Co-ordination & Harmonisation of Advanced e-Infrastructures
Research Infrastructures – Support Action
Grant Agreement n. 260011
www.chain-project.eu

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

Deliverable D4.4
Roadmap of Follow-up Activities

Version 12

Due on: 30/11/2012 Submitted on: 30/11/2012

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

Abstract:

This deliverable provides a roadmap of the follow-up activities as identified during the CHAIN project. Starting with overview of the current status of European Grid Initiative (EGI), it contains a set of recommendations divided into four main Sections: (i) on organizational and non-technical matters, (ii) on technical matters, (iii) related to virtual research communities and, finally, (iv) region-specific ones. The recommendations together form a clear and concrete roadmap of common and specific activities that stays ahead for grid and research communities in the medium and long term future.

Document log

Issue	Date	Comment	Author
V1	25/10/2012	First draft of regional Sections, circulated for internal comments	Aleš Křenek, Petr Hanousek
V2	29/10/2012	Comments processed, ready to circulate to regional partners.	Aleš Křenek
V4	16/11/2012	Preliminary material to the common recommendations Sections. Merged in regional contributions received so far.	Aleš Křenek, Petr Hanousek, Eric Yen, Salem Alagtash, Luis Nunez
V5	21/11/2012	Comments by O.Prnjat, merge of Chinese contribution. VRC Section. Revisited regional Sections except of Africa and India. Reviewed common recommendations.	Ognjen Prnjat, Sijin Qian, Aleš Křenek, Petr Hanousek
V6	29/11/2012	Regional contribution from Africa and India amended. EGI Section. Consolidation of all regional Sections.	Margeret Ngwira, Neeraj Sinha, Luděk Matyska, Aleš Křenek
V7	30/11/2012	Reflected O. Prnjat's comments to VRC Section. Consolidated common recommendations and remaining regional Sections.	Petr Hanousek, Aleš Křenek
V8	30/11/2012	Finalizing Introduction, Executive Summary and Conclusion	Luděk Matyska
V9	30/11/2012	Cleanup, pre-final release	Aleš Křenek
V11	30/11/2012	Additional Cleanup	Sijin Qian, Ludek Matyska, Petr Hanousek
V12	30/11/2012	Final revision	Federico Ruggieri

Table of contents

1. Introduction

The CHAIN project, started on the 1st of December 2010, aims to coordinate and leverage 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. CHAIN uses their results with a vision of a harmonised and optimised interaction model for e-Infrastructure and specifically Grid interfaces between Europe and the rest of the world. During its 24 months, CHAIN project collected serious pile of information related to commonalities as well as differences between individual regions in their way toward national and regional distributed computing (grid) infrastructures. CHAIN project also extensively studied the EGI organizational and operational model (as EGI is the most advanced and most extensive grid infrastructure worldwide) and applied the experience in proposing an organizational model for interoperable world-wide grid infrastructure in its D4.3 deliverable. CHAIN project also extensively worked with international research communities to analyse and propose models for the proper interaction between grid infrastructure and virtual research communities spanning whole globe or at least several regions. These communities also served as a driver for further adaptations of preliminary organizational models, guaranteeing a wide applicability of collected results.

This deliverable represents a culmination of one thread of CHAIN efforts, namely to elaborate a strategy and define the instruments necessary to ensure coordination and interoperation of the European Grid Infrastructure with those emerging in other regions of the world (Asia, including India and China, Mediterranean, Latin America and Sub-Saharan Africa).

1.1. Purpose

The primary goal of this deliverable is to provide a summary of achievements and results of the CHAIN project in a form of a Roadmap, a set of recommendations for the follow up activities to build, manage and extend a world-wide interoperable grid infrastructure. The Roadmap is expected to serve as a guideline for all actors involved in grid-related activities, covering organizational (non-technical) and operational (technical) matters as well as dealing with specific research communities. To setup the reference frame, this deliverable also includes a short description of the current EGI status, its organizational and operational principles.

This document also demonstrates that there is a huge set of activities that are common to all the regions and countries, providing thus a strong support for the optimistic assumption (that lay in the basis of the CHAIN project itself) that there is high chance to build an interoperable world-wide grid infrastructure sharing many common principles.

On the other hand, this document also clearly demonstrates that difference that exists between individual regions and their countries can be clearly identified, described and taken into account in a form of specific recommendations that aim to minimize the unnecessary differences while strengthening aspects that are unique and important for the world-wide collaboration.

Last but not least, this deliverable also serves as an input and light guidance for the CHAIN-REDS project activities (CHAIN-REDS project will start immediately after end of the CHAIN project), explicitly naming CHAIN-REDS workpackages and tasks that are expected to take care of some of the follow up activities specified in this roadmap.

1.2. Terminology

This subsection provides the definitions of terms, acronyms, and abbreviations required to properly interpret this document.

1.3. Glossary

CHAIN	Co-ordination and Harmonisation of Advanced e-Infrastructures
CHAIN-REDS	CHAIN for **R**esearch & **E**ducation **D**ata **S**haring
DoW	Description of Work – Annex I to the GA
EC	European Commission
EGI	European Grid Initiative
EGI-InSPIRE	European Grid Initiative – Integrated Sustained Pan-European InfrastructuRE
EMBL	European Molecular Biology Laboratory
EPIKH	Exchange Programme to advanced e-Infrastructure Know-How
FP7	European Commission's Framework Programme Seven
GA	Grant Agreement
GGUS	Global Grid User Support
GPGPU	General Purpose Graphic Processing Unit
HPC	High Performance Computing
MoU	Memorandum of Understanding
NREN	National Research and Education Network
OLA	Operation Level Agreements
OSG	Open Science Grid
POP	Point of Presence
ROC	Regional Operation Center
SLA	Service Level Agreements
xGUS	External Grid User Support

2. Executive Summary

This deliverable, the “Roadmap of the follow up activities” can be considered as one of the most important outcomes of the CHAIN project, to serve as a basis for the next period of work on creating, managing and further extending the world-wide grid infrastructure. Building on the collection of work performed within the CHAIN project, the Roadmap, in its current form, provides concise and clear guidelines to actors willing to participate in a setup and operation such infrastructure.

A short EGI status overview helps to provide a proper perspective for this Roadmap, ensuring an anchor for the recommended follow up activities and steps to relate to work and plans of EGI, currently the largest and most extensive grid infrastructure worldwide.

The actual Roadmap deals separately with Organizational (and in general, non-technical) as well as Operational (i.e., technical) matters and issues.

As part of the organizational recommendations, the Roadmap discusses implications stemming from the proposed world-wide collaboration model, developed by CHAIN and proposed in its Deliverable D4.3. It also discusses necessary activities for relationship formalization, coordinated user support and dissemination; all these activities are considered general in nature that they can be discussed without explicit relationship to any particular country or region.

The technical follow up activities cover networking, user authentication and authorization through PKI (including setup and operation of Certification Authorities) and identity federations, interoperability (through standards) and Science Gateways as means to overcome the current and eventually also future heterogeneity at the application level. Hardware is currently limited not only to the physical resources, but virtualized ones, provided as part of the cloud activities, are also discussed and recommendations formulated to help in establishing some common strategy to this novel technology and its potential.

As the infrastructure does not have any meaning without users, the Roadmap works with the non-substitutable role of large international research communities to oppose any diverging activities that could emerge in individual regions and that could endanger the setup and operation of world-wide interoperable grid infrastructure. Through specific support of selected large Virtual Research Communities that need interoperable infrastructure, pushing for standards adoption, the converging path towards full interoperability is sketched.

Last part of the Roadmap deals with specificities of individual regions, taking into account their different maturity and cultural and technological background. Each individual region that has been targeted by the CHAIN project is individually discussed and set of recommendations for their specific follow up activities presented. These specific recommendations again follow the pattern of organizational, operational and VRC issues, but deal with unique situation in each individual region or large country (in case of India and China).

