EGI-InSPIRE

SUSTAINABILITY PLANS FOR THE ACTIVITIES OF THE HEAVY USER COMMUNITIES

EU DELIVERABLE: D6.5

Document identifier:	EGI-741-V7.doc
Date:	**13/06/2012**
Activity:	**SA3**
Lead Partner:	**CERN**
Document Status:	**DRAFT**
Dissemination Level:	**PUBLIC**
Document Link:	https://documents.egi.eu/document/741

Abstract

The Heavy User Community activities are planning for a sustainable future – either within their own community, as part of the generic production infrastructure, or through some other mechanism. This annual report exposes the sustainability plan for each task within the SA3 activity and assesses the progress made to date towards sustainability and the plans for improving sustainability in 2012/13.

II. DELIVERY SLIP

	Name	Partner/Activity	Date
From	SA3	Jamie Shiers/CERN	1/6/2012
Reviewed by	**Moderator:** **Reviewers:**	Steven Newhouse/EGI.eu	6/6/2012
Approved by	**AMB & PMB**		12/6/2012

III. DOCUMENT LOG

Issue	Date	Comment	Author/Partner
1	1/6/2012	First draft	Jamie Shiers/CERN
2	8/6/2012	Revised following review	Steven Newhouse/EGI.eu

IV. APPLICATION AREA

This document is a formal deliverable for the European Commission, applicable to all members of the EGI-InSPIRE project, beneficiaries and Joint Research Unit members, as well as its collaborating projects.

V. DOCUMENT AMENDMENT PROCEDURE

Amendments, comments and suggestions should be sent to the authors. The procedures documented in the EGI-InSPIRE "Document Management Procedure" will be followed: https://wiki.egi.eu/wiki/Procedures

VI. TERMINOLOGY

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

VII. PROJECT SUMMARY

To support science and innovation, a lasting operational model for e-Science is needed − both for coordinating the infrastructure and for delivering integrated services that cross national borders.

The EGI-InSPIRE project will support the transition from a project-based system to a sustainable pan-European e-Infrastructure, by supporting 'grids' of high-performance computing (HPC) and high-throughput computing (HTC) resources. EGI-InSPIRE will also be ideally placed to integrate new Distributed Computing Infrastructures (DCIs) such as clouds, supercomputing networks and desktop grids, to benefit user communities within the European Research Area.

EGI-InSPIRE will collect user requirements and provide support for the current and potential new user communities, for example within the ESFRI projects. Additional support will also be given to the current heavy users of the infrastructure, such as high energy physics, computational chemistry and life sciences, as they move their critical services and tools from a centralised support model to one driven by their own individual communities.

The objectives of the project are:

1. The continued operation and expansion of today's production infrastructure by transitioning to a governance model and operational infrastructure that can be increasingly sustained outside of specific project funding.
2. The continued support of researchers within Europe and their international collaborators that are using the current production infrastructure.
3. The support for current heavy users of the infrastructure in earth science, astronomy and astrophysics, fusion, computational chemistry and materials science technology, life sciences and high energy physics as they move to sustainable support models for their own communities.
4. Interfaces that expand access to new user communities including new potential heavy users of the infrastructure from the ESFRI projects.
5. Mechanisms to integrate existing infrastructure providers in Europe and around the world into the production infrastructure, so as to provide transparent access to all authorised users.
6. Establish processes and procedures to allow the integration of new DCI technologies (e.g. clouds, volunteer desktop grids) and heterogeneous resources (e.g. HTC and HPC) into a seamless production infrastructure as they mature and demonstrate value to the EGI community.

The EGI community is a federation of independent national and community resource providers, whose resources support specific research communities and international collaborators both within Europe and worldwide. EGI.eu, coordinator of EGI-InSPIRE, brings together partner institutions established within the community to provide a set of essential human and technical services that enable secure integrated access to distributed resources on behalf of the community.

The production infrastructure supports Virtual Research Communities (VRCs) − structured international user communities − that are grouped into specific research domains. VRCs are formally represented within EGI at both a technical and strategic level.

VIII. EXECUTIVE SUMMARY

The Heavy User Community activities are planning for a sustainable future – either within their own community, as part of the generic production infrastructure, or through some other mechanism. This annual report exposes the sustainability plan for each task within the SA3 activity and assesses the progress made to date towards sustainability and the plans for improving sustainability in 2012/2013.

The basic model for sustainability of the solutions and services that are currently supported by SA3 is that of collaborative development. This isolates individual projects from funding changes at individual participating institutes, but cannot easily cater for significant reductions in funding, such as that which will occur when SA3 terminates in April 2013. In other words, it mitigates, but does not eliminate, these risks.

However, as the main objective of the work package has been to ensure a smooth transition, to prepare for the longer term by identifying commonalities and other synergies, the work package is well on target in meeting its objectives.

The primary goal of the final project year for this work package must therefore be to complete any remaining development and ensure that any tools / services are adequately documented and that the support is passed over to the units that will carry it on in future years.

TABLE OF CONTENTS

1 INTRODUCTION

The SA3 activity within EGI-InSPIRE provides continued support for activities previously supported by EGEE while they transition to a sustainable support model within their own community or within the production infrastructure by:

- Supporting the tools, services and capabilities required by different Heavy User Communities (HUCs);
- Identifying the tools, services and capabilities currently used by the HUCs that can benefit all user communities and promoting their adoption;
- Migrating the tools, services and capabilities that could benefit all user communities into a sustainable support model as part of the core production infrastructure;
- Establishing a sustainable support model for the tools, services and capabilities that will remain relevant to single HUCs;
- As a further step, seeking additional areas of commonality that would result in lower on-going support costs through reduced diversity.

All of the above mentioned tasks aim to reduce the cost of new development and support, and have as a common objective improving the sustainability of the model. Working towards sustainability is therefore a key element of the strategy of this activity and this document describes the progress to date in achieving this objective together with plans for the coming year. The document is organised by HUC – a description of the communities involved is given in the section below.

For the Heavy User Community (HUC) workshop held at the EGI TF 2001, each community was asked to come prepared with answers to three questions regarding the sustainability of their respective discipline, Distributed Computer Infrastructure (DCI) and the support the community received. Below is a table summarising the key points that were highlighted by each community.

