SEVENTH FRAMEWORK PROGRAMME
Capacities specific program - Research Infrastructures
FP7–INFRASTRUCTURES–2010-2

Project acronym: *CHAIN*
Project full title: Co-ordination & Harmonisation of Advanced e-Infrastructures
Grant agreement no.: 260011
Start date of project: 1 December 2010
Duration: 24 months

<div style="border:1px solid black;">

Deliverable 3.3
"Trans-continental scientific and technical communities updated"

</div>

Version 0.7

Due on: 31/05/2012   Submitted on: 31/05/2012

Organisation name of lead beneficiary for this deliverable: CIEMAT
Dissemination Level: Public

Abstract:

This deliverable presents the results of the investigation made regarding the existing and emerging scientific and technical communities that already have trans-continental activities or have potential to have them in the future (CHAIN D3.1), and the work performed during last year since the candidates for testing the roadmap have been identified in the first trimester.

Document log

Issue	Date	Comment	Author
V0.1	07/05/2012	First draft	R. Muñoz and R. Mayo (CIEMAT)
V0.2	15/05/2012	Small Modifications	F. Ruggieri (INFN)
V0.3	29/05/2012	Small formatting	R. Mayo (CIEMAT)
V0.4	29/05/2012	Small formatting	S. Qian (IHEP)
V0.5	30/05/2012	Review	T. Banda (UbuntuNet)
V0.6	26/02/2013	Text revision	R. Mayo (CIEMAT)
V0.7	29/05/2012	Small formatting	O. Prjnat (GRNET), R. Mayo (CIEMAT)

Table of contents

1. **Executive Summary**

The first study of the state-of-the-art in Grid developments across the world regions (those that have direct relationship to the European Grid activities) performed by the CHAIN project, was carried out throughout 2011 (approximately, the first year of the project). This study was done by means of an on-line survey that aimed to overcome the lack of indexed and exploitable information (see D2.1 "State of the art questionnaire"); and specifically targeted to obtain information about the scientific communities - the so called VRCs - for standardising the Grid interactions and developments between those groups in Europe, Africa, Asia and Latin America.

As a result, CHAIN delivered D3.1 "Trans-continental scientific and technical communities", which reported on the communities surveyed. This document describes the advances made since the publication of D3.1, specifically regarding VRCs. Summarising, it can be mentioned that all of them are keeping a seamless collaboration with CHAIN, but specifically we point out:

- We-NMR has signed an MoU with CHAIN, and is studying the SG paradigm and has been provided new contacts by CHAIN;
- WRF4G has signed an MoU, has defined a work plan for the integration of the SG paradigm with DRMAA, and has been provided new contacts by CHAIN;
- jModelTest/ProtTest has signed an MoU with CHAIN, is working on the integration of its sequential and distributed versions into the SG paradigm, and has been provided new contacts by CHAIN;
- INDICATE (DC-NET) has signed an MoU with CHAIN and has adopted the SG paradigm;
- DECIDE has signed an MoU with CHAIN and has adopted the SG paradigm;
- Climate Change VRC is being formed, since several researchers have shown their interest and are looking for synergies with WRF4G;
- LSGC has signed an MoU with CHAIN;
- SuperB is being formed (a VO is already operational), looking for researchers outside Europe, and the MoU with CHAIN is under preparation.

Around all these VRCs, a first draft for a model for evaluating services required by these communities was proposed in D3.2 "Road-map of trans-continental e-infrastructures for virtual communities", which also described the platform adopted by CHAIN as a proof of concept for testing all the collected and processed specificities: the SG.

The new information described in this document has been partially obtained from events organised by CHAIN during this reporting year (Jun 2011-May 2012). This includes the direct interaction between VRCs and Infrastructure Providers that took place in "Resource Infrastructure Providers meet VRC (II)", held as part of the EGI Technical Forum (Lyon, September 2011) and "VRCs on EGI and Regional Infrastructures", held as part of the EGI User Forum (Garching, March 2012). Other meetings supported by CHAIN include the UbuntuNet-Connect Conference (Nairobi, November 2011), the e-AGE 2011 conference (Amman, December 2011), the "Sustainability & Interoperability - Step 2" held as part of ISGC 2012 (Taipei, March 2012), as well as the EGI-CHAIN meeting (Amsterdam, January 2012). The latter were also very useful for CHAIN to collect valuable feedback from VRCs and their relationship with sustainability issues.

In addition to these activities, WP3 has submitted a new shortened survey to the National representatives in March 2012, which only contained the VRCs-related part of D2.1 and a list of the identified research communities. The aim of was to get updated information about the status of e-Science in the different countries as well as to obtain new coordinates of researchers interested in the VRCs identified. The former will be used to redefine the road-map of services and the latter to enhance the collaboration with the associated communities.

In this sense, continuous contacts with the representatives and/or members of these VRCs have been kept in order to be up-to-the-date regarding their advances and initiatives and to get their feedback about the recommendations proposed in CHAIN.