With separate discussion of common and region specific issues the Roadmap covers both general converging matters where difference could damage the ultimate goal of a setup of a world-wide grid infrastructure and region and country specific steps that are factual enough to be immediately taken by actors in individual regions and countries and used as clear guidance for their future work.

While the Roadmap is one of the final CHAIN deliverables, a follow up project – CHAIN-REDS – has been already approved to start immediately after end of the CHAIN project. The Roadmap therefore clearly identifies issues and recommendations relevant directly for the CHAIN-REDS projects, giving an additional input (in parallel to the CHAIN-REDS description of work) for the new project activities. The Roadmap CHAIN-REDS related recommendations focus specifically on areas where new activities are needed and already foreseen, like more user friendly authentication through identity federations, further development of Science Gateway to leverage the interest of large international VRCs to use interoperating grid infrastructures, to cloud infrastructures and their future role, and last, but not least, on dealing with big data, especially those collected through world-wide international collaboration and expected to serve human wellbeing (e.g. sensor data collected to better predict climate change and related change in pattern of natural disasters).

3. EGI Status Overview

EGI is a collaboration and community of independent resource providers. Coordinated by EGI.eu, an organization founded in 2010, its primary aim is to build, operate and evolve the European grid infrastructure as a large scale distributed computing and storage facility. EGI works with specific research communities to achieve this goal.

Resource providers and eventually also the research communities are organized on a national basis through NGIs—National Grid Initiatives. These create the formal building blocks of EGI—NGI representatives are members of the EGI Council, the supreme governing body of the EGI Collaboration. The membership is restricted, only NGIs from EU member or affiliated states or from states with specific relationship with EU can become members. Apart from countries, also EIROForum organizations (like CERN or EMBL) can become EGI Council members. Each member has one representative in the EGI Council, however voting is weighted and the number of votes is proportional to the GDP of a particular country (using the so called TERENA key). While the EIRO members can select their membership fees and thus the number of votes (it must lie within the limits set by the smallest and largest country), they cannot achieve a majority of votes.

The coordinating body—EGI.eu—is located in Amsterdam (The Netherlands) and has a form of a Dutch Foundation (or Stichtung). It does not own any resources (i.e. it is not competing with any national or EIRO resource provider) and it provides services with added value to the community. The primary is the strategic planning and coordination of the implementation of agreed (by EGI Council) strategies. EGI.eu is also responsible for the communication, events and outreach at the international level, supplementing similar national activities realized by NGIs. The major service delivered by EGI.eu is operations and management of the international grid infrastructure, together with the necessary technology provision.

Recently, EGI.eu started a new set of activities, not directly related to the Operations. Through a Virtual team framework, EGI.eu provides a mechanism for international collaboration on any non-operational matter. Only international activities, covering at least three NGIs, are supported, ranging from outreach and specific work with research communities, to collection of information (e.g. creating the recently published EGI Compendium). This framework encourages collaboration among NGIs, mobilizing additional man power (coming from individual NGIs).

A strong aspect of the EGI Collaboration is its flexibility allowed for the NGI internal organization. There are no restrictions; individual countries could select organizational models that best fit their internal rules for infrastructures and their support. While some NGIs follow the EGI and EGI.eu model—namely that the national coordinating organization does not own any resources—in some countries NGIs are moderate or even strong resource provider, being able to offer reasonable resources to the research communities at national and international levels. The model not only provides wide freedom to the national organization, it is also able to deal with different maturity of individual NGIs.

EGI uses two basic platforms—the EGI Infrastructure Platform and the EGI Collaboration Platform—as primary architectural layers for its operations. EGI Infrastructure Platform provides core infrastructure services, necessary for running and managing large grid infrastructure. These services collect, keep, and provide status information about the grid and its services, the general information about resource owners (the grid topology) and also provide monitoring and accounting tools, together with the provisioning of a helpdesk. The EGI Collaboration Platform extends these services with new, domain neutral ones. This includes training, application database, workflows and workflow systems and also science gateways, usually a community-specific set of tools, applications and data integrated in a web portal. The science gateways can also span several infrastructures (e.g. connecting EGI grid infrastructure with a community-specific and operated one).

With the increasing interest in cloud infrastructure and EGI own work on cloud interfaces, the EGI Cloud Infrastructure Platform will offer capabilities necessary for future EGI cloud services.

The platform concept supports development of operations tools that could be offered in a reusable way to address typical resource and technology independent needs of different grid infrastructures. Through variety of configuration options EGI is able to provide solutions that can be partially or fully adopted outside the EGI scope, paving thus way to fully interoperable grid infrastructures.

Since 2010, EGI Collaboration and EGI.eu are co-funded by the EU project InSPIRE, that covers part of the cost of EGI.eu and also of international activities of individual project partners (the NGIs joined in the EGI Council plus several other countries). For EGI.eu, project funding is complemented by an income based on the membership fees—like votes, these are derived from the GDP using the TERENA key. The membership fee (for the year 2013 it has been accepted by the EGI Council at the level of 1.5 million Euro) is the primary source for financial stability (and therefore sustainability) of the EGI.eu and the whole EGI Collaboration. As the EGI InSPIRE project will end in May 2014, discussion about the 2014 budget will be decisive for the shape and future of the EGI Collaboration under the Horizon 2020.

EGI InSPIRE project does not directly deal with the world-wide collaboration and relies on other EU co-funded project to cover the international aspects of the Grid collaboration (however, EGI InSPIRE is directly involved in collaboration with USA—interoperation with TerraGrid and recently XSEDE are taken care of by EGI InSPIRE operations themselves). Apart from the operations (where interoperability is already taken care of, at least partially, by the adopted platform-based architecture), the EGI Collaboration organizational structure is not suitable for extensive enlargement over the current size (34 EGI Council members and 5 associate members).

4. Common Recommendations on Organizational and Other Non-Technical Issues

In this Section common recommendations and steps factorized out of the regional Sections are presented and discussed in more detail. The issues are referenced from the regional Sections then, with specifications of concrete actions and responsibilities.

4.1. Peer-to-peer world-wide collaboration model

Due to the known limits on increasing the number of EGI Collaboration, for the world-wide collaboration, a peer-to-peer model based on individual regions is proposed (in the CHAIN Deliverable D4.3). In this model, each region is expected to take care of its own organization, the technology and middleware and the interoperability is dealt with at the peer-to-peer technology level. EGI offers a set of tools, services and middleware together with an organizational model (the ROC-based one) whose adoption by other regions practically solves the world-wide interoperability problem. However, even if different technology is used in individual regions, EGI provides a list of services that must be taken care of to create truly interoperable grid infrastructures. In these cases, scientific gateways are the recommended means to access different grid infrastructures through a common interface. They also offer bridges between the core interoperability services, allowing pass status (and other meta-) information between grid infrastructures, creating a virtual interoperable grid infrastructure even on top of technologically very different ones.

4.2. Role of CHAIN-REDS

All work-packages of the starting CHAIN-REDS project are highly relevant for virtually all the recommendations throughout this document. However, not all the work required to achieve the final goal of global interoperable e-Infrastructure can be done within the scope of a single project with limited budget. Therefore wherever we further refer to CHAIN-REDS work-packages, we mostly assume coordinating, supporting and dissemination role in case of WP2, and provision of technical expertise, consulting, and service prototypes in case of WP3—WP5.

4.3. Dissemination

Proven in the series of grid-centric collaborative project between EU and various regions of the world, including the harmonization effort of the CHAIN project, communication, dissemination, and knowledge transfer are essential concepts in achieving the goals of building worldwide interoperable e-Infrastructures.

Typically, this task is accomplished by organizing the conferences, workshops and training sessions. During these meetings collaboration among VRCs, between VRCs and DCIs are discussed, and even the support from the local governments and possibly from commercial subjects can be involved. It is recommended to hold one regional conference per year and several workshops and trainings with all involved VRCs and DCIs. The responsibility for organisation of such actions is upon the regional leader (ROC, JRU). Training organisation is not taking the significant part of CHAIN-REDS, specialized projects are expected to emerge, following mainly the EPIKH project and other initiatives, such as BrainGain from UNESCO, that CHAIN is already collaborating with and other initiatives, such as BrainGain from UNESCO, that CHAIN is already collaborating with.