HUC	SUSTAINABILITY OF:		
	DISCIPLINE	DCI	SUPPORT
Life Sciences	Life Sciences is a very broad “discipline”. • Ranging from fundamental research to direct clinical applications. • Every potential patient may be interested: really everybody. • Strong societal impact.	Foundation of the Life Sciences Grid Community. • http://wiki.healthgrid.org/LSVRC:Index • 5 NGIs (Dutch, French, Italian, Spanish, Swiss). • 5 VOs (biomed, lsgrid, medigrid, pneumogrid, vlemed). • 2 projects (EGI-InSPIRE, Lifewatch ESFRI).	Services: • LSGC dashboard and support team. • Hydra encryption / decryption service. • GRelC database interface service. • Taverna workflow manager. • EBI core databases.
Astronomy & Astrophysics	• Worldwide, distributed community that operates large experiments and observing facilities. • Observation-driven science based on national and international ground	• A&A Grid/DCI community funded so far through short/medium term national and international projects (e.g. EGI-InSPIRE and former EGEE projects) and funds typically from	• The most relevant SA3 tool directly provided by A&A is VisIVO.

	and space-based observatories.	Europe and national governments.	
Earth Sciences	Fundamental topics: • Understanding of the Sun-Earth system. • Societal needs. • Civil protection. • Natural resources identification / exploration. • Reaction to natural and human-induced changes. • Long-term Climatology • Multi-disciplines: to observe and interpret one event. • Cross scientific domains: Physics, chemistry, fluid mechanics. • International societies and programs.	Resources: • Resources from ES related sides directly at NGIs. • Resources are not only provided by EGI. • Support • Community building. • Community tools. • Publications: see actual papers from e.g. climate, hydrology, and satellite data exploitation.	Services: • Organisational: achieved collaboration e.g. with GENESI-DR or ESGF –on-going applications. • Software. • Generic VO and services.
High Energy Physics	• HEP is clearly a long term effort, the main experiments are becoming bigger and bigger.	• Grids have become integral part of HEP computing. • WLCG has a MoU with long term financial commitments by the participating countries.	• A strategy following multiple paths. • More commonality = Lower maintenance burden. • HEP experience is a key resource here.

2 HEAVY USER COMMUNITIES

Heavy User Communities (HUCs) are Virtual Research Communities (VRCs) that have been using EGEE and now EGI routinely and thus have become more structured and advanced in terms of their grid usage. These communities focus on domain specific issues, such as how to access High Energy Physics (HEP) applications on EGI, how to enable new physics experiments on EGI and so on.

On the one hand these teams are operated by external projects, such as the Worldwide LHC Computing Grid (WLCG), but on the other hand have members in the SA3/WP6 work package of EGInSPIRE. The effort of the distributed WP6 team of EGI-InSPIRE is targeted towards the provision of shared services that will ease the porting of new applications from these scientific domains to the wider grid by detecting and exploiting commonalities between virtual organisations (VOs) and driving the implementations to a generic direction.

At the same time inter-VO collaboration typically not only result in more powerful solutions, but also saves significant amounts of manpower in the long run. Such benefits would be unlikely to be achieved with generic support structures, both for individual large communities such as HEP (which could otherwise develop multiple similar solutions to basically common problems), as well across disciplines (e.g. the usage of Dashboards [DASHBOARD], Ganga [GANGA] and HammerCloud [HAMMERCLOUD] across communities).

In conclusion HUCs can offer benefits not only to new adopters of grid technology but also to each other. This continues to be demonstrated, both by the adoption of tools initially developed for one community spreading to others, as well as at the conceptual level: which offers a more pragmatic solution for existing communities.

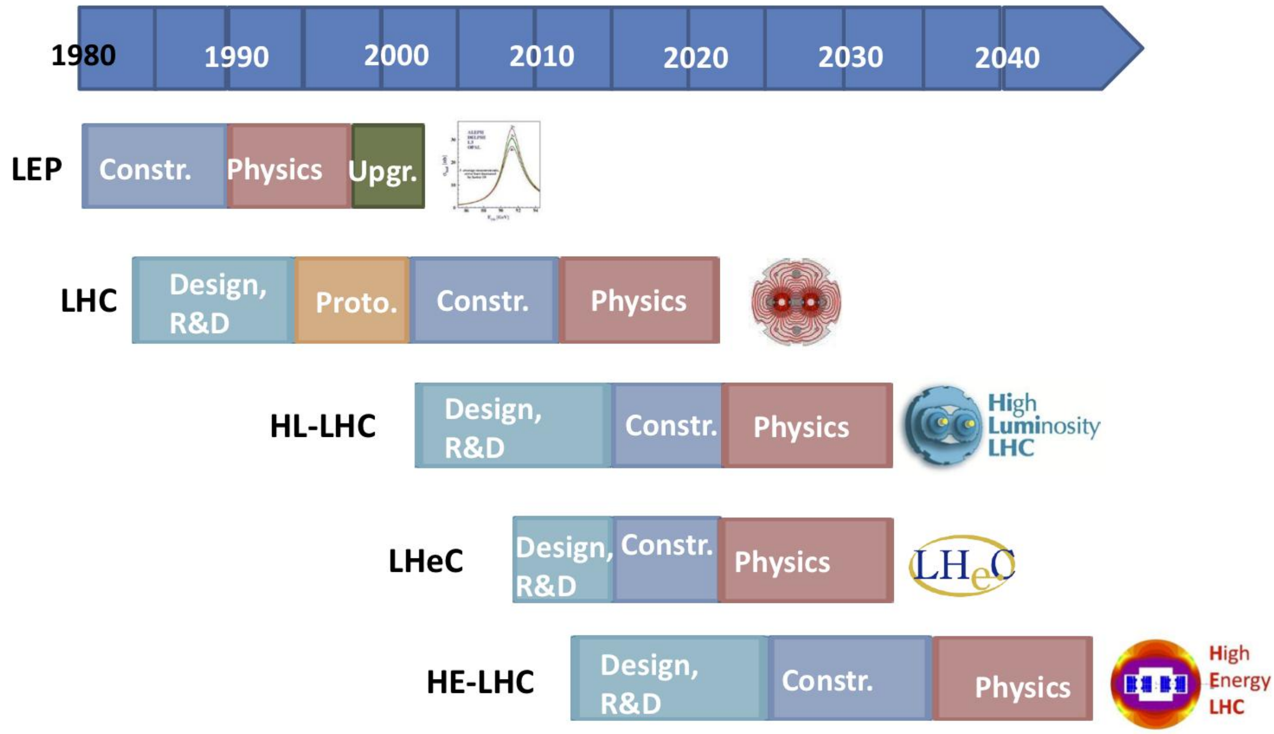
3 HIGH ENERGY PHYSICS

3.1 *Introduction*

The Services for High Energy Physics (HEP) task continues to focus primarily but not exclusively on the 4 LHC VOs (experiments) centred at CERN: ALICE, ATLAS, CMS and LHCb (and hence for and via the WLCG project and collaboration). Services and tools developed or extended by these VOs are also used by other HEP experiments and/or are under consideration for the future.

HEP is clearly a long term effort, the main experiments are becoming bigger and bigger. Hence, some appropriate level of funding is expected to be available long term to continue to support this work. The community of physicists is well connected and organized at the national and international level even when different “data producing” labs are used. However the HEP computing community has in the last 10 years counted on projects (EU and others) for gaining extra funds and for sharing the development and maintenance effort with other communities within and outside HEP.

