MCTDH*: The multiconfigurational time-dependent Hartree (MCTDH) method is nowadays considered as one of the most powerful tools for thermally averaged quantum dynamics simulations of multidimensional systems.

*FLUSS*: The fluss code performs a modified Lanczos iterative diagonalization of the thermal flux operator.

*VENUS96*: This application is a modified version of the VENUS96 program by W.L.Hase (QCPE-671). It calculates the trajectory for two reactants (atoms or molecules) by integrating the Hamilton equation in cartesian coordinates.

*RWAVEPR*: This application integrates rigorously the three-dimensional time-dependent Schroedinger equation for a generic atom-diatom reaction by propagating wave packets.

*GAMESS*: This is a program for ab initio molecular quantum chemistry computation which can be performed using direct techniques, or in parallel on appropriate hardware.

*DL\_POLY*: This is a package of subroutines, programs and data files, designed to facilitate molecular dynamics simulations of macromolecules, polymers, ionic systems, solutions and other molecular systems on a distributed memory parallel computer.

*GROMACS*: This is a versatile package to perform molecular dynamics, i.e. simulate the Newtonian equations of motion for systems with hundreds to millions of particles.

*NAMD*: Recipient of a 2002 Gordon Bell Award, is a parallel molecular dynamics code designed for high-performance simulation of large biomolecular systems.

*SC-IVR*: Semi-classical (SC) initial value representation (IVR) methods are used to calculate the thermal rate coefficients for the gas-phase reactions

## Investigated EGI tools

The same approach has been followed to list the in-use EGI tools for the support of the various communities operating on the European grid for their work:

*APPDB*

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 filed in the AppDB is ready to be used on the European Grid Infrastructure.

*DASHBOARDS*

* EDMS(Experiment Dashboard Monitoring System) a system to monitor, transfer data, commission sites and provide as well assistance and Virtual Organizations management. EDMS can operate on several Grid.

*USER INTERFACES AND FRAMEWORKS*

* GANGA an easy to use front end for job definition and management offering a uniform environment across multiple distributed computer systems;
* DIANEa lightweight task processing framework utilizing an application aware scheduler allowing an efficient and robust execution of large number of computational tasks on heterogeneous computing infrastructure.

*WORKFLOWS*

* Tools developed 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 ASKALON, KEPLER, K-WF GRID, MOTEUR, PEGASUS, P-GRADE, TAVERNA ,TRIANA, UNICORE WORKFLOW**.**

*GATEWAYS*

* Tools offering the service of routing packets outside the local network providing not only the basic functions but also a series of services which are often specific of a community. Among them SOMA2 is specific of the molecular science community).

*DATA MANAGEMENT*

* GREIC (Grid Relational Catalog) a tool providing a set of advanced data grid services aimed at transparently, efficiently and securely managing databases on the Grid;
* HYDRA a file encryption/decryption tool developed as part of the gLite middleware,
* MPI (Message Passing Interface) is a library of routines providing concurrent execution of parallel programs**;**
* FTS (File Transfer Service) is a lightweight but fullyfunctional set of services supporting data management;
* DPM (Disk Pool Manager);
* LFC (LCG File Catalog).

## Additional tools

There are also some additional tools (framework) developed internally by some members of the community and actually in use:

*GRIF*: Grid Framework enabling efficient and user-friendly scientific massive calculations

*GCRES*: Quality of Users (QoU), Quality of Services (QoS) evaluation Framework

*GGAMESS*: Front-end script for submitting multiple GAMESS-US jobs

G-LOREP: Grid LOs Repository project is its focus on large communities that implies both a complex and a distributed nature of the repository

# computational and Management aspects

The formal acceptance of a community in EGI is the signature of a Memorandum of Understanding (MoU). 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, 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).

For this reason appropriate answers to the following questions have to be provided:

1. How to exploit the capabilities of the existing EGI tools for distributing runs of CMMST applications on EGI and PRACE platforms (HTC and HPC platforms).
2. How to attract more CMMST users into a common endeavour offering the possibility of assembling higher level of complexity applications and services.
3. How to utilize a credit system to encourage CMMST users to cooperate in developing higher level of complexity applications (GriF and GCreS).
4. How to develop a coordinated management body for such endeavours and configure a Virtual Research Community (VRC).
5. How to operate the EGI VRC in a sustainable way.