Holding the regular meetings should be supplemented by proper and actual documentation on provided applications and services.

To attract and inform the new users should also serve the proper information about available software in a SW catalogue, preferably EGI APPDB[1]. The service itself is maintained by EGI, however, inputs and maintenance of the data there are the responsibility of the application communities. The role of dissemination is explaining the value of the service and encouraging the communities to use it.
These dissemination activities should necessarily involve the participation of ROC/NGI staff on the EGI and maybe OSG meetings. Such involvement will lead to staying aware of the tools used for building and maintaining e-Infrastructures and to adopt and discuss the best practises raised during the past development of these bodies.
CHAIN-REDS WP2 is focused on dissemination, and it is expected to encourage and to help coordinating activities of this kind.
Collecting of best practises is a part of CHAIN-REDS WP4.

4.4. Formalisation of relationships

The scientific community usually tend to build projects on the enthusiastic base, solving problems in best effort way. To be more dependable for users, the routine operation should be secured in better than "best effort" way. Generally, for all regions, is thus recommended the establishing of some Service Level Agreements (SLAs) between users (NGIs) and DCIs and Operation Level Agreements (OLAs) inside the DCIs and NGIs. These SLAs should also cover the theme of intellectual properties raised from using the DCIs hardware and the proper using of the software. Gradual development of more formalized relationships along these lines should be among activities of all the emerging NGIs. CHAIN-REDS WP2 can help with provision of appropriate contacts to EGI and European NGIs from where examples and templates can be drawn.

4.5. Coordinated user support

Due to participation in the international Grid structures, it is sometimes necessary to escalate the problems to the proper administrators. Clear procedures of escalation the problem from national to international level have to be defined.
EGI uses the GGUS system for Europe-wide integration of user support. Various national ticketing systems are integrated with GGUS at the technical level. In general, for the other regions use of xGUS is recommended. It is not necessary to install additional instance of ticketing system, integration with EGI is available while preserving separate administration of the regional system; there is a common know-how about connecting other, more light-weight, trouble ticketing systems (like Request Tracker) to xGUS in EGI.

[1] http://appdb.egi.eu/

5. Common Technical Recommendations

Similarly to the previous Sections, common recommendations are presented here in more detail. The issues are referenced from the regional Sections then, with specifications of concrete actions and responsibilities.

5.1. Networking requirements

The general acceptance of the grid paradigm requires certain critical threshold of network bandwidth and reliability; failing to achieve those prevents the users to rely on remote resources, shifting their focus to purchase of local, frequently dedicated storage and computing equipment. These services must be provided by NRENs. The network should aim at least 2.5 Gbit/s network backbone and international uplinks in short term, with foreseen upgrade to 10 Gbit/s in medium term. Similarly, grid sites should be connected with at least 1 Gbit/s links. Reliability of the network connection should go above 99.9% (i.e. less than 1 hour outage in month).

Unfortunately, these requirements may still not be realistic in some of the target regions. Therefore we understand them as the desired, idealistic goal, and we define relaxed criteria in some of the regional Sections bellow.

5.2. Interoperability and standard adoption

Use of different middleware stacks, and even adoption of different computing paradigms besides the grid model – cloud and HPC in particular, yields man interoperation issues. A thorough but very long and effort consuming way is development and wide adoption of recognized standards.

On the other hand, some time ago a science gateways paradigm has been adopted as a way to bridge the gap among different systems. CHAIN Science Gateway[2] was developed in 2012 and supports all the middleware used in regions and EGI. Use of this paradigm is a pragmatic approach to address the interoperability issues.

Using the science gateways won't make the interoperability problem to disappear, it only moves it to another place. Development effort is relocated from standard adoption by various middleware stacks to application porting to the portal. However, based on long experience we believe the latter way to be more realistic in short to medium time frame. One of reasons is also more direct involvement of the user communities – VRCs, rather than the infrastructure only.

Therefore the general recommendation goes towards acceptance of the portal technology, and to joining the development efforts of both VRCs and the infrastructure providers around this paradigm. Science gateways are the concern of WP5 of CHAIN-REDS which will provide both coordination and some central development effort as well.

Standards are concern of CHAIN-REDS WP3, and development of appropriate strategies to approach the cloud and HPC resources is addressed there too.

5.3. Application integration with Science Gateway

Integrating an application to a portal (e.g. Science Gateway) requires non-negligible effort at the beginning; however, it enables the application to a wider scattered community of the users for whom the technological barrier of using the raw grid services could be too high otherwise. In short term, the portals, which are provided by collaborating European institutions, can be used by users in the target regions to access regional services. In medium term, an independent

[2] http://science-gateway.chain-project.eu/

regional portal(s) should be provided (probably by one of the more mature NGIs in the region). Specific applications, once identified (there must be indications of sufficiently large set of potential users) can be integrated through targeted short-term projects with participation of portal experts, experienced users of the application (or even its developers), and infrastructure experts (recruiting from NGIs typically). Finally, the ported applications should be widely disseminated to get the expected users base. CHAIN-REDS WP5 will work on evolution of the portal technology as well as it will contribute to porting specific applications (together with the user communities).

5.4. Certification authorities and Identity federations

The paradigm of authentication through certificates is widely accepted in the Grid, and the technology highly relies on it. However this approach has limits – non-trivial organizational issues on accrediting a CA and a steep learning curve for the end users in particular – which cannot be easily avoided with the concept of certificates.

The problem of accrediting CAs is gradually addressed by IGTF (International Grid Trust Federation, http://www.igtf.net/) which sets up CA associations (Policy Management Authorities – PMAs) defining common requirements on the CA policies. However, the associations still don't have global coverage (they are present in America, Europe, and Asia-Pacific region). CHAIN-REDS should support activities leading to extending the coverage to the target regions in cooperation with other projects (e.g. AfricaConnect) and organisations (e.g. TERENA).

On the other hand, the technology of identity federations could help overcoming the steep learning curve of certificate usage, technically hiding the internal certificate use behind mechanisms that are more users friendly. It is fair to say that it has its own limitations so a development is still needed[3]. Identity federations are the concern of CHAIN-REDS WP5, including the technical work targeted to overcome the known limitations (the need of M x N trust relationship between service and identity providers in particular).

5.5. Recommendations on hardware

5.5.1. Minimal hardware requirements

Specification of global minimal requirements on hardware resources is a delicate task, nearly impossible given the heterogeneity of the CHAIN target regions.

In developed regions it is feasible to recommend sites of at least 100 CPU cores and 50 TB of disk storage in order to keep the ratio of computing power vs. human effort at the optimal level. Similarly, the technology (CPU in particular) should not be older than 5—6 year in order to keep optimal performance to power consumption ratio.

However, such requirements may not be realistic in less developed regions. Therefore stating the minimal hardware requirements on a site to join the infrastructure must be left to regional decision – the only common statement is that there should be such limit specified. The specification should be also negotiated with the present user communities. Their needs may be varying, and they should be reflected in the hardware requirements. VRCs also may specify additional requirements on specific hardware, e.g. GPGPUs.

Similarly, certain levels of availability and reliability should be specified. The relaxed EGI criteria are 75% of month reliability and 70% of month availability[4], and those can be taken as the global minima.

[3] https://documents.egi.eu/public/ShowDocument?docid=1178
[4] https://documents.egi.eu/public/ShowDocument?docid=31

5.5.2. Provision of seed resources

In some countries in the target regions the resources available to the grid users are very limited. However, older European experience shows this to be a chicken-and-egg problem. Once the available resources reach certain (not extremely high) critical level, they start attracting users, and the users help bringing further resources to the shared infrastructure. NGIs in countries which need such resources, with the help of regional initiatives as well as CHAIN-REDS, should look for funding sources to set up several "seed" sites providing at least 500—1000 CPU cores and storage in the order of 100 TB. These resources should be interfaced to the grid, and widely advertised for general use of the scientific community. In this way the critical mass for the further infrastructure development will be gained. Specific actions are identified in the regional Sections bellow wherever applicable.