Last but not least, WP3 has been working on the recommendations made by the panel after the first project review and has reported about the quality metrics laid out in the DoW: all of these have been fulfilled.

2. Introduction

e-Science has been strongly funded over the last decade by the EC through FP6 and FP7. Many computational developments have been supported and they have been transferred to other regions (mainly Asia, Latin America and the Mediterranean Regions) through projects which linked different world regions with the European e-Infrastructures actions. This effort has been carried out in three areas:
- Research and Education Networks;
- Computational Infrastructures; and,
- Virtual Research Communities.

Nevertheless, a slow progress has been registered until now to link e-Infrastructures at the intercontinental level due to their specific requirements that depend on the targeted region – despite the fact that basic operational, organisational and technological know-how transfer/exchange has been performed (specifically valid for Grid computing).

The CHAIN project is working on coordinating and leveraging the efforts made over the past 6 years to extend the European e-Infrastructure (and particularly Grid) operational and organisational principles to a number of regions in the world, mainly those identified for this purpose by the EC. The CHAIN project aims to elaborate a strategy and define the instruments in order to ensure coordination and interoperation of the European Grid Infrastructure with those emerging in other regions of the world (Africa, Asia, Latin America, Mediterranean and Middle East).

This vision has been structured in several lines of action: study of the state of the art of the current e-Infrastructures and e-Science status per region, including needs and commonalities of these computing platforms and VRC requirements and services. CHAIN WP3, "Present and emerging needs of trans-continental scientific communities", devotes its framework to the latter. As a first step, a coordinated data collection on the existing Grid state of the art across the world was carried out (see D2.1 "State of the art questionnaire"[1]).

The first result of such a study from the WP3 point of view was the identification of several VRCs that were of interest for several reasons:
- they covered several scientific fields;
- they showed a clear sustainability plan;
- they had presence in the Regions of interest to CHAIN;
- they presented a different level of maturity regarding Grid know-how; and,
- they ranged in their level of maturity: from already formed to nascent, needing strong CHAIN support.

Working that way, CHAIN pursued to get a broad feedback from different disciplines that was expected to produce a real general first version of the road-map of services required by the VRCs to the infrastructure providers.

---

[1] CHAIN D2.1, http://documents.ct.infn.it/record/486/files/CHAIN-D2.1-V01-Final.pdf?version=1

The first list of identified communities can be found in D3.1 "Trans-continental scientific and technical communities" [2] (which was subsequently expanded), and the first service recommendations in D3.2 "Road-map of trans-continental e-infrastructures for virtual communities" [3]. This document presents, in the 18th month of the project, the updated information of the identified VRCs and the actions that have been performed jointly with them during the last year (D3.1 was submitted in 2011).

2.1. Purpose

The purpose of this deliverable is to describe the updated information regarding the collaboration between CHAIN and the emerging scientific and technical communities that have trans-continental activities and have been identified by the project during the first 18 months of its lifetime. Thus, the WP3 actions devoted to enhance the liaison with the VRCs and the results that have been produced are presented, as well as a summary of the road-map of services required by these communities. In addition, a brief description of the platform selected by CHAIN to test the road-map, i.e. the SG paradigm, is also documented.

2.2. Glossary

CHAIN	Co-ordination and Harmonisation of Advanced e-Infrastructures
DoW	Description of Work – Annex I to the GA
DRMAA	Distributed Resource Management Application API
EC	European Commission
EGI	European Grid Initiative
EGI-InSPIRE	European Grid Initiative-Integrated Sustained Pan-European Infrastructure
EPIKH	Exchange Programme to advanced e-Infrastructure Know-How
FP6/FP7	European Commission's Framework Programme Six / Seven
GA	Grant Agreement
HEP	High Energy Physics
HPC	High Performance Computing
JRU	Joint Research Unit
KoM	Kick-off Meeting
MoU	Memorandum of Understanding
MS	Milestone
NREN	National Research and Education Network
ROC	Regional Operation Centre
PMB	Project Management Board
SAGA	Simple API for Grid Applications
SG	Science Gateway
VO	Virtual Organization
VRC	Virtual Research Community
WP	Work Package

---

[2] CHAIN D3.1, http://documents.ct.infn.it/record/500/files/CHAIN-D3.1-V06.pdf
[3] CHAIN D3.2, http://documents.ct.infn.it/record/510/files/CHAIN%20D3.2-V03.pdf

3. Direct actions performed to enhance the VRCs collaboration

This section describes the actions that have been carried out by CHAIN since the first VRCs were identified. This includes the report on how the collaborations with them have been enhanced, which steps have been taken in order to support new initiatives that were not previously formed, which and what kind of events have been organised in order to get their feedback, which contacts with e-Science actors have been kept, and what new communities have been found.

