**EGI-InSPIRE**

TOWARDS A CMMST VRC virtual team project report

M2 outcomes

[**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 incorporates the first and the second Milestone outcomes (M1 and M2) 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 goal of the VT project is the elaboration of a proposal guiding the set up of a CMMST VRC on the ground of the present situation of the CMMST community and of the advantages that a VRC status would offer for what concerns the satisfaction of the members requirements and the access and use of the computing resources federated in EGI. Such proposal, detailed in the present document, enumerates the technical and non technical aspects related to the building of the CMMST VRC and analyses technology, structure and organization evolution (at national and European level) of distributed computing (Grid and Cloud) enabling the set up of the CMMST VRC. The document engages also the community in the development forms of cooperative computing based on the selection of the resources (from personal systems to supercomputers) and services (from number crunching to massive data handling on heterogeneous platforms). On this basis, the CMMST VRC (as well as other communities which use distributed computing resources) will be able to share hardware and software that (thanks to the expertise offered by the various groups) can be integrated to address complex problems and computing simulations of high social impact. |

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** |
| 0.1 | 18/06/2013 | Draft0 | A. Costantini |
| 0.2 | 19/06/2013 | Draft1 | A. Costantini |
| 0.3 | 26/06/2013 | Draft1 | A. Laganà |
| 1.0 | 30/06/2013 | Draft | A. Costantini |
| 1.1 | 01/07/2013 | Draft | D. Cesini |
| 1.2 | 05/07/2013 | Draft | M. Ceotto |
| 1.3 | 12/07/2013 | Draft | A. Costantini |

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 outcomes 6

3 the CMMST-VRC proposal 6

3.1 Current scenario of computational resources and their usage 6

3.2 The synergistic Grid model for the CMMST VRC 7

3.3 Coordination and Management 7

3.3.1 The coordinated management body for the VRC 8

3.3.2 TASK A1: User support and training 8

3.3.3 TASK A2: Services and operations 8

3.3.4 TASK A3: Requirement gathering 8

3.3.5 TASK A4: Communication and Dissemination (A4) 9

3.3.6 Sustainability 9

4 Technologies and services offered by the CMMST-VRC through EGI 9

4.1 Provided Computational Chemistry applications 9

4.2 Provided Grid tools and services 10

4.3 Additional tools 12

5 HPC-HTC integration 12

5.1 The CMMST activity 13

5.2 The EGI/EUDAT/PRACE collaboration 13

5.2.1 Objectives and related pilot activities 13

6 Conclusions 13

Appendix A – The “full name”VT membership 15

Annex 16

References 16

# 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 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. Each topic has been allocated either to an individual or to a team leader from the VT.

## M2 outcomes

M2 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).

# the CMMST-VRC proposal

## Current scenario of computational resources and their 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 one-site (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 one-site 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

A synergistic innovative model of allocating resources was developed within a series of European projects and collaborations (DATATAG, EGEE-I-II-III, WLCG, etc) mainly under the pressure of the research needs of High Energy Physics. Such synergistic model, based on the aggregation of a large number of geographically dispersed CPUs and users connected over the public network and managed through the use of appropriate middleware and tools, has made possible the execution of distributed programs on a large number (over hundreds of thousands) of processors, reaching the maximum performance for those 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. Within EGI-Inspire European project the Grid model has undergone, in fact, a significant increase in networked resources, software and skills (otherwise confined in the local sites) as well as in the proliferation of collaborations among scientists.

To preserve the peculiarities of the different research areas within the Grid and strengthen their role their members are being articulated as VO (http://www.egi.eu/community/vos/) and VRCs (http://www.egi.eu/community/vrcs/) by gathering together individual researchers, laboratories and virtual organizations operating in them.

The specific commitment of the COMPCHEM VO is to join its forces with those of other actors of the CMMST community (especially those of similar VOs like GAUSSIAN, WeNMR, etc.) and structure a VRC able to manage related resources, 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 VRC scientists should be able to compose such simulations a la carte from standard components made available to the members of the community by other members exploiting user friendly tools that will guarantee flat access to the most suited computing resources (within the limits of availability). Through the VRC the work done on behalf of the community and the services provided to its members could be recognised and rewarded.