5.5.3. Keeping resources up to date

Typically, new hardware resources are purchased with funding of short-term projects. However, when such funding stops (the project is over), sustainability becomes an issue. Due to the Moors law, 5 years old computing equipment becomes essentially useless. The resource centres, with the help of NGIs and regional organizations should aim at getting sufficient sustainability-related funding, allowing not only operation of the purchased equipment but also necessary replacements.

5.6. Grid interfaces to existing resources

In many countries significant computing and storage resources exist but they are used exclusively by a single user community (or few ones only), with no interconnection to distributed infrastructure via standardized interfaces. Owners and current users of such resources should be approached, and interfaces to the grid infrastructure should be negotiated. There must be mutual benefits for this process to be successful; therefore advantages for both sides must be carefully and individually identified. To large extent, this is responsibility of CHAIN-REDS WP2.

From the technical viewpoint, rather strict requirements of the UMD-based software stack can be inconsistent with current requirements of the resource users. In this case, virtualization technologies should be leveraged, allowing co-existence of different software environments as well as dynamic assignment of hardware resources to the software environments. [CHAIN-REDS WP3]

The process should be initiated by NGIs, with the goal to attract majority of sites in medium term. Specific actions are identified in the regional Sections bellow wherever applicable.

6. Global VRCs

6.1. Specific support of selected VRCs

WP3 of the CHAIN project concentrated on work with global virtual research communities. One of this work results is a list of VRCs which signed (or are about to sign) a MoU with the project, and which deserve specific support because partners in these VRCs come from the regions of interest of CHAIN. In this Section we comment specifically on these VRCs, relating them specifically to roadmap goals described elsewhere in this document.

6.1.1. We-NMR

The We-NMR community provides gridified access to a set of approx. 15 widely used life-science applications through a custom web portal. The community is organized in a dedicated virtual organization recognized by EGI, and it provides all required VO-central services. Integration with Science Gateway is not foreseen in near future due to the lack of interest in the community, not willing to migrate from the existing portal solution, but an interest has been expressed towards some of the main SG components i.e. IdP Authentication and JSAGA with Robot Certificates.

The community has existing contacts in Latin America and Africa; specific actions to follow this collaboration are described in Sections 7.1.4 and 7.3.4 in "support existing collaborations". Due to the expressed interest in life science (see D4.1) in other regions, similar specific collaboration should be sought for as well.

6.1.2. Climate change and WRF4G

The climate and weather modelling community developed a specialized workflow engine and related software (WRF4G and DRM4G) to run the main applications relevant for the community on the grid. The current implementation is command-line only; however, integration to Science Gateway is desirable. This is an action for CHAIN-REDS WP5.

Moreover, in the target regions (India and Asia-Pacific in particular) there are strong communities in this area. Those should be contacted and further collaborations discussed [CHAIN-REDS WP2].

6.1.3. jModelTest and ProTest3

jModelTest and ProText3 are specific molecular modelling applications used in life sciences. The applications were successfully integrated in the Science Gateway. Due to general interest in the life science in most of target regions, specific groups should be looked for, contacted, and the use of the applications discussed [CHAIN-REDS WP2].

6.1.4. INDICATE and DC-NET

The INDICATE and DC-NET projects focused on preserving the cultural heritage in digital form. Within their framework, specific contacts to China were made and a Chinese-specific portal was provided. Preservation of this work is addressed in Section 7.5.4 "Support existing collaborations".

Digitization efforts in the cultural heritage fields will yield non-trivial requirements on long-term data storage, being the subject of CHAIN-REDS WP4.

6.1.5. DECIDE

The medical image processing, one of the topics of the DECIDE project, was integrated in a Science Gateway successfully. Interest in these applications is reported from Nigeria, addressed in Section 7.3.3 “Supporting user communities”.

6.1.6. LSGC

LSGC, the EGI community in life-science, is well established in all the target regions, therefore no specific actions, besides the general support by CHAIN-REDS is foreseen.

6.1.7. SuperB

SuperB is another VRC in the HEP framework and received particular attention in all the target regions. The original project is currently under revision by the Italian Ministry of Education, University and Research. Through naturally emerging channels in INFN specific contacts should be established and plans for collaboration discussed in CHAIN-REDS WP2.

6.2. VRC roadmap comments

Besides identifying specific VRCs, WP3 also distilled a list of quite generic requirements on the underlying infrastructure, from both technical and organizational point of view (CHAIN D3.4 and preceding documents). In this Section those requirements are related to the roadmap goals.

1. **Further consolidation of NGIs, for which the in-depth documentation has been published through CHAIN website and circulated among the NGIs in CHAIN target regions** – global recommendation 4.3 and region specific recommendations in Software and services Sections (7.1.3, 7.2.3, 7.3.3, 7.4.3, 7.5.3, 7.6.3). Documentation support is in CHAIN-REDS WP2.
2. **To form VRCs that would group the related researchers of a common field and its liaison with similar communities in Europe and the other regions avoiding duplicated efforts** – dissemination support is in CHAIN-REDS WP2 and here in Section 6.1.
3. **To use the VRCs as a key point for the use of the infrastructure and the best test for demonstrating its sustainability** – recommendation is in Section 5.2, also some sort of support is in CHAIN-REDS WP4.
4. **For regions where this is relevant, supporting and/or establishing a dedicated regional body for Grid coordination** – this is relevant for all regions except India, mainly for Asia-Pacific and Africa. See relevant Sections (7.1.5, 7.2.5, 7.3.5, 7.4.5, 7.5.5). CHAIN-REDS WP2 is also particularly useful.
5. **To prepare a cook book of recipes on how to set-up and manage fully fledged ROCs, possibly complemented by Wiki pages or community tools** – cookbook of recipes is under preparation within CHAIN and will be released by CHAIN-REDS WP3.
6. **To ensure a service quality by means of SLA and SOP** – SLA and OLAs are part of CHAIN-REDS WP3 and here of Sections 4.4, 7.5.5 (point 4) and 7.6.5 (point 3).
7. **Interoperation of helpdesks** – Section 5.1 of this document. Partly CHAIN-REDS WP3 is suitable for this.
8. **Support for interoperable monitoring and accounting mechanisms, core services, job submission and data management.** – Section 5.2 of this document and CHAIN-REDS WP3.
9. **To coordinate the security related efforts and to participate in the corresponding well established bodies** – security is concern of global Section 5.4 and region specific recommendations in Software and services Sections (7.1.3, 7.2.3, 7.3.3, 7.4.3, 7.5.3, 7.6.3). Also is related to SLAs (point 6 of Section 6.2).
10. **Adoption of standards** – Section 5.2 targets this recommendation