All this information is related to a specific period from June 2011 to May 2012 since D3.1[2] was submitted one year ago. This information is also related to a concrete scenario: to closely collaborate with VRCs that are related to different scientific fields, so they identify various requirements, and with several degrees of expertise in what e-Science (and specifically Grid) concerns. Nevertheless, an aspect is present in all of them: their sustainability and their presence and interest in every Region targeted by CHAIN.

Working this way, CHAIN expects to benefit from a wide range of scenarios in order to test and validate its proposed model (the major and final goal of WP3) with several communities. The first outcome for that model has been the publication of D3.2[3], which describes a first draft for the specificities and capabilities that VRCs require from the infrastructure providers. It could be assumed that collaborating with similar communities demanding a (broad) common portfolio of requirements would make easier the adoption of standards or the way that a road-map of services could be defined. On the other hand, further consequences and adoption of policies or recommendations that could be derived from such a work would have for sure a scarce impact.

3.1. Precedents

As stated in the previous CHAIN deliverables, WP3 started working from the KoM[4] held in Rome in December 2010. Some communities were approached then, and the VRC part of the Regional and National questionnaires contained in D2.1 "State of the art questionnaire"[1] looked for information about the scientific disciplines of interest, and already proposed the identified VRCs in order to look for researchers working in those fields.

With the analysis of the collected results of such a survey (see D2.2 "Interoperability and interoperation guidelines"[5]), a search on the web about EU funded projects, the aim of which were aligned with the CHAIN objectives, and the organisation of events, WP3 was able to define an internal schedule for achieving its objectives as declared in the DoW.

The first events promoted by CHAIN and already described in D3.1[2] were the CHAIN & GISELA workshop "Resource Infrastructure Providers meet VRC"[6] held as part of the EGI User Forum 2011 and the "Conference on the Role of e-Infrastructures for Climate Change Research"[7], where several communities and e-Infrastructures providers presented to each other their requirements as well as what they have to offer. It can be mentioned that several

---

[4] CHAIN KoM http://agenda.ct.infn.it/conferenceDisplay.py?confId=464
[5] CHAIN D2.2, http://documents.ct.infn.it/record/506/files/CHAIN_D2.2_V1.8.pdf
[6] CHAIN & GISELA workshop held within EGI User Forum 2011 'Resource Infrastructure Providers meet VRC', http://agenda.ct.infn.it/conferenceDisplay.py?ovw=True&confId=561
[7] Conference on the Role of e-Infrastructures for Climate Change Research 2011, http://users.ictp.it/~smr2238/

communities and initiatives attended these events and there was a fruitful exchange of information.

As a result, the first list of identified VRCs was completed: WeNMR, WRF4G, jModelTest, DC-NET/INDICATE, DECIDE and Climate Change. It is worth mentioning at this point that WP3 adopted the VRC definition proposed by GISELA[8]: VRCs range from a group of researchers working together, sharing the means to carry out common investigations such as distributed resources (instruments, computing, storage and software tools), up to large international well-structured collaborations. VRCs cover small user communities (several groups with or without Grid knowledge) up to large collaborations (looking for more computing resources and/or advanced services).

3.2. Follow-up of WP3 milestone "MS07 Agreements with reference communities signed"

WP3 had to fulfil during the first nine months of the projects its three milestones:
- "MS05 - Call for interest for reference communities";
- "MS06 - Shortlist of reference communities ready"; and,
- "MS07 Agreements with reference communities signed".

All of them were completed in due time (see D3.1[2] and D3.2[3] for details about them), but in order to enhance the collaboration between CHAIN and the identified VRCs, work has been carried out to establish an official framework where synergies were easily identified.

Thus, after the signature of the MoU with WeNMR and WRF4G that fulfilled MS07, new agreements have been signed:
- jModelTest initiative, signed on February 21st 2012;
- LSGC initiative, signed on March 27th 2012;
- INDICATE project, signed on March 28th 2012; and,
- DECIDE project, signed in April 13th 2012.

These six documents can be found in the CHAIN Document Server[9] and more detailed explanations about these VRCs can be found in Section 4.

With this well established scenario, the validation model, which will reflect the road-map of services required by the VRCs and how the infrastructure providers are providing them, is expected to be properly tested.

3.3. Summary of the first version of the road-map and recommendations

Once the communities were identified in May 2011, the next step was to make a draft of the requirements that these VRCs had regarding the infrastructure. Based on the CHAIN deliverables and outcomes, and from the information collected by means of the events held and the direct contacts with the coordinators of these initiatives, D3.2 "Road-map of trans-continental e-infrastructures for virtual communities"[3] was submitted to the EC. This document describes such a road-map of recommendations (since the reader can consult in the CHAIN web page this deliverable, in this subsection only the structure will be documented).