The timeline of the LHC and related accelerators / colliders is given in the figure below. This shows a tentative programme lasting for the next quarter of a century that is obviously subject to funding and approval. However, the WLCG Virtual Research Community and / or its successors will need to be sustained for at least this duration, followed by up to a decade of further analysis that typically follows after the machine has finished its exploitation phase.



3.1.1 Distributed Computer Infrastructure (DCI)

Grids have become an integral part of HEP computing models of the LHC experiments and other big accelerator experiments like CDF and planned ones such as SuperB. WLCG has a MoU with long term financial commitments from the participating countries, however only part of the grid services are really relevant for WLCG, and the LHC experiments have in many cases developed complex and different frameworks for accessing grid services.

3.1.2 Support

For HEP support there is a strategy following many paths. Within the LHC the plan is to harden and simplify the services and increase their commonality (WLCG TF [1]being launched). As there is more commonality now, less maintenance is needed. Within HUCs commonality is increasing and within EGI they are participating in the phases of design and then execution of a renewed effort for attracting new commonalities. HEP experience is a key resource here. With any interested community they will look for innovative projects that may make Grid/Cloud use easier and more widespread and attract new project funds to these projects.

3.2 Analysis and Tools

3.2.1 HammerCloud

HammerCloud (HC)[2] is a Grid site testing service developed around Ganga. HC uses frequent short jobs to validate a site's availability and functionality, and also delivers on-demand stress tests to aid in site commissioning or general benchmarking. HC was developed with the ATLAS experiment but is used today by CMS and LHCb experiments.

Regarding organization and sustainability, the HC project is currently led by an EGI-InSPIRE funded FTE, and developments and operations are complemented by at least 1 partial FTE from each user VO (ATLAS has 2 extra support persons (partial)). It is foreseen that the continued life of the project does not depend on any one of these partners.

3.2.2 CRAB Client

The CMS Remote Analysis Builder (CRAB) was the first analysis tool in CMS to aid users in configuring CMS applications for distributed use, by discovering the location of remote datasets and submitting jobs to the Grid infrastructure. CRAB has progressed from a limited initial prototype nearly 5 years ago to a fully validated system that is critical to the production of published physics results. CMS currently observes more than 400 unique users submitting CRAB jobs per week, with close to 1000 individuals per month. The CMS Computing Technical Design Report (CTDR) estimated roughly 100k Grid submissions per day. In the second half of 2010 the job submissions routinely exceeded the estimate by 40-50% and CRAB coped well with the increased load.

After the development of the CRAB submission client a CRAB server was developed, which has increased the scalability of submission and added capabilities of automatic resubmission. The CRAB server also provides a development platform for additional capabilities.

During the past year effort has been spent both to maintain the production version of CRAB (CRAB2) and to develop & commission a new CRAB3.

CRAB is a joint project among various institutes, mostly from the Experiment Support group at CERN, from Fermilab and the Italian Institute for HEP (INFN). If any of these teams decrease the effort spent on the project, although there will be an impact on CRAB, it will not affect the sustainability issue.

---

[1] Worldwide LHC Computing Grid Task Force

[2] https://savannah.cern.ch/projects/hammercloud/

3.2.3 Ganga

Ganga is an end-user tool for creating and managing computational tasks. The mature Python codebase provides for a stable yet extensible framework, which is consistently used by over 400 end-users per month, distributed over 100 individual sites and submitting in excess of 250,000 jobs per week. It remains a popular environment for running grid analysis jobs within the LHCb and ATLAS experiments, and is known to be used by a further 10 communities across a range of scientific disciplines.

Future developments will focus on increasing the adoption of Ganga as an analysis platform. For example, inclusion of Ganga within a recognised Linux distribution would provide both a familiar installation procedure for new users, and reinforce the fact that Ganga is a mature and established tool.

The Ganga project is led by Imperial College, UK with contributors from other universities in the UK, Norway and Germany plus effort provided through EGI-InSPIRE and CERN. The non-project effort will remain even after EGI-InSPIRE effort is no longer available.

3.3 Data Management

All of the data management activities described below are support by collaborative effort from institutes that are part of the corresponding collaboration (e.g. ATLAS, CMS or LHCb) and / or major laboratories such as CERN or Fermilab and / or major national institutes such as INFN.

3.3.1 ATLAS Distributed Data Management

ATLAS, one of the LHC experiments, fully relies on the use of Grid computing for offline processing and analysis. This processing is done worldwide using the well-known tier model across heterogeneous interoperable Grids and the ATLAS Distributed Data Management (DDM) project is responsible for the replication, access and bookkeeping of ATLAS data across more than 100 distributed Grid sites.

The current ATLAS DDM software is now in a mature state and the present work is focused on maintenance and support operations. Inside the ATLAS Distributed Computing community there is an on-going discussion about the proposal on the development of a new DDM system (the Rucio project) to solve the current shortcomings and scalability issues in the Central Catalogues. The details about the future of the project and the implications in other groups are unclear at this point.

3.3.2 CMS Data Management

Building on the previous experience acquired by the ATLAS experiment, the CMS Popularity Service has been developed to monitor the experiment’s data access patterns (i.e. frequency of data access, access protocols, data tiers, users, sites, CPU usage). The understanding of this framework provides a crucial step ahead towards the automation of data cleaning and data placement.

3.3.3 DIRAC

The DIRAC framework provides a complete solution for using the distributed computing resources of the LHCb experiment. DIRAC is a framework for data processing and analysis, including workload management, data management, monitoring and accounting (more details have been given in document [MS610]). LHCbDIRAC framework is the DIRAC extension specific to the LHCb experiment, which has been formally separated from DIRAC in order to streamline the implementation of features requested by the LHCb community.

The support of the LHCbDIRAC Data Management system (DMS) was started under the EGI-InSPIRE project in October 2010.

The LHCb DIRAC project has made considerable progress towards a sustainable future by trying to improve the model for developing, certifying and releasing software. Regarding software development, a modular approach is preferred, in order to avoid any duplication of code and to make the program more easily maintainable. For the process of certification and releases; a more systematic process has been recently defined where more careful testing is required from developers. This has the positive effect of reducing the time required for certification and for producing new releases. For next year, the objective is to consolidate these practices which have been recently introduced.

3.4 Monitoring and Dashboards

During PY1 the Experiment Dashboard applications have been heavily used by the LHCVOs, in particular by ATLAS and CMS. The system plays an important role for everyday operations, site commissioning activity and for distributed computing shifts. More than 200 CMS physicists daily access CMS Dashboard task monitoring in order to follow processing of their tasks on the distributed infrastructure. ATLAS DDM Dashboard is being actively used for monitoring of ATLAS Data transfers. All LHC experiments use the Dashboard SAM portal which provides information for evaluating site usability from the VO perspective.