## Additional points

To the end of building the VRC proposal document the following points have also to be specified:

*On the provision of technology side*

-Contribute to enable the vision of providing European scientists and international collaboration for sustainable distributed computing services to support their work.

-Provide robust, well designed, user centric services to scientific user communities

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

-Accelerate the development of standards within production grid infrastructures

-Disseminate results of the collaboration

Exchange ideas and collaborate on the definition of sustainability models

-Collaborate in business relationships development;

*On the provision of infrastructure side*

-Enhance the capacity of the infrastructure

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

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

-Comply with the operation interfaces needed to ensure seamless and interoperable access to resources

-Participate in the Operations Management Board to contribute to the EGI operations agenda

-Participate to security police team to contribute to the development to the security police fabric of the infrastructure

# Conclusions

This document provides a list of technical and non technical topics to be investigated by the VT and assigns their elaboration either to individuals or teams of the VT.

The topics proposed for investigation are:

- C1. the consistency of the available application patrimony of CMMST. This will allow to plan its evolution higher complexity applications

- C2. the abandoning of the present ways of accessing compute resources by exploiting EGI technologies, resources and services. This will allow to build workflows and workflows of workflows for the existing applications and distribute them as well for execution on EGI and PRACE in order to making combined usage of HTC and HPC

- C3. the use of Quality of User (QoU) and Quality of Service (QoS) parameters to build a credit system. This will allow to foster collaboration, develop a community grid economy, help with the VRC sustainability and ground a robust coordinated management body

- C4. the attraction of more CMMST users. This will be based on training (including an extension of CMMST applications to use in research based education)

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 |
| MCTDH | MultiConfigurational Time Dependent Hartree method | Academic |
| FLUSS | Lanczos iterative diagonalization | 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 |
| ggamess | Front-end script for submitting multiple GAMESS-US jobs | Free |

# Appendix A – The “full name”VT membership

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

|  |  |  |  |
| --- | --- | --- | --- |
| Country Code: | Name: | e-mail: | Role: |
| IT | Antonio Lagana (UNIPG) |  | VT Leader |
| IT | Daniele Cesini (INFN) |  |  |
| IT | Alessandro Costantini (INFN) |  |  |
|  | Stavros C. Farantos, IESL/FORTH and Chemistry UoC |  |  |
|  | Stamatis Stamatiadis, IESL/FORTH and Materials UoC |  |  |
|  | Manos Giatromanolakis, IESL/FORTH |  |  |
|  | Giannis Remediakis, Materials UoC |  |  |
|  | George Lendvay (MTA TTK Research Center for Natural Sciences) |  |  |
|  | Anna Vikar (MTA TTK Research Center for Natural Sciences) |  |  |
|  | Peter Szabo (MTA TTK Research Center for Natural Sciences) |  |  |
|  | Tibor Nagy (MTA TTK Research Center for Natural Sciences) |  |  |
|  | Akos Bencsura (MTA TTK Research Center for Natural Sciences) |  |  |
|  | Remco Havenith (RUG) |  |  |
|  | Fermin Huarte Larrañaga (IQTCUB) |  |  |
|  | Amaia Saracibar (University of the Basque Country) |  |  |
|  | Ernesto Garcia (University of the Basque Country) |  |  |
|  | Marco Verdicchio (CNRS) |  |  |
|  | Petr Hanousek (CESNET) |  |  |
|  | Peter Oliver (STFC) |  |  |
|  | Alan Sill (CERN) |  |  |
|  | Sergio Maffioletti (UZH) |  |  |
|  | Mariusz Sterzel (CYFRONET) |  |  |
|  | Elena Tamuliene (VU) |  |  |
|  | Gergely Sipos (UCST) |  |  |
|  | Richard McLennan (UCST) |  |  |
|  | Karolis Eigelis (UCST) |  |  |

# Annex

None

# References

[1] https://wiki.egi.eu/w/images/5/5f/VT\_CMMST\_Proposal\_v1.pdf