The angular stones of such endeavour are both the willingness of the members of the community to feed the VRC with the results of their work to be offered as a service to the community (new algorithms and applications, new validated data, etc.) and the exploitation of the functionalities embodied in two tools of the COMCPHEM VO (the evaluation of the quality of both the services (QoS) and the users (QoU) /rif GriF/) and the awarding of credits /GCres/)

## 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). 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 the following activities have been formalized:

### The coordinated management body for the VRC

The VRC management body is composed by

- VRC leader: is the person representing the VRC and is the MoU contact point

- VRC deputy leader: is the person acting on the behalf of the VRC leader

- VRC task coordinators: are the persons representing the tasks of the VRC and acting as experts in the assigned roles

 - User support and training Coordinator: is the person taking care of managing user support and training activities (A1)

 - Operations Coordinator: is the person taking care of managing the services and operation activities (A2)

 - Technical Coordinator: is the person taking care of managing requirements gathering activities (A3)

 - Dissemination Coordinator: is the person taking care of managing communication and dissemination activities (A4)

### TASK A1: User support and training

-provide dedicated user support activities for code porting multi platform computing model 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

-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.

### TASK A3: Requirement gathering

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

-participate to 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 policies, procedures and Operation Level Agreements (OLA)s.

### TASK A4: Communication and Dissemination (A4)

-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.

### Sustainability

The VRC management body collaborates 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.

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

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*: 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*: performs a modified Lanczos iterative diagonalization of the thermal flux operator.

*VENUS96*: 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*: 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

## Provided Grid tools and services

The same approach has been followed to list the in-use EGI tools and connected initiatives and projects 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
	+ SHIVA SIMULATION PLATFORM provides 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.
	+ ASKALON is a tool aimed at simplifying the development and optimization of applications that can harness the power of Grid and Cloud computing.
	+ GC3PIE is a suite of Python classes (and command-line tools built upon them) to aid in submitting and controlling batch jobs to clusters and grid resources seamlessly. GC3Pie aims at providing the building blocks by which Python scripts that combine several applications in a dynamic workflow can be quickly developed.
	+ KEPLER is free and open source application for scientific workflow management and designed to help scientists and developers to create, execute, and share models (workflows) that can refer to broad range of scientific and engineering disciplines.
	+ MOTEUR is a data-driven workflow manager for enacting scientific applications on the EGI.
	+ PEGASUS is a project encompasses a set of technologies that help workflow-based applications execute in a number of different environments including desktops, campus clusters, grids, and clouds.
	+ gUSE WS-PGRADE 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 of various DCIs.
	+ SOMA2 is a molecular modelling workflow environment used with WWW-browser and it allows users to combine scientific applications into unique application workflows, which are automatically executed in the underlying computing system.
	+ TAVERNA is an open source and domain-independent Workflow Management System – a suite of tools used to design and execute scientific workflows and aid InSilicoexperimentation.
	+ TRIANA is a workflow-based graphical problem solving environment, which was initially developed to provide a quick-look analysis tool for gravitational wave data in association with the GEO 600 project.
	+ UNICORE WORKFLOW SYSTEM allows to run complex workflows involving UNICORE execution services, offering a wide range of control constructs (for-each, if-else, ...) workflow variables and other features.

*SCIENCE 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.

* + SOMA2 gateway is a molecular modelling workflow environment developed and deployed by CSC - IT Centre for Science Ltd. The SOMA2 environment is used with WWW-browser and it allows users to combine scientific applications into unique application workflows, which are automatically executed in the underlying computing system
	+ CHAIN has been built in the context of the EU CHAIN project to demonstrate how Science Gateway paradigm and standard adoption can make e-Infrastructures worldwide, based on different middleware, interoperable among each other at user level.
	+ The Catania Science Gateway Framework (CSGF) has been developed by INFN, Division of Catania (Italy), to provide application developers with a tool to create Science Gateways in short time and in a very easy way.
	+ GriF [4] is a Service Oriented Architecture (SOA) Collaborative Framework designed to facilitate the use of the Grid by non specialists aimed at optimizing the selection of different computing elements for running single and parameter study applications.
	+ InSilico Lab 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.
	+ IGI Portal is a science gateway providing a workflow environment enable to access Grid resources and supporting the EGI Federated Cloud