One of the important goals of the Experiment Dashboard development is to offer common monitoring solutions for the LHC VOs. Achieving this permits the monitoring infrastructure to be simplified and contributes to its sustainability. All Dashboard development in 2011 and the beginning of 2012 followed this strategy. During the referenced period, the global transfer monitoring system which provides cross-VO monitoring of all data transfers on the WLCG infrastructure was prototyped and deployed on the preproduction server. New versions of the job monitoring applications were developed based on the common hBrowse framework. The hBrowse framework offers common client-side implementation of the user interfaces with generic look and feel and with the advanced plotting, sorting and filtering functionality. After successful validation by the user community, new versions of the job monitoring applications were deployed in production. The Site Status Board (SSB) which provides a common solution for monitoring of the distributed sites and services was redesigned in order to improve performance of the data queries and the quality of the user interface. New SSB versions were successfully validated by the LHC VOs, were deployed in production and received a positive feedback of the user community.

The LHC VOs rely on the Service Availability Monitoring service (SAM) for remote testing of the distributed sites and services. Therefore one of the important WLCG monitoring tasks during 2011 and the beginning of 2012 was migration to a new SAM architecture. The Experiment Dashboard provides a portal for navigating through SAM test results and site availability calculated. The new portal compatible with the new SAM architecture was validated and deployed in production.

Further development of the Experiment Dashboard system is being aligned with the LHC computing needs and will follow the requests of the LHC virtual organizations.

Support will continue to be provided through the Experiment Support group, where effort will concentrate on the framework and generic components, together with the experiments and sites that use these tools.

3.5 Persistency Framework

Persistency and detector conditions cover two essential areas for HEP data processing and analysis, namely the handling of the event data and detector conditions data of the LHC experiments. Event data, that record the signals left in the detectors by the particles generated in the LHC beam collisions, are generally stored in files, while conditions data, that record the experimental conditions (like voltages and temperatures) at the time the event data were collected, are commonly stored using relational database systems. In three of the LHC experiments (ATLAS, CMS and LHCb), some or all of these types of data are stored and accessed inside data processing jobs using one or more of the three software packages developed by the Persistency Framework [PERS-FRAME] project: CORAL, COOL and POOL.

The Persistency Framework software has been developed over several years (since 2003 for POOL, 2004 for CORAL and COOL) through the well established collaboration of developers from the LHC experiments with a team in the CERN IT department (now partly funded by the EGI-InSPIRE project), which has also managed the overall project coordination. Within the common project, the personnel pledged by ATLAS, CMS and LHCb, coming from a large number of institutes in several countries, have contributed to the development and continue to support the components used by their experiment. The common project, in particular the effort from the CERN IT Department (and from EGI-InSPIRE through it) only deals with components that concern (or, initially, that show the potential to concern) more than one experiment: in the past, individual experiments have already taken up the sole responsibility of components that have been proven to be relevant to that single experiment. This is in line with the more general focus on common solutions as a strategy for the sustainability of the HEP community. The usage of the three packages is periodically reviewed with the relevant stakeholders in the CERN IT Department and the experiments. In particular, the sustainability of the support model for POOL, CORAL and COOL, taking into account the upcoming termination of funding from EGI-InSPIRE, has been discussed within the context of the WLCG Technical Evaluation Groups in PQ4 2011.

Regarding POOL, these discussions have already successfully converged on an agreed sustainable support model. LHCb has already stopped using POOL, after replacing it by a new software layer, inspired from POOL but maintained internally, which is also able to read existing data stored in POOL format; as a consequence, LHCb no longer needs support from POOL through EGI-InSPIRE and its software has no pending dependencies.

ATLAS will continue to use POOL for as long as the 2012 production version of the ATLAS software is actively used, but it will no longer use it or need POOL support through EGI-InSPIRE for the 2013 production version, where a custom software package derived from POOL will be built and maintained by ATLAS as part of their internal software. A development version of the 2013 ATLAS software release already exists; this will eventually become the production version of the ATLAS software, by the end of 2012 or beginning of 2013.

CORAL and COOL, a sustainable support model for the future is still being discussed with all its users (ATLAS and LHCb for both, CMS only for CORAL) and other relevant stakeholders. While the software is by now mature in its development cycle, a large development and support effort (approximately 3 FTEs) is still required for user support, service operation and maintenance tasks. An important step towards sustainability was achieved in 2011 by transferring the full responsibility for CORAL and COOL release build and validation, from the EGI-InSPIRE funded IT team to the team in the Physics (PH) department that takes care of software infrastructure issues common to all LHC experiments. Understanding service operation issues for CORAL-based applications almost always requires detailed troubleshooting on the Oracle database servers where the data is stored; these tasks are performed by the Persistency Framework team in collaboration with the relevant database administrator team (funded by CERN IT).

The Frontier/Squid system is also relevant to CORAL. Both the software and its deployment are stable and mature, as they have been critical to the operations of the CMS experiment for several years. For CMS, the Frontier/Squid deployment is monitored and supported by a small team composed primarily of institute-contributed effort. The system is now becoming more critical also to the operations of ATLAS, which has recently adopted it to provide data distribution for remote analysis on the Grid. This is likely to require additional operational effort, which may eventually be managed in the context of the common project as this activity is relevant to more than one experiment.

3.6 Summary

HEP has been using EGI-InSPIRE effort to refactor its software by exploiting commonalities to reduce its long-term operation maintenance and support costs. Wherever possible these tasks are being transferred to established teams at CERN and by relying on a collaboration model of software development activity distributed across a number of different HEP related institutes. Sustainability will be achieved by reduced costs and leveraging the work of financially decoupled teams.

4 EARTH SCIENCES

4.1 *Introduction*

The fundamental topics in Earth Sciences are the understanding of the Sun-Earth system through observations and models, and the societal needs of the discipline.

Key areas in the Earth Sciences discipline include; natural resources identification and exploration, reaction to natural and human-induced changes, long term climatology – observations and modelling, multi-disciplines to observe and interpret one event, cross scientific domains (physics, chemistry, fluid mechanics, etc) and Civil Protection which involves risk assessment, disaster forewarning, impact studies and mitigation of natural hazards.

Furthermore there are international societies and programs such as the United Nations Environment Program which is a UN space-based platform for many organisations such as; Information for Disaster Management and Emergency Response, International Union of Geology and geosciences, Global Monitoring for Environment and Security programme, Intergovernmental Panel on Climate Change, and the European Geosciences Union.

4.1.1 Distributed Computer Infrastructure (DCI)

Resources come from ES related sides directly at NGIs and from, but not solely, EGI.

The partners of the VRC[3] are engaged in different projects or have institutional funding; large variety of funding and person-power conditions.

In terms of community building, there is COST action as well as other proposals to support the ES "Grid" community. There's also participation in new projects such as Verce[4].

Community tools consist of web presence, collaboration tools and portals. These are committed by core VRC members for at least the next few years.

For publications, see actual papers from; climate, hydrology, satellite data exploitation.

4.1.2 Support

Organisationally, collaboration has been achieved with GENESI-DR and ESGF (on going applications).

