

TOWARDS A CMMST VRC

PROPOSAL FOR AN EGI VIRTUAL TEAM PROJECT

by

Antonio Lagana, Gergely Sipos, Richard McLennan

Date: **12/12/2012**

Document Status: **Draft**

Dissemination Level: **Internal**

Abstract

This document proposes the establishment of an EGI Virtual Team (VT) project to investigate services and organisational possibilities for a sustainable Virtual Research Community that could represent the Chemistry, Molecular & Materials Science and Technology community within the European Grid Infrastructure collaboration. The final output of the Virtual Team project would be a proposal that can guide the setup of the VRC. The project is expected to last for 6 months and include experts from all specialisations within this scientific field besides members of the EGI ecosystem.

I. DOCUMENT LOG

Version	Date	Comment	Author/Partner
0.8	3 Oct 12	1 st internal draft	Antonio Lagana / University of Perugia
0.9	11 Nov 12	2 nd internal draft	Gergely Sipos, Richard McLennan / EGI.eu
0.99	8 Dec 12	3 rd internal draft	Antonio Lagana / University of Perugia
1.0	17 Dec 12	Minor corrections, draft for EGI-InSPIRE NA2 leaders	Gergely Sipos, Richard McLennan / EGI.eu

II. APPLICATION AREA

This document is a project initiation document written under the EGI-InSPIRE NA2 virtual team framework. Further information is available at https://wiki.egi.eu/wiki/Virtual_team.

TABLE OF CONTENTS

1	INTRODUCTION / BACKGROUND	3
2	AIMS	3
3	TASKS.....	3
4	PARTICIPANTS / REQUIRED RESOURCES.....	4
5	JUSTIFICATION / VALUE.....	5
6	PROJECT LENGTH & DELIVERABLES.....	5
7	REPORTING.....	6

1 INTRODUCTION / BACKGROUND

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 VRCs¹ 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 will take the first step towards the setup of a CMMST VRC, by documenting:

- i. the structure that such a VRC should have to represent the Chemistry, Molecular & Materials Science and Technology community in EGI;
- ii. the technologies, resources and services that already exist within EGI and could be used to satisfy the requirements of the CMMST VRC;
- iii. 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.

2 AIMS

- i. To 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.
- ii. 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.
- iii. On the basis the above two, 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.

3 TASKS

The VT will investigate the following items, then will document its findings in a VRC proposal:

- i. How to exploit the capabilities of the existing EGI tools in building distributed workflows and “workflows of workflows” from various software packages, including:
 - a. Molecular Simulators (like GEMS the Grid Empowered Molecular Simulator).
 - b. Electronic structure computation (like GAUSSIAN, GAMESS, CRYSTAL, etc.).
 - c. Quantum and classical molecular dynamics computation (like ABC, GROMACS, MCTDH, VENUS96, DL_POLY, RWAVEPR, etc.)
 - d. Statistical averaging to produce physical observables

¹ <http://www.egi.eu/community/vrcs/>

- e. distributed database and knowledge repositories (G-LOREP)
- ii. 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).
- iii. How to attract more CMMST users into a common endeavour offering the possibility of assembling higher level of complexity applications and services.
- iv. How to utilize a credit system to encourage CMMST users to cooperate in developing higher level of complexity applications.
- v. How to develop a coordinated management body for such endeavours and configure a Virtual Research Community (VRC).
- vi. How to operate the EGI VRC in a sustainable way.

Note:

A presentation² by Antonio Lagana at the EGI Technical Forum 2012 event (within the EGI/EUDAT/PRACE workshop) provided information on current activities within the CMMST community. The following items were covered:

- the development of tools, to select the computing resources better suited to run the different tasks of a complex application both in EGI and PRACE platforms,
- the evaluation of the quality of users and services (GriF),
- the reward with credits of the work done by the community members (GCreS)
- the complementary focus on educational aspects (Insilico Lab)
- is a suitable ground for developing a plan for the assemblage of the intended CMMST VRC

4 PARTICIPANTS / REQUIRED RESOURCES

Because the investigations that need to be carried out by the VT cover a wide spectrum of technical and non-technical areas that are all key to the definition of the VRC, it is crucial that the VT has members who are knowledgeable about these areas. This could be achieved with the support of EGI.eu, the support of the NILs, the support of Antonio Lagana, who will be the contact person between the VT and key institutes/organisations from the CMMST domain. The representation of these institutes is very important because without them the VT is unlikely to be able to propose a sustainable structure and a scope that is relevant scope for the CCMST community.