D3.2 firstly describes the capabilities that would be of use for the VRCs if they were fully provided by the infrastructure providers; thus, thirteen (13) recommendations are presented

---

[8] GISELA D3.1 http://documents.gisela-grid.eu/collection/WP3?ln=en
[9] CHAIN MoUs, http://documents.ct.infn.it/collection/CHAIN%20-%20MoUs?ln=en

from the interoperation and interoperability point of view. Next, fifteen (15) items describing the status of the infrastructures on November 2011 are described.

Subsequent subsections are devoted to specifically document the road-map of recommendations for existing and emerging e-Infrastructures for VRCs. Such recommendations are grouped in three major areas: Technical, which counts for thirty two (32) items; Training, administration and use, which describes six (6) items; and, Collaborative scenario, which presents nine (9) items.

Regarding the way how providers are giving response to the VRC requirements, the relevant presentations were made in the "Heavy Users Community", "VRC" and "Portal technologies" sessions held within the EGI Technical Forum in Lyon (September 2011) and also the CHAIN & GISELA workshops. In these presentations, several Infrastructures presented their catalogue of services.

3.4. Actions performed during the reporting period

In the last year (from June 2011 to May 2012), direct actions have been performed by WP3 in order to enhance the collaboration between the project and the VRCs, all of them aligned with the general CHAIN strategy.

As a major and continuous task, conversations with the identified communities were carried out in order to get updated information about how they were evolving. In addition, the following actions were also promoted.

After the events already mentioned [4,6,7], the following ones were also held as part of major conferences, profiting in this way of a major international attendance. Thus, during the EGI TF 2011 held in Lyon (France), the workshop "Resource Infrastructure Providers meet VRC (II)"[10] was organised and even a third phase was held in Garching (Germany) as part of the EGI UF 2012 in the session "VRCs on EGI and Regional Infrastructures"[11].

Besides these workshops directly devoted to the interaction between CHAIN and the VRCs, the project organised a second high level conference held in Amman (e-AGE 2011[12]), several workshops as part of the UbuntuNet-Connect Conference 2011[13] held in Nairobi (Kenya), and the events on interoperation and sustainability issues as the ones held as part of ISGC 2011[14] and 2012[15] in Taipei (Taiwan) or the meetings with EGI[16] in Amsterdam (Netherlands). From all of them, WP3 has extracted useful information.

Another important action performed during this reporting period to support VRCs was the signature of MoUs that have been described in Subsection 3.2.

In order to present a scenario in as much detail as possible, the CHAIN survey has been critical. The overall analysis of the survey can be found in D4.1 "Specificities of the various

---

[10] CHAIN & GISELA workshop held within EGI Technical Forum 2011 "Resource Infrastructure Providers meet VRC (II)", http://agenda.ct.infn.it/conferenceDisplay.py?confId=561

[11] CHAIN & GISELA workshop held within EGI Community Forum 2012 "VRCs on EGI and Regional Infrastructures", http://agenda.ct.infn.it/conferenceDisplay.py?confId=761

[12] E-AGE 2011, http://eage2011.asrenorg.net/

[13] UbuntuNet Connect 2011, http://www.ubuntunet.net/uc2011_cfp

[14] CHAIN workshop at ISGC 2011, http://agenda.ct.infn.it/conferenceDisplay.py?confId=553

[15] CHAIN workshop at ISGC 2012, http://agenda.ct.infn.it/conferenceDisplay.py?confId=704

[16] CHAIN-EGI meting, http://agenda.ct.infn.it/conferenceDisplay.py?confId=701

regional e-Infrastructures"[17] and the specific part for VRCs in D3.2. Later on in March 2012, the project published D4.2 "Report on the applicability of the preliminary results of the Organisational Study"[18], where updated information can be found on the general plan to test such a study and on the status of the VRCs. As a wrap-up summary, CHAIN has recently published on its website the status of the recommendations.

Beyond concrete CHAIN work, the EGI report "Seeking new horizons: EGI's role for 2020" has been also consulted. All these documents have been very useful for WP3 and even the activity itself has participated with contributions.

Nevertheless, and due to the evolving character of e-Science, WP3 has sent its related part (see Subsection 3.3 in D3.1[2]) of the CHAIN general survey to the National representatives in order to get updated information from different countries. The first e-mail asking for such collaboration was sent on April 25th, 2012 and further requests have been sent weekly to those unresponsive contacts. As it has been stated, this e-mail contained a file where the pre-filled WP3 part of the CHAIN survey was present in order to get updates, but it also had a specific demand for reporting researchers working in the scientific fields covered by the eight (8) CHAIN collaborative VRCs.

Moreover, the road-map of recommendations, i.e. the list of services appearing in Subsection 4.3 of D3.2[3], was also attached to the mentioned e-mail in order to get the feedback from the National representatives. The same was sent on the same day to the coordinators of the VRCs supported by CHAIN in order to get their feedback.