11. **Interoperation and interoperability between different e-Infrastructures** – Sections **Error! Reference source not found.**, 5.2 and 5.4. Also some region specific recommendations: Software and services Sections (7.1.3, 7.2.3, 7.3.3, 7.4.3, 7.5.3, 7.6.3),
12. **Queues for the submission and scheduling of jobs should …** – this is raised from the needs of VRCs. Therefore is valid the comment for Section 6.2 point 3.
13. **Computing Elements, Working Nodes and sites should provide**… - Sections 0 and 6.1 in this document. CHAIN-REDS supports only storage needs in WP4.
14. **Storage should provide…** - Sections 0 and 6.1 in this document. CHAIN-REDS WP4.
15. **Data...** – partially Sections 0 and 6.1, some bullets of D3.4 are not concerned.
16. **Plans for scalability issues** – this should be dependent upon VRC needs. Section 6.1 targets needs of VRCs.
17. **Accurate information about the sites should be published** – this is the scope of Section 0 point **Error! Reference source not found.**.
18. **OLA and SLA to be agreed among interested actors** – Section 6.2 point 6.
19. **Simple mechanisms to deal with troubleshooting and monitoring (GGUS)** – Section 6.2 point 7.
20. **Provision of intellectual property rights (when some data/information demand it) and secure mechanisms for the data transmission** – this is the part of SLAs and thus the Section 4.4 applies here. Also the CHAIN-REDS WP3.
21. **There are still communities and/or groups of researchers who are not sufficiently ICT skilled** – this is the scope of CHAIN-REDS WP2 and Section 4.3 of this document.
22. **Sustainability of the VO management is a key point due to its cost** – sustainability is the concern of region specific Funding Sections (7.1.6, 7.2.6, 7.3.6, 7.4.6, 7.5.6, 7.6.6).
23. **Implementation of user friendly front-ends for the submission of jobs (Science Gateway, gUse, etc.)** – concern of Section 5.2 and CHAIN-REDS WP3. Also the Software and services Sections (7.2.3, 7.3.3, 7.4.3, 7.5.3, 7.6.3).
24. **Implementation of library packages and services to access DCI for application developers** – it is not in the scope of this roadmap. However, some activities in this direction may be initiated within CHAIN-REDS WP3.
25. **More accessible authentication and authorisation methods** – partly Section 5.4 and CHAIN-REDS WP5.
26. **Drawing up of best practice Handbooks, Technical References and Wikis** – Section 4.3, CHAIN-REDS WP2 and partly WP4.
27. **Infrastructures more user community driven** – whole this document, mainly Sections 0 and 6.1.
28. **Agile tools for VO administration, easily allowing creation of new groups and roles** – not directly in the scope of this roadmap.
29. **Collaborative scenario** – concern of Section 4.3 and CHAIN-REDS WP2.

7. Region Specific Recommendations

Recommendations in this Section are classified as short term with the meaning that the goals are expected to be achieved by the end of 2013, medium term – in 2015, and long term – by 2020. In addition, some of the recommendations hold generally, independently on the time frame (those are not classified explicitly). Tentative responsibilities to institutions and/or workpackages of CHAIN-REDS project are assigned as well. Wherever the recommendations emerge from related ones in CHAIN D2.2, those are indicated in parentheses, e.g. (59-LA). Other relevant CHAIN documents are referenced as well.

7.1. Latin America

7.1.1. Networking

1. **Network speed and reliability.** With possible exceptions (Cuba in particular) the NREN backbone is operated by RedCLARA (Latinamerica Cooperation of Advance Networking) with bandwidth and connection to Europe above the critical level for grid computing. However, due to the growing demand the network backbone bandwidth should be upgraded to 2.5 Gbit/s in short term, aiming at 10 Gbit/s in medium term. Make sure it is achieved in the whole region (with possible relaxations where not realistic but not bellow 155 Mbit/s. Starting in 2013, all the REdCLARA partners will assume all the costs of the networking operations. Some its NRENs (México, Costa Rica, Colombia, Ecuador and Argentina) have agreed to sustain and to develop advanced computing services of LatinAmerica. [short/medium, RedCLARA]

7.1.2. Hardware Resources

1. **Ensure seed resources.** (*59-LA*) In countries where the grid resources are below the critical level (Panama, Paraguay, Costa Rica, Bolivia, Peru, Cuba, and Guatemala) the grid initiatives should look for attracting at least one site in short term, and to help it to provide access to the resources through grid interfaces in order to attract the local users. Alternatively, one generic national grid site should be set up with the help of centralized funding to serve the same purpose. Particularly, the region has devised a strategy based on Advanced Computing Services (ACS) where Grid is only one of the possible services, adapted to LA needs. [short, RedCLARA]
2. Avoid building parallel infrastructures. The scientific communities may tend to build parallel "grid" infrastructures specific for the needs of individual scientific disciplines. The permanent interaction among RedCLARA and its associated NRENs is preventing to replicate the same efforts in each country. The e-infrastructure (hardware, software and manpower) is shared across disciplines by researchers form different countries. [long, RedCLARA and CHAIN-REDS WP2]

7.1.3. Software and Services

1. **Consolidate certification authorities.** Currently most of the users in the region use certificates issued by UFF Latin American and Caribbean Catch-all Grid CA. This works for the time being but it is not a scalable model for the future. Even for the medium term, infrastructure of national certification authorities, approved by TAGPMA seems not to be viable in the region. Therefore the catch-all CA should be capable of handling the increased load. [medium, RedCLARA]
   RedCLARA and its associated NRENs are moving towards the Identity Federation scheme, in order to provide ACS, through Science Gateway interface in a more user-friendly way. [short, RedCLARA and CHAIN-REDS WP5]

2. **Deploy region-centralized services.** The region is quite large, full integration of individual countries and sites with EGI, as a straightforward extension of the European setup, would reach scaling limits. Instead, fully operational set of EGI-like services, but completely independent on the European infrastructure should be set up in the region. This can be achieved by following and extending the outcomes of the GISELA project. The exact roadmap depends on the definition of organization structure, cf. Section 7.1.5. [short/medium, RedCLARA]
3. **Primary middleware stack.** As a natural outcome of the collaboration with Europe many countries in the region, including all those with the driving potential, deploy UMD services. Therefore UMD should be encouraged as the primary middleware stack. Obviously, where there is the need for other stacks (Globus based services for smooth interoperation with OSG and local development-implementation like OurGrid from Brazil), parallel deployment should not be prevented. When necessary, virtualization technologies should be leveraged to allow co-existence of multiple middleware stacks on the same physical resources. [medium, RedCLARA and CHAIN-REDS WP3]
4. **Support interoperability with US.** There are user communities in the region who collaborate with US partners traditionally. These communities should be contacted, and work on interoperability between the LA grid infrastructure and OSG started. The proven interoperability pattern between EGI and OSG, as well as experience from this collaboration, can be followed. Coordination of the activity is in the scope of CHAIN-REDS, implementation of RedCLARA. [short, RedCLARA and CHAIN-REDS WP2]

7.1.4. User Communities

1. **Support small and emerging user communities**. In the region, there are either countries with large number of users in established communities (e.g. Brazil, México Colombia, Chile, Argentina and Venezuela), or countries with virtually no national users. In the latter ones, smaller user communities should be actively sought and provided with support to establish themselves on the grid. These communities are likely to represent non-negligible mass of users when aggregated. However, due to their diverse requirements they need to be approached individually. A regional plan to approach these communities should be developed in short term and executed in medium term, resulting in visible increase of the user base. Emerging from the GISELA project, five countries (México, Costa Rica, Colombia, Ecuador and Argentina) have agreed to support the regional ACS, and promote social impact projects which could profits from it. [short/medium, CHAIN-REDS WP2/RedCLARA]
2. **Maintain discipline-neutral infrastructure.** Existing user communities in the region cover many scientific areas, there seems to be no particular focus. Therefore the infrastructure must be kept strictly neutral w.r.t. the scientific discipline. When contradicting requirements (e.g. deployment of incompatible software, different versions of operating systems etc.) appear, they should be resolved using e.g. virtualization techniques, not favoring any of the user communities. [short, CHAIN-REDS WP3/4]
3. **Support existing collaborations.** (*62-LA*) There are large international scientific projects active in the region (e.g. Pierre Auger Cosmic Ray Observatory and ESO Observatories). Links with regional representatives of these collaborations should be established and their needs supported directly. Mechanisms similar to the EGI virtual teams can be leveraged to establish a working collaboration of the infrastructure experts who are close to the specific scientific community in their country. [short, CHAIN-REDS WP2 and the involved NRENs]

   There is participation in We-NMR global VRC in Costa Rica. The team should be contacted together with eligible resource providers, and a plan for resource commitment to the We-NMR VO should be discussed. At the technical level the integration will go

smoothly then due to the UMD-compatible middleware and use of PMA-approved catch-all CA. [short, RedCLARA and Costa Rica NREN].