In terms of software, small self-contained open source tools with no central maintenance foreseen are in use, and there is organization of software exchange and knowledge between ES projects and/or researchers (e.g. 8 different ports of WRF[5] to gLite based grids, lots of complementary hydrological codes).

For generic VOs and services it depends on the evolution of the usage model.

4.2 *GENESI-DR*

The GENESI-DR infrastructure provides a standardized data discovery interface based on OpenSearch and metadata standards for a federation of data repositories. Meanwhile, the European project behind it (GENESI-DEC[6]) focuses on a central portal as an interactive entrance point, the usage on EGI

[3] Virtual Research Community

[4] http://www.verce.eu

[5] Weather Forecasting Model

[6] http://www.genesi-dec.eu

requires versatile clients such as a non-interactive, bulk oriented, tool. The work on such a command line client[7] that facilitates the usage in Grid jobs is ongoing. Because of changes to the underlying GENESI infrastructure, the tools have to be adapted. The command-line tool was extended with an ncurses based text user interface for interactive command-line usage over a shell, e.g. in the typical situation of a terminal SSH connection to a gLite User Interface. Documentation with a description of the functionality and a few standard use cases and examples to help users getting started with the utility has been prepared.

Additionally, a flexible web GUI was designed. It uses the OpenSearch interface provided by a running GI-cat instance. The interface leverages the OpenLayers capabilities for defining the geographical bounding box. The list of results is subsequently being used to generate search queries for the respective datasets such as AIRS, MODIS, GOMOS or GOME (mainly depends on the catalogues that are managed by GI-cat). Mostly, these datasets provide OpenSearch interfaces as well. The user defined search parameters from the first steps are then delegated to these second layer interfaces. The data is presented to the user, who is able to select the required files. Based on this step, a job description file to be submitted to Grid resources is generated.

4.3 Earth System Grid Federation

The Earth System Grid Federation (ESGF/ESG) is a distributed infrastructure developed to support CMIP5 (The Coupled Model Intercomparison Project, Phase 5), an internationally co-ordinated set of climate model experiments involving climate model centres from all over the world. Data access within ESGF is provided with two main services: OpeNDAP and GridFTP. A site that hosts these services is called a “Data node”. The team that works on ESG interoperability is developing and testing a scenario based on an application from IPSL which uses CMIP5 data (climate model data stored on the ESG). The MPI code of the application is now running successfully on EGI. An important activity was the implementation of a multi-threaded climate data transfer program to download the data from the ESG data nodes. One of the team members regularly participates in ESGF developer meetings.

The EGI Earth Science community is in contact with the FP7 project Virtual Earthquake and seismology Research Community e-science environment in Europe (VERCE), which aims at integrating a service-oriented architecture with an efficient communication layer between the Data and the Grid infrastructures, and HPC. A second novel aspect is the coupling between HTC data analysis and HPC data modelling applications through workflow and data sharing mechanisms. Discussions about effective and goal-oriented exploitation of Grid resources in the projects infrastructure have taken place. The projects members were suggested to register for the ESR general purpose Virtual Organisation to immediately be able to use computational resources. The setup of a dedicated VERCE VO or VRC is being discussed at present.

4.4 Summary

The Earth Science community is working on a few smaller development projects to improve access to Earth Science data infrastructures, mainly the exploitation of the GENESI-DR infrastructure and the Earth System Grid (ESG) through the use of community standards. For EGI users, the GENESI-DR infrastructure provides a standardized data discovery interface based on OpenSearch and metadata standards for a federation of data repositories. While, in the context of the European project behind it (GENESI-DEC), it focuses on a central portal as an interactive entrance point, its usage within EGI

[7] https://appdb.egi.eu/?#!p=L2FwcHMvZGV0YWlscz9pZD03MDg

requires versatile clients such as a non-interactive, bulk oriented tool, the provision of which will help sustain ES relations with ESG.As the developed software is rather focused and consists of atomic applications developed for specific use-cases of the Earth Science community, we fear that the prospective gains for other communities are rather limited. Nevertheless, the inquiries and technical developments may be of assistance, or even serve as groundwork for other communities. Although the specific technical conditions for the application areas, such as the data handling the tools deal with, discern seriously throughout the different communities, common requirements, issues, and principles could be identified. The underlying basics are however not part of this task, that is more practice-oriented. As the tools are self-contained, we foresee that they can survive as open source projects.

The investigation on authentication schemes, or put more generally, bridging infrastructure security functions, are of general interest. The development is aligned and communicated to both the ESGF as well as EGI. Members of the task have participated in identity management workshops and plan to continue to engage in further developments. We hope that this issue will be solved in a larger scale than the ES task of SA3, such as in the infrastructure projects, as they can be helpful in numerous situations.

5 LIFE SCIENCES

5.1 Introduction

Life Sciences are a very broad discipline. It ranges from fundamental research to direct clinical applications. Every potential patient may be interested in this discipline in some way or another as LS has a strong societal impact.

There is a strong competition between daily healthcare and preventive measures (including research), healthcare is increasingly becoming based on early diagnosis through human body data acquisition, modelling, and data analysis. LS are adapting towards the digital world with data now being acquired through digital sensors and health records being increasingly archived on digital storage facilities. Computer aided intervention and diagnosis assistance is becoming much more relied upon and key in many areas.

5.1.1 Distributed Computer Infrastructure (DCI)

The Foundation of the Life Sciences Grid Community (http://wiki.healthgrid.org/LSVRC:Index) consists of 5 NGIs (Dutch, French, Italian, Spanish, Swiss) as well as 5 VOs (biomed, lsgrid, medigrid, pneumogrid, vlemed) and 2 main projects (EGI-InSPIRE, Lifewatch ESFRI). Regarding the financial support for Life Sciences, it will mainly come through human resources dedicated by NGIs. Funding also comes from EGI-InSPIRE funding. Life Sciences are also represented by the HealthGrid association. This was set up in 2003 to support LS activities related to Grids. It also recently signed MoU with EGI.eu. However, recently the HealthGrid Association went bankrupt and will be dissolved shortly and new legal representation will need to be identified.

5.1.2 Support

Life Science services include; LSGC dashboard, LSGC support technical team, GReIC database interface service, Hydra encryption/decryption service, EBI core databases. The LSGC dashboard was under the responsibility of HealthGrid Association, however its development has been compromised with the end of the HealthGrid Association and is being transferred within the project. The LSGC support technical team is mostly funded by NGIs and is, for the foreseeable future, sustainable. The Hydra encryption/decryption service is currently supported by the French NGI and this will continue. The GReIC database interface service is not a LS-specific service, it re-uses effort invested internally. It receives funding through applying to national and international calls and through collaborative projects (e.g. Climate-G) and through future adoption by other communities. The EBI core databases are maintained through their own funding.