*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,
* FTS (File Transfer Service) is a lightweight but fullyfunctional set of services supporting data management;

*MIDDLEWARE COMPONENTS*

* MPI (Message Passing Interface) is a library of routines providing concurrent execution of parallel programs**;**
* DPM (Disk Pool Manager) is a lightweight solution for disk [storage management](https://www.gridpp.ac.uk/wiki/Grid_Storage%22%20%5Co%20%22Grid%20Storage). If offers the required SRM [5] interfaces, hopefully without being complicated by other modes of access or complications such as tape storage systems.;
* LFC (LCG File Catalog) is a high performance file catalog containing logical to physical file mappings.

## Additional tools

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

*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

# HPC-HTC integration

Even if the grid approach meets the needs of the CMMST community in overcoming the limitations of both the grant model and the purely opportunistic computing models, work is in progress in order to overcome the present highly unsatisfactory situation in which neither HPC nor HTC are completely fit alone to meet the requests of complex CMMST applications. 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.

## The CMMST activity

An advantage of such distributed model consists in the possibility for the user to choose the platform better fitted for his/her applications and for the computer centres to qualify the usage of theirs resources. This provides the CMMST members with the possibility of combining different pieces of software to the end of assembling complex realistic simulators (like GEMS [2]) 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 [3]. This has prompted the development of tools (like the framework GriF) enabling the redirecting of computer applications to run on the best suited sites (including HPC machines) by properly redirecting the jobs to the most appropriate architecture. This enhances cooperative compute capabilities by opening the perspective of combining different complementary know how into single (higher level of complexity) realistic applications and paves the way to apply for more ambitious research grants to be shared within the community.

## The EGI/EUDAT/PRACE collaboration

The goal of this EGI activity [6] 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.

### Objectives and related pilot activities

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 [7] 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

# 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 |
| 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 |
| InSilicoLab | Application portal designed to support in silico experiments | Free |
| IGI Portal | science gateway providing a workflow environment and a Cloud access | Free |
| SOMA2 | gateway for molecular modeling workflow environment | Free |
| GC3PIE | Framework providing building block to build dynamic workflows | 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)  |  |  |
| IT | Carlo Manuali (UNIPG) |  |  |
| IT | Michele Ceotto (UNIMI) |  |  |
| GR | Stavros C. Farantos, IESL/FORTH and Chemistry UoC  |  |  |
| GR | Stamatis Stamatiadis, IESL/FORTH and Materials UoC  |  |  |
| GR | Manos Giatromanolakis, IESL/FORTH  |  |  |
|  | Giannis Remediakis, Materials UoC  |  |  |
| HU | George Lendvay (MTA TTK Research Center for Natural Sciences)  |  |  |
| HU | Anna Vikar (MTA TTK Research Center for Natural Sciences)  |  |  |
| HU | Peter Szabo (MTA TTK Research Center for Natural Sciences)  |  |  |
| HU | Tibor Nagy (MTA TTK Research Center for Natural Sciences)  |  |  |
| HU | Akos Bencsura (MTA TTK Research Center for Natural Sciences) |  |  |
| NL | Remco Havenith (RUG) |  |  |
| ES | Fermin Huarte Larrañaga (IQTCUB)  |  |  |
| ES | Amaia Saracibar (University of the Basque Country)  |  |  |
| ES | Ernesto Garcia (University of the Basque Country)  |  |  |
| FR | Marco Verdicchio (CNRS)  |  |  |
| CR | Petr Hanousek (CESNET) |  |  |
| UK | Peter Oliver (STFC) |  |  |
| USA | Alan Sill (CERN)  |  |  |
| CH | Sergio Maffioletti (UZH)  |  |  |
| PL | Mariusz Sterzel (CYFRONET) |  |  |
| EGI | Elena Tamuliene (VU) |  |  |
| EGI | Gergely Sipos (UCST)  |  |  |
| EGI | Richard McLennan (UCST)  |  |  |
| EGI | Karolis Eigelis (UCST)  |  |  |

# Annex

None

# References

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

[2] 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)

[3] 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.

[4] 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)

[5] https://www.gridpp.ac.uk/wiki/Grid\_Storage

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

[7] https://indico.egi.eu/indico/conferenceDisplay.py?confId=1228