For improving the presence and reach of CHAIN in the targeted Regions, the project is collaborating with the Abdus Salam International Centre for Theoretical Physics (ICTP) in order to get a new list of contacts from researchers that could be interested in joining the communities promoted by CHAIN. Since ICTP is mainly focused on advanced scientific expertise in the developing world, a profitable liaison with it is expected in the near future building on the successful organization of the "Conference on the Role of e-Infrastructures for Climate Change Research"[7] in May 2011.

The analysis of the new collected information from National representatives, VRCs coordinators and ICTP contacts is expected in the near future.

Last, but not least, the CHAIN plan to address the reviewers' recommendations from the first project review has been a new incentive to enhance the impact of WP3 activities. From the VRC point of view, a strategy to approach stakeholders and uptake the recommendations addressed in D2.2 has been defined.

A third recommendation, "Quality Metrics", is detailed in this document in Section 5 in what regards the scientific communities. WP3 has collaborated in the new outreach activities, which fulfil the fourth one, and has contacted the two proposed projects JISC-Engage and e.nventory. However, both are about to finish and neither extension nor second phases are, to our knowledge, planned to be funded. So, CHAIN has focused on the legacy they transfer in order to include it in the CHAIN Knowledge Base.

---

[17] CHAIN D4.1, http://documents.ct.infn.it/record/499/files/CHAIN_D4.1%20V1.3.pdf
[18] CHAIN D4.2, http://documents.ct.infn.it/record/515/files/CHAIN_D4.2%20V1.pdf

3.5. The Science Gateway paradigm

Using Grid is not always straightforward. Several barriers can be identified, but undoubtedly one major drawback is the relatively high level ICT skills requirement for beginners and the way of accessing the Grid (authentication & authorisation). Users have to cope with complex security procedures, execution scripts, job description languages, command line based interfaces and the lack of standards. This makes the learning curve very steep and keeps non IT-experts away.

TeraGrid project defined a Science Gateway (SG) as a community-developed set of tools, applications and data that is integrated via a portal or a suite of applications, usually in a graphical user interface, that is further customized to meet the needs of a specific community. This way, it would be possible to abstract the final user from the technological complexity.

The second step is to get an easier access to the SG. The adoption of robot-certificates for managing the jobs to be executed and the authorisation to get into the SG by means of Identity Federations (see for example eduGAIN[19]), which count on millions of users (so, for example, any researcher belonging to a R&D Centre or to a University can immediately login by his/hers own credentials), is an outstanding asset for this paradigm.

Last action is then to develop a computational framework where available and new developments could be coupled (added) as independent linked modules. Doing so, it would be easier to integrate new applications and codes in the framework.

Because of all these reasons, SG is supposed to provide a short term answer to interoperability when several different e-Infrastructures have to cooperate in order to fulfil the requirements of intercontinental VRCs and is being proposed and fostered by CHAIN.

The framework for SG is fully web-based and adopts official worldwide standards and protocols, through their most common implementations. These are:

- The JSR 168 and JSR 286 standards[20] (also known as "portlet 1.0" and "portlet 2.0" standards);
- The OASIS Security Assertion Markup Language [21] (SAML) standard and its Shibboleth and SimpleSAMLphp implementations;
- The Lightweight Direct Access Protocol and its OpenLDAP implementation[22];
- The Cryptographic Token Interface Standard[23] (PKCS#11) standard and its Cryptoki implementation; and,
- The Open Grid Forum (OGF) Simple API for Grid Applications[24] (SAGA) standard and its JSAGA implementation.

In addition, SG is built using the Liferay portal framework and can be downloaded and installed by a virtual machine containing the development environment and examples of basic template portlets that can be customised to integrate specific applications.

---

[19] eduGAIN, http://www.geant.net/service/edugain/pages/home.aspx

[20] JSR 168 and JSR 286, http://developers.sun.com/portalserver/reference/techart/jsr168/

[21] SAML, N. Ragouzis *et al.*, Security Assertion Markup Language (SAML) V2.0 Technical Overview. OASIS Committee

[22] OpenLDAP, B. Arkills, LDAP Directories Explained: An Introduction and Analysis. Addison-Wesley Professional. (2003)

[23] PKCS#11, http://www.rsa.com/rsalabs/node.asp?id=2133

[24] SAGA, https://forge.ogf.org/projects/saga-rg/

More detailed technical information about SG and its current status about connections to different middleware and integration with different initiatives can be found in Subsection 4.5 of D4.2[18]. In addition, the CHAIN web page offers continuous updated information about all these topics on the "Applications" section[25].

In this document, Section 4 describes how the identified VRCs are adapting the SG paradigm.

[25] "Applications" Section in the CHAIN webpage, http://www.chain-project.eu/applications

4. Updated list and status of scientific and technical communities

This section describes the scientific and technical communities that have been selected by CHAIN in order to propose a possible model of services to be implemented by the e-Infrastructures and the new information that has been produced regarding these, since the submission of D3.1[2].