7.1.5. Organization

1. **Define suitable organization model for the region.** A straightforward mirroring of the EGI-NGI pattern may not work in the region. Despite of fairly high level of the grid development, many countries are rather reluctant to establish NGI as a formal body. In LA even the NRENs typically depend strongly on funding and sustainability of their members. It seems to be more appropriate to look for closer regional collaboration in the region, with driving strength of countries which already have the critical mass of resources. Investigate the reasons further, and propose a suitable alternate organization model in short term to be deployed in medium term. This model is slowly emerging after the GISELA project with the active participation of the GISELA-CLARA Transition team and five interested NRENs. This organization will implement a business model in order to share costs across countries and disciplines. The model should also define the role of RedCLARA in the organization of the grid activities in LA. [short, RedCLARA]
2. **Integrate regional grid operations.** (*61-LA*) The current regional grid operations were run with the support of GISELA project, and the plans assumed they will be fully integrated into the infrastructure of RedCLARA. However, the transfer is not taking place. A clear alternate plan for the ROC-LA follow-up should be defined, and the support of task forces coming from different countries should be leveraged. [short, RedCLARA]
3. **Transfer expertise among sub-regions.** (*63-LA*) Level of maturity of acceptance of the grid paradigm, deployment of the services, and collaboration between the infrastructure providers and the user communities is not homogeneous in the region. On the other hand, international collaboration among the NRENs is well established. Strategies to transfer the know-how from the more mature sub-regions to the less mature ones, reusing the existing communication channels, should be defined and followed subsequently. [medium, RedCLARA]

7.1.6. Funding

1. **Ensure sustainable funding.** (*60-LA*) Following the defined organizational model for the region, develop the right lobbying strategies to ensure sustainable funding of such international collaboration. The strategy is based first on the commitment of the above mentioned NRENs to support the operation of the regional ACS; secondly selecting a project with a high social impact in order to gain some external funding and to illustrate the usefulness of the e-infrastructure. Finally to implement the business plan to make the ACS sustainable. [medium, RedCLARA and NRENs]

7.2. Mediterranean, Middle East and Gulf

7.2.1. Networking

1. **Network speed and reliability.** (*72-MED*) The NREN backbone bandwidth and connection to the global research and education network, mainly in Europe, is at different level of implementation, but in general above the critical level for grid computing. Algeria ARN and Palestine PalNREN are currently connected through the EUMEDCONNECT3 through dedicated links operating at speeds: 622 Mbit/s and 45 Mbit/s, respectively. EUMEDCONNECT3 has dedicated points of presence (PoPs) at Catania (Sicily) and Nicosia available for direct interconnection by EUMEDCONNECT3 partners and also options for connecting via hubs in Europe on the GEANT network. In addition to

networking the EUMEDCONNECT3 partner countries together, the links to GÉANT also provide connections to research networks in other world regions:

- Europe - Research and education networks in over 30 European countries
- North America (INTERNET2, National Lambdarail, CANARIE, ESnet)
- Asia Pacific via TEIN3
- The South African research network, TENET
- Latin America via RedCLARA
- Sub Saharan African via UbuntuNet and Africa Connect
- Commercial Internet (optional service, terms and conditions apply).

The other Mediterranean countries: Morocco, Egypt, Tunisia, Syria, and Jordan have been connected to EUMEDCONNECT2, but are now disconnected because of today's dominating political instability and financial crises. Yet grid applications are being the driving force in this region. The United Arab Emirates, Qatar, and Saudi Arabia are connected to the US Internet2 with gigabits links. The remaining Arab countries are in the process of developing NRENs. Further improvements towards 2.5 Gbit/s of the NREN backbone in short term, upgrading to 10 Gbit/s in medium term are desirable. Similarly the connectivity of the grid sites should be upgraded to 1 Gbit/s to avoid bottlenecks at the last mile.[short/medium, ASREN]

2. **Support developing NRENs.** (*72-MED*) In some countries in the region NRENs are at early stages of implementation. ASREN has been supporting development of NREN in Lebanon, Libya, and Bahrain, and in some cases helping to look for suitable funding sources to acquire sufficient but still realistic connection quality. ASREN aims at developing a Pan-Arab operational network covering the whole region and all NRENs established in medium term. The network development is being coordinated with the grid infrastructure so that grid sites and the emerging user communities are connected with sufficient bandwidth and reliability. [short—long, ASREN]

7.2.2. Hardware Resources

1. **Grid interfaces to existing resources.** Identify not yet connected resources and negotiate their connection to the grid infrastructure (Section 5.6) [short/medium, ASREN and CHAIN-REDS WP2]
2. **Provide seed resources.** Identify countries where seed resources (Section 5.5.2) should be provided, and develop appropriate plans, also according to the local political situation. [short—long, ASREN and the affected NGIs/NRENs]

7.2.3. Software and Services

1. **Fully accredited CA infrastructure.** (*65-MED, 66-MED*) The process of establishing accredited certification authorities already advanced in the region, with CAs accredited in Morocco, Algeria, Egypt, Jordan and Syria. In short term, CAs will be possible in all countries with NGIs established. The infrastructure for identity federations is also at different level of implementations, but still did not reach maturity levels.[short, ASREN and NGIs]
2. **Deploy region-centralized services.** Central services necessary for independent operation of the grid infrastructure in the region (BDII, accounting portal, monitoring, GocDB, ticketing system, and more) following the EGI pattern have been deployed in Africa ROC and should be extended in the region in the short term. Africa ROC is the right body to coordinate this activity. Similarly, some optional but desirable services (VOMS for VOs emerging from the region, WMS for supported VOs, data catalogues, FTS) have been

installed in some countries but should gradually extended in medium term. Typically, this is done by institutions participating in scientific activities that require such services. However, NGIs should offer provisioning for some of them due to concentrated expertise. Outcomes of the EUMEDGRID-Support project have put emphasis on these topics and are being coordinated with the NGIs. [short/medium, ASREN and NGIs]

3. **Primary middleware stack.** With the collaboration with European partners, preference of UMD middleware (former gLite components in particular) has become visible in the region. UMD is a naturally emerging de-facto standard of the middleware stack, and therefore, facilitating the deployment of UMD-based services in all the countries. This de-facto standard should be further promoted in the region. [short—long, CHAIN-REDS and ASREN]
4. **Integrate applications in portals.** (*70-MED*) A EUMEDGRID Science Gateway has been deployed by the project, but it is currently based in Europe. Negotiate resources and required technical staff to run a regional instance of the Science Gateway portal. Identify suitable applications and contribute to their porting to the portal framework (Section 5.3) [medium/long, CHAIN-REDS WP2/5 and ASREN].

7.2.4. User Communities

1. **Support better organization of user communities.** (*71-MED, 72-MED*) In the region there are multiple user groups of the same scientific discipline, already using significant computational resources, however, they are rather isolated from one another. NGIs should map such existing user groups, interchange the information, and approach them with targeted support, helping them identify how they can benefit from collaboration and sharing resources over the grid. Once such regional per-discipline communities are established, NGIs can also help them establish worldwide contacts. [short/medium, CHAIN-REDS WP2 and NGIs]
2. **Support specific disciplines.** In the survey (D4.1) computer science and mathematics are reported to be significant user communities in the region. In short term, NGIs should approach this community specifically, and develop plans for their support leading to establishing a coordinated collaboration. [short, CHAIN-REDS WP2 and NGIs]