To a large extent, the VRC is expected to exist as a collaboration of some of the existing EGI Virtual Organizations (VO), namely COMPCHEM, GAUSSIAN, CHEM-IBERGRID, and of some of the communities that already expressed interest in participation (WeNMR³, European Chemistry Thematic Network⁴). Besides them, the following entities should be also directly or indirectly be involved in the VT:

- i. **Research Laboratories:** CNAF (I), UNIPG (I), UPV (ES), UB (ES), TUW (PL), Univ. Toulouse (FR), FORTH (GR), Univ. Groningen (NL), Univ. Thessaloniki (GR) and some other user clusters including Brazil, US (Houston, Atlanta)
- ii. **Computer Centres:** WCSS (PL), CINECA (IT), CESCA (ES), GC3 (CH), FORTH (GR) (other possible support is likely to be obtainable from CYFRONET (PL), CESGA (ES), SARA(NL) and other centres)
- iii. **NGIs:** IGI, IBERGRID plus any other NGI interested in the activities of the VT
- iv. The computational chemistry division of **EUCHEMS** (<http://www.euchems.org/>)

² <https://indico.egi.eu/indico/contributionDisplay.py?contribId=157&confId=1019>

³ WeNMR: <http://www.wenmr.eu/>

⁴ European Chemistry Thematic Network: <http://ectn-assoc.cpe.fr/network/index.htm>

The project will be managed by Antonio Lagana from the University of Perugia. He will dedicate on the average 3 person hours/week of his time on VT activities. Administrative support for the VT project will be provided by Richard McLennan (User Community Support Officer at EGI.eu). The EGI contact of the VT will be Gergely Sipos (Technical Outreach Manager at EGI.eu), who will be responsible for the delivery of activities from the EGI side.

Other project members are expected to be proactive in their contribution. Depending on their specific roles in the VT (resource provider, application developer, tool provider, CMMST institute/community representative, etc.) they shall normally:

- attend VT teleconference meetings (typically 1 hr/week)
- share experience and requirements on the VT topics
- share tools and dedicated applications
- contribute to setting strategies and targets of the community
- contribute to defining metrics to evaluate the work done by the VT members on behalf of the community
- contribute to the documentation of the projects findings and the follow-on proposal(s).

5 JUSTIFICATION / VALUE

The potential value for the EGI VT project includes:

- i. extension of the CMMST international (EU and non EU) community
- ii. cooperation with the European Computational Chemistry Division of EUCHEMS
- iii. cooperation with the European Chemistry Thematic Network ECTN
- iv. the growth of the area for innovation and technology transfer
- v. the definition of new requirements for molecular and materials science and technology grid/cloud applications
- vi. definition of a sustainable representation of the CMMST community in EGI

6 PROJECT LENGTH & DELIVERABLES

The expected duration of the project is 6 months. The planned outcomes of the project are:

Milestone number	Outcome	Means of delivery	Time of delivery (after VT start)
M1	The list of technical and non-technical topics that the project should investigate (finalise list from Section 3). Each topic is allocated to an individual or to a team with a leader from the VT.	• Investigation planning document	1 month
M2	Outcome of the technical and non-technical investigations are integrated into the first draft of the VRC plan. The draft is restricted to the project leader.	• VRC proposal document (1 st draft)	4 month
M3	An improved draft, with chapters reviewed and integrated by the VT leader. The draft is shared with VT members.	• VRC proposal document (2 nd draft)	4.5 month
M4	Integrated draft reviewed and commented by VT members.	• VRC proposal document	5.5 month

		(3 rd draft)	
M5	Final draft – a proposal for the setup of a CMMST VRC in EGI.	<ul style="list-style-type: none"> • VRC proposal document (public version) 	6 month

7 REPORTING

The VT project manager will provide a short emailed progress report on a weekly basis to NA2 Leaders⁵ (inspire-na2-leaders@egi.eu). The report will be due by 17:00 on Fridays and is to contain details of:

- Work achieved that week
- Work planned for next week
- Progress against the goals in the project plan
- Issues that the virtual team leader needs help with (e.g. non-responsive partners, more resources, support from EGI.eu teams, etc.)

The VT project manager will also provide a more comprehensive input (around half page) for the NA2 section of the EGI-InSPIRE quarterly report. This input is required every three months during the VT project activity.

⁵ This is the normal reporting process for EGI VT projects.