These communities have different levels of Grid technical maturity, and are structured differently, ranging from well-established communities, even funded by external entities, to a group of users who employ a common application or have a synergy in their scientific interest. This heterogeneity is as an asset for CHAIN in order to propose a new portfolio of services that could fit a broad range of groups. Even more, their ICT interests move from applied Science to Humanities, so researchers with different scopes can be addressed. A risk actually lies in this approximation: but only acting on similar and well-developed Grid initiatives (either from the same scientific area or not) would make the final proposed model useless for many people and/or e-Infrastructures providers.

In addition, they have also been selected because of their sustainability and widespread continental presence, which is a must for CHAIN.

The following order of the VRCs is chronological on a first-contacted first-served basis.

4.1. We-NMR

We-NMR (http://www.wenmr.eu/) is a project which aims to optimize and extend the use of the NMR and SAXS research infrastructures through the implementation of an e-Infrastructure in order to provide the user community with a platform integrating and streamlining the computational approaches necessary for NMR (Nuclear Magnetic Resonance) and SAXS (Small Angle X-ray Scattering) data analysis and structural modelling.

An MoU was signed with this initiative[9] on September 21st, 2011 and its presence has been continuous in the workshops organised by CHAIN during the reporting period as a continuation of the first contacts held in the KoM. Several discussions have been held with the coordinators of WeNMR and their feedback has been very useful for the drawing-up of D3.2[3]. Once this document was published, it has been presented in several forums, but the summary of its recommendations has been directly analysed together with WeNMR in order to improve the quality of such recommendations – in order to appear in D3.4 by the end of 2012.

By means of the CHAIN WP2 survey, new contacts of researchers interested in the areas covered by WeNMR were provided to the VRC. The countries where these scientists were settled were Burundi and Costa Rica. As aforementioned in this document, a new round of e-mails has been sent by CHAIN WP3 to the National e-Science representatives, so new researchers are expected to join WeNMR, and this information will be part of D3.4.

WeNMR counts on an own-developed web access to submit jobs to the Grid, but discussions have been held to test and adopt the SG paradigm or benefit from its methodologies in order to improve the WeNMR submission.

4.2. WRF4G

WRF4G (http://www.meteo.unican.es/es/software/wrf4g) is a Grid version of the well-known Weather Research and Forecasting (WRF) modelling system application. It is widely used by the meteorological agencies and many other groups in the Earth Science domain. Its Grid-based version has increased the resources where tasks such as Idealized simulations, Regional and Global applications, Parameterization and Data assimilation research or Forecast and hurricane research can be performed.

An MoU was also signed with this initiative[9] on September 19th, 2011 and its presence has been continuous in the workshops organised by CHAIN during the reporting period too, as a continuation of the first contacts held in the KoM. Several meetings have been held with the coordinators of WRF4G and their feedback has been useful in drawing up D3.2[3]. The most important information has been also directly consulted with WRF4G representatives in order to improve the quality of recommendations to appear in D3.4 by the end of 2012. Even more, joint activities are foreseen for 2012.

By means of the CHAIN WP2 survey, new contacts of researchers interested in the scientific applications of WRF4G were provided to the VRC. The targeted countries were Burundi, China, Costa Rica and Cuba.

WRF4G has been working on a similar basis in order to port to the Grid the CAM model, so a new Grid release called CAM4G[26] is available. CHAIN has also supported this new version and has disseminated it in the events it has organised. A direct liaison between the WRF4G initiative and the Climate Change community is being supported by CHAIN (see Subsection 4.6).

As also described in the previous Subsection, the new round of e-mails was sent by CHAIN WP3 to the National e-Science representatives, looking for new researchers interested in WRF4G and CAM4G, which will be documented in D3.4.

WRF4G relies on a command line method for the submission of jobs, so its incorporation to the SG paradigm has been foreseen. Nevertheless, since WRF4G is mostly based on GridWay metascheduler[27] (which distributes tasks by means of DRMAA) and SG works with SAGA[24], an actor for communicating both methodologies should be developed. This work has been already planned by the WRF4G and CHAIN initiatives and is expected to be finished by the end of 2012.

WRF4G has kept disseminating CHAIN plans and outcomes inside the Climate Change CORDEX project[28].

4.3. jModelTest / ProtTest3

These two applications belong to the Life Sciences domain and, in particular, to the Evolutionary Biology. Both are freely available on-line (http://darwin.uvigo.es/) for the statistical selection of best-fit models of nucleotide substitution (jModelTest) and amino-acid (ProtTest3) replacement for a given set of aligned sequences. Thus, many researchers

---

[26] CAM4G, http://meteo.macc.unican.es/es/node/73051

[27] GridWay, http://www.gridway.org/

[28] CORDEX initiative, http://www.meteo.unican.es/en/projects/CORDEX

interested in molecular systematics, phylogenetics, phylogenomics, molecular evolution and/or bioinformatics use them continuously all around the world.