5.2 LSGC User Management Tools

The Life Sciences HUC steers the LSGC ("Life Sciences Grid Community" VRC[8]) effort to organize the community and deliver new services. A particular effort is invested in assisting users to better exploit the Grid and rationalizing Grid usage. In this context, several services to the HUC users have been provisioned:

- Web gadgets listing Life Sciences applications and community requirements were set up by User Communities Support Team have been added to the LSGC wiki;

[8] LSGC wiki, http://wiki.healthgrid.org/LSVRC:Index

- A HUC support service has been delivered. A technical team of expert users has been set up to address the difficulties reported by users on the VRC mailing lists or through GGUS. Bi-monthly phone conferences are scheduled and shifts have been organized to ensure that there is always a team on duty tackling the problems. See technical team wiki for details[9];
- The technical team pro-actively monitors the infrastructure health at a VO level, to identify the problems occurring. The French NGI deploys a dedicated Nagios server[10] for that purpose. New probes to monitor all VO SEs, WMSs and CEs were developed;
- On-line reporting tools easing the monitoring of SEs space management have been added to the technical team wiki page.

Thanks to the tools set up, a large-scale action to clean up all obsolete files left behind by years of Grid usage from a heterogeneous community with 200+ registered users is currently being conducted.

The development of a HUC users database and management tools to assist VO administrators in their daily task is planned. This service will interface to Virtual Organization Membership Service (VOMS) servers as well as the EGI applications database, to avoid replicating existing information. It will complement the VOMS and applications database with extra-information on the users and their affiliations. It will be interfaced to the LSGC Dashboard.

Despite the unexpected dissolution of the HealthGrid association, the LSGC organization proved to be resilient. In the future, technical services hosting requests will be made out to the NGIs to improve their sustainability.

5.3 GRelC

The GRelC service is a Grid database management service aiming at providing access and management functionalities related to relational and non-relational databases in a Grid environment.

Additional support to the HUC will be provided to address new user needs and requirements. In particular the UNIPROT and IAS in the Life Sciences domain will be further supported. A major goal regarding the support for the LS community will be to create in Lecce (SPACI node) "*a Grid database node hosting several data banks addressing LS needs*". Concentrating at the same place many different biological databases will be relevant for the LS community and could be also crucial to attract new users.

Integration of GRelC in the European Middleware Initiative release can be considered an important milestone towards sustainable software.

For more information please refer to the GReIC section of the HUC Software Roadmap: https://documents.egi.eu/document/746.

5.4 Hydra

Hydra[11] is a file encryption/decryption tool developed by EMI to enable the protection of sensitive files stored on Grid storage resources. The service comprises a distributed encryption key store, and

9 Biomed technical team wiki, http://wiki.healthgrid.org/Biomed-Shifts:Index

10 Biomed Nagios server, https://grid04.lal.in2p3.fr/nagios

11 Hydra service overview: https://twiki.cern.ch/twiki/bin/view/EGEE/DMEDS

client command lines that can (i) upload/fetch keys to/from the key store and (ii) encrypt/decrypt data files using these keys.

Currently, to deliver a production service, Hydra depends upon:

- Server's provision for Hydra key stores;
- Standard client packages installation on all EGI computing resources;
- Service monitoring and maintenance.

A 3-server based Hydra key-store will be deployed as a service for the life sciences community. This task, due in the first year of the EGI-InSPIRE project, has been delayed until gLite 3.2 (64 bit) servers could be provisioned and further slowed down by the end of the HealthGrid association which was supposed to host one of the servers. An attempt will be made to produce client packages compatible with the current middleware release and deploy them. If it proves impossible, it will be needed to wait for the first deployment of the EMI release on the EGI infrastructure. The Life Science Grid Community (LSGC) technical team will handle the service client installation and monitoring for more sustainability. The service monitoring procedure will be integrated in the Nagios box-based monitoring service offered by the LSGC technical team.

5.5 Summary

The LSGC relies on project based funding (such as EGI-InSPIRE to deliver the new tools and services it needs to run for its communities. For the sustainability of the services and resources used by the community it relies on the agreed collaboration of a few NGIs committed to supporting the various Life Sciences VOs.

6 ASTRONOMY & ASTROPHYSICS

6.1 Introduction

Astronomy & Astrophysics (A&A) is a worldwide, geographically distributed community that operates large experiments and observing facilities. A&A has new and increasingly sophisticated instrumentation facilities with enhanced capabilities and capacities. The discipline has developed so that in silico experiments and simulations modelling physical and chemical processes in the Universe complement traditional ways of conducting A&A. A huge amount of legacy software in terms of applications and libraries has been produced to support this analysis. High data volumes that are coming from experiments, simulations and observing facilities present challenges for the storage systems. There is a need for facilities to store and share software and data and a need for computing facilities to fulfil the data processing requests which is an extremely demanding job. There is also a need to handle data in a quick and efficient manner to enable the maximisation of the scientific return.

There are currently no user-friendly A&A-specific environments currently in place that are able to globally meet all user's expectations and to hide the complexity of the underlying DCIs. The A&A community can count now on the Virtual Observatory and on its infrastructure for a full exploitation of large astronomical data collections distributed worldwide.

Typically, funding for A&A research is allocated to international organizations or to national/local institutes by national governments on a regular basis. Funds coming from specific scientific or technological projects are in turn typically funded by national governments, European Commission, international organisations or by profit entities (business, industry, etc). However recent funding cuts imposed by national governments force communities to look for other possible funding channels (technological, transfer, etc).

6.1.1 Distributed Computer Infrastructure (DCI)

The A&A Grid/DCI community are currently funded through short/medium term national and international projects (e.g. EGI-InSPIRE and former EGEE projects). Funds typically come from Europe and from national governments. Recent project proposals to sustain the A&A Grid/DCI community by means of a general purpose, pan-European A&A SSC were unsuccessful. Single institutes and organizations now ensure a survival level of funds to support activities related to Grids/DCIs. NGIs play a key role in providing tools, services, training, dissemination and other important services. Past experience with A&A HUC has shown that setting up a unique large infrastructure doesn't work and a unique general purpose VRA failed in achieving its objectives.

Grid/DCI activities have to be carried out in the framework of well-focused scientific or technological projects with a dedicated amount of funds allocated to data processing aspects (with corresponding dedicated WPs). Multiple VRCs gathering a more restricted and well-focused community seems now the most promising and viable way to create sustainability for the community.

6.1.2 Support

The most relevant SA3 tool directly provided by A&A is VisIVO. Other activities of A&A in SA3 focus on the provision of requirements, use-cases, test-beds to EGI to create tools and services to meet A&A's needs.

6.2 Coordination of A&A

This activity, although not officially part of the work plan of task TSA3.5, is related to it given that one of the most important objectives of the task is the provision of requirements, use-cases and test-beds to EGI; they concern: a) interactivity between e-Infrastructures based on different technologies (Grid, HCP and Cloud); b) support to access and manage astronomical databases from Grid Infrastructures. To fully achieve this, objective people working in task TSA3.5 (who are also in charge of the coordination of the European A&A community at large in EGI) need to interact with as many European A&A research groups and Institutes as possible.