7.2.5. Organization

1. **Agree on coordinating body of the grid activities.** (*64-MED*) ASREN is recognized to be the coordinating body of the regional NRENs. It is emerging that it can also take the coordinating role of the grid activities in the region. The exact relationships should be defined, and formalized to some extent, e.g. with a MoU. [short, CHAIN-REDS WP2 and ASREN]
2. **Evaluate the model of interaction with EGI.** A nontrivial number of countries in the region have already established NGIs. These NGIs are approved through the EGI process, even though they are still not full members of the EGI council. This high level of maturity is the outcome of the EU co-funded projects, including the EUMEDGRID-Support. In short term, it is desired to evaluate the pros and cons of linking these countries directly to EGI rather than joining them with the emerging regional grid infrastructure (which is a preferred model in other regions). [short, CHAIN-REDS WP2]
3. **Clarify the role of Africa-Arabia ROC.** (*69-MED*) The role of Africa-Arabia ROC on whether or not to continue to coordinate the whole region, together with Sub-Saharan Africa, or whether to split the ROC in two, will need to be analyzed, discussed, and agreed upon. [short, CHAIN-REDS WP2 and Africa-Arabia ROC]
4. **Support to establish missing NGIs.** (*67-MED*) In countries where NRENs are still in the implementation phase, establishment of NGIs is a challenge. ASREN, being the regional coordinating body, will work with these countries (their NRENs in particular), and transfer

the necessary expertise to help establishing also NGIs (either joint with NREN or independent), as soon as the local political situation permits, eventually. [short—long, ASREN]

7.2.6. Funding

1. **Find suitable funding schemes.** (*68-MED*) ASREN together with NGIs should push on political consensus to provide funding schemes that would encourage universities and research centers to share resources in larger collaborations, rather than using them exclusively. The Association of Arab Universities and the League of Arab States can play important roles in developing these funding schemes. [medium/long, ASREN and NGIs]
2. **Follow successful EU co-funded projects.** (*71-MED*) NGIs together with their European partners will look for continuation of projects like EUMEDCONNECT, which push on extending the global collaborations in the region. NGIs will also need to seek funding from local and regional stakeholders and semi-governmental institutions, and respond to calls for proposals published by the Islamic development bank, the UAE Maktoum foundation, the Saudi AGFund, the Qatar Foundation, and more. [short/medium, CHAIN-REDS WP2, ASREN and NGIs]

7.3. Sub-Saharan Africa

7.3.1. Networking

1. **Network speed and reliability.** Although the AfricaConnect network has just been launched in Lisbon, in fact this is the launch of the rollout. Several UbuntuNet member NRENs have direct links to GEANT, however apart from TENET South Africa and KENET Kenya, at speeds of 155 Mbit/s. Seaboard countries have a lower cost threshold for entry. The reality is (particularly at the landlocked countries) some time will be needed before the recommended threshold of network bandwidth and reliability are reached. Realistic plans can aim at 155 Mbit/s. [medium, UbuntuNet and NRENs]
2. **Network development coordination. (***29-AF*) Regional and international reliable, high speed network backbone is critical for the grid acceptance for the same reasons as above, as well as enabling seamless cross-border collaboration. The regional backbone with 80% funding from EU, AfricaConnect, is already past the design stage and at the procurement stage. The NRENs in the region should coordinate in designing reliable regional backbone. UbuntuNet Alliance is recognized to be coordinating body in this field. Clear relationships of UbuntuNet are identified by the NRENs as giving leadership in GRID initiatives in the region. Each member country will have an UbuntuNet POP, so all the NRENs in the region will be established in the short term together with a coordinated strategy on the architecture of reliable peering of the NRENs in the region; however, achieving at least 2.5 Gbit/s connectivity to Europe (and USA preferably) will be a medium term objective. [short/medium, UbuntuNet]
3. **Interconnect isolated sites.** Isolated groups of researchers or small sets of collaborating institutes appear in the region. Thus isolated points or isolated patches of infrastructure occur, while good network connection to the NREN may not be affordable for those small groups. A targeted program to map the needs of such groups, and to find means (both technical and budgetary) should be developed in short term, with the goal to provide the connection to most such groups in medium term, including also newly appearing groups. CHAIN-REDS dissemination activities can help here. [short/medium, UbuntuNet, NGI, and CHAIN-REDS WP2]
4. **Demonopolize network provisioning.** Full monopoly, or market with only very few internet providers is rather typical in the countries of the region, yielding high prices of the connection, which block further development of the grid in the region. Effectively,

work at the policy level with the Regulators has been part of UbuntuNet's and the Association of African Universities' strategy for several years and has yielded some fruit. For example, TENET in South Africa had a critical role in breaking a telcom monopoly through forward purchase of IRUs on new fibre. Several NRENs such as Malawi and Mozambique have International Gateway Licenses from their Regulators allowing their infrastructure to cross borders without charge. Coordinated lobbying strategy to follow at both national and international level should continue, coordinated by UbuntuNet Alliance, and the local governments continue to be encouraged towards opening the internet market to competition. The strategy to be developed in short term to bring some results in medium term. [short, UbuntuNet]

7.3.2. Hardware Resources

1. **Grid interfaces to existing resources.** (*31-AF*) The general strategy (Section 5.6) should be followed [short/medium, NGIs and UbuntuNet]
2. **Provide seed resources. C**lusters exist in some countries and the work of UNESCO BrainGain has also added resources e.g. in Kenya. The general strategy (Section 5.5.2) should continue [short/medium, NGIs and UbuntuNet]
3. **Interface to supercomputers.** Super Computers that may be available for GRID applications in the region exist in Tanzania, Sudan and South Africa. Modalities for opening these resources to the wider GRID community were developed at the CHAIN round table in Dar es Salaam at the UbuntuNet-Connect conference in November 2012. This process should continue. [short, UbuntuNet and the relevant NGIs]

7.3.3. Software and Services

1. **Identity management infrastructure.** (*28-AF, 34-AF*) Management of trusted user identities is essential for smooth operation of the grid. Technologically it is achieved via approved certification authorities, and recently with the emerging identity federations (typically connected to human resources management systems of the involved institutions). UbuntuNet Alliance, as the representing body, should negotiate with IGTF the general approach for the region – either associating with one of the existing PMA, or setting up a specific one for the region. As the final goal, NGIs in each country in the region should setup certification authorities approved by the relevant PMA. The negotiations should start in short term, and full setup of CAs is realistic in medium term. In the meantime, the SAGrid CA can serve as catch-all CA for the region. Realistic options for setting up identity federation(s) in the region should be investigated. [short/medium, UbuntuNet]
2. **Primary middleware stack.** (*31-AF*) All countries in the region are on the way to accept the UMD middleware stack as the primary de-facto standard. Also for the user communities already collaborating with European partners this is a natural choice. Therefore UMD should be encouraged as the primary middleware stack for the grid sites in the region. Obviously, where there is the need for other stacks (Globus based services for smooth interoperation with OSG), parallel deployment should not be prevented. When necessary, virtualization technologies should be leveraged to allow co-existence of multiple middleware stacks on the same physical resources, CHAIN-REDS can help with provision of the technology roadmap [medium, CHAIN-REDS WP3]
3. **Deploy region-centralized services.** (*36-AF*) Central services necessary for independent operation of the grid infrastructure in the region (BDII, accounting portal, monitoring, GocDB, ticketing system, …) following the EGI pattern have been deployed in the northern region in Africa ROC, but a larger effort is needed to support the many countries of Africa in the short term. Africa ROC is the right body to coordinate this activity. Similarly, optional but desirable services (VOMS for VOs emerging from the region,

WMS for supported VOs, data catalogues, FTS, …) should be gradually deployed in medium term. Typically this is done by institutions participating in scientific work which requires the service; however, NGIs should offer provisioning of some of them as well, due to concentrated expertise. [short, Africa ROC and UbuntuNet]

4. **Integrate applications in portals.** (*32-AF*) The general strategy (Section 5.3) should be followed. In short term, portals provided by collaborating European institutions can be used by users in the region and to access the regional services, in medium term an independent regional portal(s) should be provided (probably by one of the more mature NGIs in the region, e.g. SAGrid). CHAIN-REDS can provide the technology support. [medium, SAGrid and CHAIN-REDS WP5]