ModelTest has around 30,000 registered users world wide while jModelTest (the Java version) has around 11,000 and ProtTest around 5,000, i.e. almost all countries in the world have scientists who are using one of these tools.

CHAIN signed an MoU with this initiative on February 21st, 2012 and has also supported it by implementing distributed versions of both codes that can be executed either on local clusters or Grid[29,30]. Thus, the CHAIN members have been able to directly provide feedback about the jModelTest requirements for D3.2 and further deliverables. Now, the aforementioned distributed versions are being integrated in the SG paradigm in both their sequential and distributed releases. For the latter, the previously described actor between DRMAA and SAGA will be an asset.

By means of the CHAIN WP2 survey, new contacts of researchers interested in the scientific applications of jModelTest and ProtTest3 were provided to the VRC. The countries involved were Burundi, Costa Rica, Democratic Republic of Congo, Ethiopia, Nigeria, Panama, Sudan and Taiwan. As in the previous VRCs cases, new researchers interested in this initiative are being polled by the latest survey.

4.4. INDICATE / DC-NET

These projects (http://www.indicate-project.eu and http://www.dc-net.org) are working on coordinating policy and best practices regarding the use of e-Infrastructures for Digital Cultural Heritage. The projects aim at establishing and stimulating a network of common interest made up of experts and researchers in all the relevant fields, whose sustainability will be planned on a long term beyond the projects' lifetimes.

INDICATE focuses its activities in countries all around the Mediterranean, both European and African. Nevertheless, contacts have been established with China and Latin America. Since INDICATE finishes after DC-NET and there is substantial overlap between the 2 projects, CHAIN has signed an MoU with INIDCATE on March 28th, 2012.

Humanities have a very different approach to ICT than Science. In addition, their interests and required services are diverse, which improves the quality of the road-map and validation tests to be done by CHAIN. In this way, new contacts to integrate a digital repository in Panama by means of the INDICATE e-Culture SG have been provided and CHAIN is also searching for new scientists through its new delivered survey. Since digital repositories are becoming more and more important, the collaboration between INDICATE and CHAIN is expected to gain from this fostered activity. As in the previous cases, the first version of the road-map of services has been sent to the INDICATE coordinator in order to get her feedback.

4.5. DECIDE

DECIDE (http://www.eu-decide.eu/) objective is to design, implement, and validate a Grid-based e-Infrastructure building upon neuGRID[31] and relying on the Pan-European backbone GEANT[32] and the NRENs. Over this e-Infrastructure, a service will be provided for the

---

[29] jModelTest distributed version: M. Loureiro *et al*., Studies Health Tech. Informatics **159**, 244-248 (2010)
[30] ProtTest3 distributed version: M. Rodríguez-Pascual *et al*. Proc. HeathGrid 2011 Conf, H5 (2011)
[31] NeuGRID project, http://www.neugrid.eu/
[32] GEANT network, http://www.geant2.net/

computer-aided extraction of diagnostic markers for Alzheimer's disease and schizophrenia from medical images.

An MoU with this VRC was signed on April 13th, 2012, but a close collaboration has been established anyway in the last months by common partners shared between the two initiatives, the direct contacts between the managerial bodies (road-map of services has been sent to the VRC coordinator requesting his feedback) and the participation of DECIDE in the events organised by CHAIN.

As a result, the SG paradigm promoted by CHAIN has been also adopted by DECIDE and the community will participate in the validation model that CHAIN is proposing.

4.6. Climate Change

A Climate Change-related VRC has been promoted by CHAIN since the very beginning (see the conference on the "Role of e-Infrastructures for Climate Change research"[33]). This topic is of outmost importance nowadays and has researchers across the world interested in it, so the work has been focused on two aspects: looking for well-established European projects that were aware of the ICT advances regarding Climate Change, and looking for researches in the Regions of interest to CHAIN who could be interested in joining either the developments offered by the European initiatives or from a new community gathered around a well-known application. In this sense, the dissemination and promotion of WRF4G among these scientists have been carried out by CHAIN.

As a result, CHAIN has attracted the interest from researchers settled in ten (10) non-European countries and has identified four (4) initiatives that could help as a starting point to definitively form a VRC. In most of these groups and with the support of ICTP, the most used application is RegCM[34]. In addition, CHAIN has also offered WRF4G code developers support to work on the applications used by these researches in order to port them to the Grid with the same methodology applied to CAM4G, the action of which could drive the creation of a VRC too. Logically, both strategies could find synergies to meet and, due to this, the new survey sent to the National representatives raises this point.

Last, groups devoted to the study of seismology have been identified in Latin America with the support of the GISELA project (they are interested in porting to the Grid the SPECFEM3D[35] code) and India through of ICTP.