During PY2 the coordination of activities within the A&A community concerning the usage of DCIs by both small-scale and large-scale projects, in particular by the ESFRI projects such as SKA (Square Kilometre Array) and CTA (Cherenkov Telescope Array) has been intensified. People have been encouraged to contribute to the process of gathering requirements to be fed to EGI; other initiatives aimed at enhancing the use of e-Infrastructures were also carried out. A coordinating workshop[12] was organized at the Astronomical Observatory of Paris on November 7th 2011, where the major astronomical projects and research areas were represented.

During the workshop the current status of the A&A HUC was discussed. The most important topics raised from the discussion highlighted the need of activating multiple astronomical VRCs each of them aggregated around a big project or Institution that would be representative of one of the major A&A research areas. A unique catch-all VRC, in fact, has proven to not be suitable as it is unable to effectively represent the whole A&A community and would impose strong limitations for a number of key aspects like the requirements gathering process and the preparation and implementation of a long-term sustainability plan. All participants agreed that the major effort within the A&A HUC should now be focused on the identification of big A&A projects and Institutions for which the adoption of DCIs could be beneficial, so they might be good candidates to lead new astronomical VRCs. The process aimed at the creation of new VRCs is currently in progress with regular checkpoints scheduled to verify its progress.

6.3 VisIVO

During PY2 significant results have been achieved regarding the porting to the Grid of VisIVO[13] (Visualization Interface for the Virtual Observatory), a visualization and analysis software for astrophysical data. It consists in a suite of software tools aimed at creating customized views of 3D renderings from many types of datasets.

Now the first grid-enabled version of VisIVO service has been deployed. They are based on a specific grid-enabled library that allows users to interact with Grid computing and storage resources. The current version of VisIVO is also able to interface with and use the gLite Grid Catalogue. Although VisIVO has been conceived and implemented as a visualization tool for astronomy, it evolved in a generic multi-disciplinary service that can be used by any other community that needs 2D and 3D data visualization.

6.4 Grid and HPC

A working environment where Grid resources and HPC resources can be combined and used together is of utmost importance for many astronomical applications. Some typical A&A applications

---

[12] http://twiki.oats.inaf.it/twiki/bin/view/AstroVRC/AstroVRCWorkshop

[13] http://visivo.oact.inaf.it/index.php

that could improve their performance in such an environment have been identified: FLY[14] (a cosmological code developed at INAF-OACt) and Gadget[15] + Flash[16], the most common cosmological codes in Astrophysics.

The FLY code (a tree N-Body code) was executed on the gLite Grid. Some problems arise when running with a high number of N-Body particles. To overcome this problem a workflow is being defined which can be adopted when the code is used in challenging scenarios where data files several tens of gigabytes in size have to be handled.

To test cosmological simulations we are now in the phase of identifying and defining some use-cases; the goal of this activity consists in preparing the environment, the input file and the watchdog procedure to verify the run and the results.

Recently the activities related to Grid and HPC are carried out in close coordination with EGI and with IGI[17] (the Italian NGI) in Italy. A tight coordination with EGI and with NGIs, in fact, is crucial given that collaborations and agreements with PRACE and with other entities that operate and maintain HPC resources (typically supercomputing centers) are very important for this activity. One of the activities recently undertaken within the task TSA3.5 aims to introduce small-size HPC resources in Grid. This is achieved by installing and configuring HPC clusters (based on low latency/high throughput networks, HPC libraries and tools, modules and compilers) and then making the Grid middleware aware of these resources. The plan is to verify such small-size HPC resources and the related middleware aware version vs. the most popular cosmological applications mentioned previously, namely FLY and Gadget + Flash.

For this reason, in the next period the activity for HPC will continue in Italy in close coordination with all those communities involved in IGI that need HPC and MPI to efficiently run their applications.

6.5 Access to Databases for DCIs

Access to databases from DCIs and interoperability with the VObs (Virtual Observatory) data infrastructure remains one of the hot topics in astronomy. One of the most important related activities deals with the identification of use-cases and test-beds (both applications and complex workflows) that require simultaneous access to astronomical data and to computing resources. The data infrastructure of reference in astronomy is the VObs; end users should be able to access astronomical data Grid through VObs standards[18] and launch computational tasks on DCIs. In this context two key issues concern a SSO (Single Sign On) mechanism able to grant access to computing and data resources by means of a single authentication (users do not need to authenticate themselves multiple times) and tools/services to access astronomical databases federated in the VObs from DCIs.

The way of accessing databases from DCIs, GRelC[19] is one of the tools under evaluation in order to verify its ability to meet the most important requirements of the A&A community.

Another important activity related to the integration of data and computing resources is the work in progress to integrate in Grid the BaSTI[20] (A Bag of Stellar Tracks and Isochrones) Astronomical

[14] http://www.ct.astro.it/fly/

[15] http://www.mpa-garching.mpg.de/galform/gadget/index.shtml

[16] http://www.flash.uchicago.edu/site/

[17] http://www.italiangrid.it/

[18] http://www.ivoa.net/Documents/

[19] http://grelc.unile.it/home.php

Database and its feeding FARANEC code. A web portal was developed to make easier the submission of FARANEC code to the Grid. The portal allows the user to define a set of parameters and to simulate stellar evolutions on the Grid, without worrying about the technical details concerning the underlying Grid Infrastructure. The portal is based on P-GRADE web portal. Since the portal is designed for the submission of an arbitrary job on the Grid, it was necessary to do some low level "tricks & hacks" to make it able to fully satisfy our needs. The work aimed at implementing a new version of the portal based on the latest version of P-Grade is now in progress. It will be based on Web Services and this should enable a low level integration with the Virtual Observatory.

6.6 Summary

Long term sustainability of A&A resides within the organisation and instruments within the community. Common standards have allowed closer integration through community developed tools and applications running on EGI resources and have demonstrated the data analysis potential of distributed computing infrastructure in this domain. The engagement of the sustainable structure within A&A with EGI will remain a focus for PY3.

[20] http://albione.oa-teramo.inaf.it/

7 ADDITIONAL COMMUNITIES

7.1 *Dashboards*

The Experiment Dashboard system provides monitoring of the WLCG infrastructure from the perspective of the LHC experiments and covers the full range of their computing activities, such as data transfer, job processing, and site commissioning. In contrast to many other monitoring systems, the Experiment Dashboard is not coupled to any particular middleware, workload management, or data management systems. It is shared by several LHC virtual organizations and works transparently across various middleware platforms.

The Experiment Dashboard has proved to be an essential component of the Grid monitoring infrastructure. The LHC experiments rely heavily on the Experiment Dashboard for distributed operations, monitoring user computing activities, and providing user support. The Experiment Dashboard servers are accessed by several thousand distinct users per month. Among the Dashboard applications which are extensively used by the LHC community are Site Status Board, Job monitoring applications, DDM (Distributed Data Management) Dashboard. Experiment Dashboard applications adhere to a set of core development principles: common technology and implementation, loose-coupling to data sources, sharing of monitoring data, and user involvement in the development process. The aims of these principles are to reduce development and maintenance overhead, to allow applications to be easily adapted for use by multiple VOs, to enable reuse of monitoring data, and to ensure that applications meet user requirements. The Experiment Dashboard Framework provides a common foundation for the development of the monitoring applications and facilitates development and maintenance tasks.

The sustainability of the system is ensured by the described development principles and by the collaborative nature of the project. The core development and support team at CERN collaborates with the developers from the LHC experiments. There are also contributions from the institutes participating in the LHC project from Russia, Taiwan, India, and the UK. The LHC experiments also take part in operating the Dashboard services, for example as a part of ATLAS computing shift procedures.

The described changes in PY2 contributed to the sustainability task through decreasing the long-term maintenance and support effort.

In 2012, priority will be given to the common applications shared by multiple VOs. Among them is the Global WLCG transfer monitor that aims to perform common tasks carried out by the LHC virtual organizations in order to monitor their data transfers on the WLCG infrastructure, namely aggregation of the WLCG transfer monitoring statistics and exposing monitoring data via user interfaces and APIs. The first prototype of the system was developed and deployed for validation in 2011. During 2012, the functionality of the system will be extended based on feedback of the user community.

7.2 *Workflow and Schedulers*

7.2.1 Kepler and GridWay

As described in MS606, Kepler is a software application for the analysis and modelling of scientific data. Kepler allows scientists to create their own executable scientific workflows by simply dragging and dropping components onto a workflow creation area and connecting the components to construct a specific data flow; GridWay [GRIDWAY] is a Metascheduler that automatically performs all the **submission steps** and that also provides the runtime mechanisms needed for dynamically adapting the application execution.

By supporting GridWay, we are improving the sustainability plans of Kepler by providing enhanced capabilities and new options for its users. GridWay supports different middlewares, hiding those middlewares from the user or from any piece of software on top of GridWay. Hence, more computing infrastructures (middleware) could be used from within Kepler.

As it has been stated in previous reports, activities have been focused on establishing collaboration with other user groups which have shown interest in our work. Also, the dissemination activities carried out in different locations and meetings show the impact of our work to other communities. We have established contact with the Computer Chemistry users, and are in the process of helping out with the development of the first use cases using Kepler (with the Serpens suite). We also plan to approach other user communities. As for the short term dissemination activities, we plan to perform hands-on training sessions during events like the EGI User Forum.

7.2.2 SOMA2

SOMA2 is a versatile modelling environment for computational drug discovery and molecular modelling. SOMA2 is operated through a web-browser and it offers an easy access to third-party scientific applications. The SOMA2 environment offers a full scale modelling environment from inputting molecular data to visualization and analysis of the results, including the possibility to combine different applications into automatically processed application workflows. During PQ6 and PQ7, CSC has maintained and operated CSC's SOMA2 service.

On a longer timescale, our goal is to expand the selection of scientific applications in the SOMA2 service, and to integrate application services from different grids hosted by different virtual organizations into SOMA2. This should be set as an important milestone because from the end users point of view, this would make using scientific applications on different grids very easy and transparent. We will continue to maintain and operate the SOMA2 service and seek possible scientific applications to be added as part of the service. During all project years, we will support using SOMA2 services. Furthermore, the development of the SOMA2 gateway will continue according to feature roadmap, including possible feature requests from the user community, bug fixes, and other enhancements.

7.3 MPI

This sub-task has produced numerous MPI workbenches of increasing complexity with specific high impact on the Computational Chemistry and Fusion communities. These products include parallel implementations of Linear Algebra routines and parallelization of the CHIMERE application.

The CHIMERE parallelization model can be reused in many other scientific domains including, among others, the Earth Science and Astronomy & Astrophysics communities. Improvements to the MPI documentation[21] were made in PY2. The EGI Wiki now provides the definitive source of information for MPI support. Documentation on the ARC and UNICORE MPI support still require extra attention. The current release of the UMD gLite-MPI 1.2.0 package was released in Project Month 21. The release of the UMD glite-WMS and the glite-MPI products are expected to contain significant middleware changes required to ensure the correct support for many MPI job types. For example, these include:

[21]https://wiki.egi.eu/wiki/MPI_User_Guide

- User defined allocation of processes/nodes
- Direct Support for OpenMP codes.

Furthermore, end-user support for both MVAPICH (an MPI implementation with advanced Infiniband networking) and support for generic parallel job support is a high-priority in PY3.

As part of User Community engagement effort, the MPI team will regularly survey Virtual Organisations, Users and Site administrators for critical feedback. This will also act as a means to gather information about current deficits and future requirements. The newly founded NA2 driven MPI Virtual Team (MPI-VT) is expected to lead to a concentrated drive to weed out and fix problems observed with some large production-level MPI jobs. This shall create a greater demand on the SA3 MPI support team, led by CSIC and TCD.

The MPI task has not yet fully tackled the issue of sustainability. The MPI community itself has two mature and major open-source framework implantations: MPICH-2 and OpenMPI. These are both well supported and are expected to be maintained and further developed in the medium term (five years). Both are also heavily used by the High Performance Computing community. The two frameworks are also actively developed, and produce several new releases per year. Methods to widely support generic parallel jobs, GPGPU, and cloud integration may be potential research projects for further exploitation on distributed computing infrastructures.

8 CONCLUSIONS

The activities within SA3 encompass many diverse activities in their support of the HUCs. Some of these have matured into activities where their sustainability plans and future are clear. Other activities are still maturing and have yet to find a clearly sustainable future after PY3. The support of these activities will be a focus for PY3.

9 REFERENCES

MS112	Quarterly Report 5:May 2011 – July 2011 https://documents.egi.eu/document/723
MS113	Quarterly Report 6:August 2011 – October 2011 https://documents.egi.eu/document/881
MS114	Quarterly Report 7:November 2011 – January 2012 https://documents.egi.eu/document/999
MS609	HUC Contact points and the support model: https://documents.egi.eu/document/419
MS610	Services for High Energy Physics: https://documents.egi.eu/document/540
MS611	Services for Life Sciences: https://documents.egi.eu/document/683
MS612	HUC Software Roadmap: https://documents.egi.eu/document/684
MS614	HUC Software Roadmap and Sustainability Plan: https://documents.egi.eu/document/746
D6.3	Annual Report of PY1: https://documents.egi.eu/document/312
D6.4	Capabilities offered by the HUCs to other communities: https://documents.egi.eu/document/472
TEG REPORTS	See http://indico.cern.ch/conferenceDisplay.py?confId=158775.
WLCG 2012	See https://indico.cern.ch/conferenceDisplay.py?confId=146547.