4.7. LSGC

In the last months, and due to its wide presence, the Life-Science Grid Community (LSGC, http://lsgc.org/en/LSGC:home;jsessionid=D1E7996F7665C3FA900B13A2116E1A36)   has been approached by CHAIN. LSGC is one of the biggest Grid users communities (HEP excluded) and counts on a solid background in this computational platform, where codes of different characteristics are being utilised nowadays. Besides, according to EGI accounting portal[36], during last year, there were jobs running in these regions: Asia Pacific, Canada, Europe, Latin America and Russia.

---

[33] Conf. Role of e-Infrastructures for Climate Change research, http://users.ictp.it/~smr2238/

[34] RegCM, http://users.ictp.it/~pubregcm/RegCM3/

[35] SPECFEM3D, http://www.geodynamics.org/cig/software/specfem3d

[36] EGI accounting portal, http://www3.egee.cesga.es/accounting/egi.php

LSGC representatives have been invited to make presentations about their communities from the first event organised by CHAIN, but it is recently, March 27th 2012, that an MoU has been signed between this initiative and CHAIN.

Since LSGC is mature enough and has very well established links with the people who participate in it, CHAIN is trying to find new users for it in the targeted Regions and, at the same time, to get its feedback for the proposed road-map of services profiting from its broad expertise.

4.8. SuperB

The SuperB flavour factory (http://superb.infn.it/home) is a major international research centre for fundamental and applied physics to be built on the campus of the University of Rome Tor Vergata. It is an accelerator that will provide complementary information to LHC, looking at rare decays with a very high luminosity electron-positron asymmetric collider.

From the CHAIN point of view, it is a novel approach to the research communities that could deserve very good results because it is based on the major Grid users group, i.e. HEP, but at the same time, is not limited to the rules and policies that the four ICT Divisions belonging to the LHC experiments impose or recommend due to its huge Grid use. In this way, groups interested in this kind of physics, but devoted to other kind of problems, can offer a new scenario and, besides, can collaborate with SuperB from their own place. SuperB VO is currently supported by five (5) National Grid Initiatives, none of them belonging to the Regions of interest to CHAIN.

Coming from outside Europe, groups belonging to twenty three (23) countries have been identified and the first discussions with the SuperB representatives have been carried out in order to sign a new MoU which better defines a strategy for collaboration. In the meantime, CHAIN is also looking for new contacts by means of its recently delivered survey.

5. Report on the initial work on the validation of the reference communities and "Quality metrics" assessment

The validation of the identified VRCs will be mainly done within CHAIN by means of the SG paradigm. The main reason for that is the development based on standards of the platform, which is expected, in addition, to enhance the use of the Grid technology worldwide because of the access to the infrastructure by using identity federations instead of certification authorities. In the previous Section, a report on how the VRCs are adapting the SG scenario has been described.

In this way, the adaptation of several applications of interest for the VRCs to this paradigm is offering to the developers, both from the SG and code developers' side, an optimum testbed where to validate the different necessities and characteristics required. Fortunately, the communities collaborating with CHAIN are involved in the adoption of the SG for their own work and the synergies are being increased.

The first results, once that several applications are currently running through the SG, are expected by the end of 2012, where a bigger use of these gateways will help in finding bugs or improving the available services.

As part of the validation model, the Quality Metrics[37] assessment is also of help. With regards to WP3, this is the status as that of March 2012:

Project Key Indicator	Metric	Target (Nov 2012)	Current status (Mar 2012)
PKI WP3.1	Number of VRCs officially supported (through MoUs)	2	5 (WeNMR, WRF4G, jModelTest, LSGC, INDICATE)
PKI WP3.2	Number of applications deployed on the participating regional e-Infrastructures	2 (1 per scientific discipline)	6 candidates (GROMACS, HADDOCK, WRF4G, CAM4G, jModelTest, ProtTest3)
PKI WP3.3	Number of new applications with respect to those already deployed in the regional projects	1	3 candidates (CAM4G, jModelTest, ProtTest3)

Table I. Quality metrics related to WP3 in the CHAIN DoW

It is important to mention that the applications appearing in Table I do it as candidates since on they are currently still in the process of being integrated in the SG paradigm.

The current planning of the CHAIN project is to have an interoperability test demo at the EGI Technical Forum 2012 in Prague in September 2012.

---

[37] CHAIN Quality Metrics, http://www.chain-project.eu/quality-metrics

6. Conclusions

This deliverable shows updated information about the VRCs collaborating with the CHAIN initiative, i.e. those communities that are benefiting from CHAIN activities and, at the same time, are participating in the validation model proposed by the project. These are:

- We-NMR;
- WRF4G;
- jModelTest / ProtTest;
- INDICATE (DC-NET);
- DECIDE;
- Climate Change;
- LSGC; and,
- SuperB.

All of them are trans-continental scientific and technical communities and the group as a whole represents a diverse sample, regarding the domains and degree of maturity represented.

Additionally, the way these collaborations have become successful has also been described. Thus, this document reports on the steps carried out up to now, which are expected to be improved once the results from the new questionnaire submitted by WP3 is collected and analysed.