7.3.4. User Communities

1. **Specific support to astronomy.** Astronomy is an outstanding scientific discipline in the region. In particular, the facilities of the Square Kilometer Array project will be hosted there. Therefore NGIs should establish close contacts with the regional scientific communities participating in these large international projects, understand their requirements on the computing and data infrastructures in detail, and provide them with the necessary technological support. In this way, the regional scientific teams are getting eligible to become equal partners of the international collaboration also from the e-infrastructure viewpoint. Coordination of these activities can follow the proven pattern of EGI virtual teams. The first contacts should happen in short term, so that in medium term the needs of the community are well understood and the regional collaboration is operational. [short/medium, relevant NGIs]
2. **Support to Atmospherics and Climate Science**. Interest in this discipline is coming from Cameroon and several countries in the Ubuntunet Alliance catchment. CHAIN has been asked to assist in research design in a major International climate project for the Lake Victoria region that covers several UbuntuNet member countries around and near Lake Victoria. Obviously it is CHAIN-REDs that will take up this opportunity [short, CHAIN-REDS WP2 and NGIs]
3. **Focus on small and emerging user communities.** Besides large, well-established user communities there are numerous small and starting communities in the region. However, when aggregated, these communities represent non-negligible mass of potential infrastructure users. Specific support is required to bring them to the use of the grid. This must be elaborated by all NGIs individually. General guidelines to this work will be provided by the CHAIN-REDS project. [medium, NGIs and CHAIN-REDS WP2]
4. **Support existing collaborations.** Large international collaborations (HEP, We-NMR, WISDOM, …), which are typical consumers of the grid resources, are present in the region. However, in many cases this is done through contribution of individuals or small groups who get associated with European or US organizations. Specific support should be provided to those regional communities to get organized at the national and regional levels and to back up them with the required technology and grid expertise in order that they become equal partners to the global collaborations. For example in Genomics Research there are research teams in several countries such as Kenya and Malawi partnering with the Sanger Institute in Cambridge UK doing genomics research on such areas as non-typhoidal salmonella. [short—long, relevant NGIs]

   We-NMR is reported to have contacts in Burundi. The team should be contacted together with eligible resource providers, and a plan for resource commitment to the We-NMR VO should be discussed. At the middleware level the integration will go smoothly then due to the UMD-compatible middleware and use of PMA-approved catch-all SAGrid CA (before PMA approval is finished, cf. Section 7.3.3. [short, UbuntuNet]

Similarly, interest in DECIDE is reported in Nigeria. The team and eligible resource providers should be contacted and plans for further collaboration – involvement of users and integration of resources through Science Gateway should be elaborated. [medium, UbuntuNet]

7.3.5. Organization

1. **Finalize NRENs establishment.** NRENs are established in most of the countries. Moreover the model of NREN taking the role of NGI was adopted by significant part of them, and it seems be appropriate for the region. UbuntuNet Alliance is in ongoing contact with potential participating organizations in countries where NRENs were not established yet, (e.g. Zimbabwe) and it will provide them with the experience to help finding appropriate organizational models for the individual countries. Due to political will, history and economic reasons establishment of NRENS (which are ready to take the NGI role in turn) in all countries in the region may be realistic in medium term. [medium, UbuntuNet]
2. **Establish NGIs as legal entities.** Aim at establishing NGIs (together with NRENs or independently) as legal entities under the law of their home countries. Such bodies become eligible for long-term funding of the centralized activities, eventually including also central resources. Also this level of organization increases the weight of the grid community in influencing the development and funding strategies of the local governments. CHAIN-REDS can help exchanging know-how on the legal and organizational issues. [medium, UbuntuNet, the NGIs, and CHAIN-REDS WP2]
3. **Formalize interoperability agreements.** (*33-AF*) Despite the grid infrastructure is mostly organized on best-effort basis, certain formal or semi-formal agreements defining policies and the level or service are necessary once the infrastructure size grows. In particular, policies of certification authorities should be coordinated and approved by PMA (see also "Identity management" in Section 7.3.3). Templates for Service Level Descriptions for the connected sites should be prepared by NGIs. EGI documents can be taken as inputs; however, more relaxed approach is appropriate for the region. [short, NGIs]
4. **Clarify the role of Africa-Arabia ROC.** Africa-Arabia ROC is operational, and currently it serves countries in North Africa, Middle East, as well as Sub-Saharan Africa. The role of the organization w.r.t. the Sub-Saharan Africa should be clarified in short term, whether it is foreseen to continue serve all the regions or whether a fork of Sub-Saharan Africa specific body is expected. In both cases the organizational scheme should be described in a semi-formal way, e.g. by MoU of the participating NGIs. [short, Africa-Arabia ROC and UbuntuNet]
5. **Training and dissemination.** (*29-AF, 30-AF*) Targeted training and dissemination events should be organized by NGIs under the coordination of UbuntuNet. In fact there are several requests for various types of training. These events should aim at both site administrators, training them in setup and operation of the grid services, as well as the user groups. In the latter case, the training should be specialized to the scientific area of the particular user group – the users should learn directly how to address their scientific problems with the help of the grid services. This approach is necessary to overcome the wide-spread "desktop centric" computing paradigm in favor of the grid one. Identification of trainers should be investigated. The dissemination opportunity of UbuntuNet annual conference has been invaluable in the two years of the life of CHAIN. The UbuntuNet e-newsletter, NUANCE, has also featured CHAIN and GRID activities regularly, and invited articles from VRCs. [NGIs, UbuntuNet, and CHAIN-REDS WP2]

7.3.6. Funding

1. **Infrastructure support projects.** HP/UNESCO Brain Gain initiative left visible traces in the region, being a huge contribution to the recent achievements of the infrastructure development. UbuntuNet, together with NGIs as well as support of the local governments should look for participation in similar programmes, as well as lobby for setting up further ones. [NGIs, UbuntuNet]

7.4. Asia-Pacific

7.4.1. Networking

1. **Strengthen inner area connection.** e-Science collaboration is an effective driver to reinforce the network infrastructure by leveraging regional and international resources for scientific, engineering, and industrial advancement. Some APGI partners do not have sufficient bandwidth connecting to the regional hub to support distributed collaborations. Continuous support and encouragement for regional e-Science applications should be provided and coordinated. Gigabit level bandwidth between APGI partners to a regional hub is the target in next three years. [medium, APGI and individual partners]

7.4.2. Hardware Resources

1. **Strengthen infrastructure in region.** There has been a strong progress in past years with developing new computing and storage resources within the region. Most progress is seen in Taiwan, Malaysia and other countries participating in EUAsiaGrid. Like in networking area, the region should focus on strengthen its infrastructure in all parts, not only in Taiwan. (*D2.2 pg. 25, D2.3 pg. 107 and D4.1 pg. 36*). This should be done in the harmony with collaboration and production applications to attract more users and communities in the medium term. [medium, APGI and its partners]
2. **Improve reliability.** Resource availability, which is now about 90% should be raised more. According to last reports this is already in progress, so APGI should keep this drive and continue in improvements. Resource centre without engagement of active user communities would not have enough motivation to achieve production quality. Besides the commitment (in SLA) and APROC coordination, the best policy is to build up close collaboration with local user communities. Site will be suspended or even expelled based on the guidelines of EGI as well as the requirements from VO. [medium, APGI]

7.4.3. Software and Services

1. **Bring services and users together.** Many production e-Science applications have been in place since EGEE Asia Federation, EUAsiaGrid and EGI era. The strategy to focus on regional common interests and to engage regional highly motivated user groups with the international VO proved to be successful. What was learned is that the user community has to be involved in the application development from the beginning. Every user group likes to have tailor-made applications. However, the resource provider aims for generic application services for as many disciplines and users as possible. The Application Portal and the Virtual Research Community (VRC) are most favourable solutions. As the sustainability described in D2.3, the user community, resource provider and application provider are all indispensable. All three parties have to work together closely all the time to build up useful e-Science application, environment and collaboration model closer to users.

   Life science, natural disaster mitigation (earthquake, tsunami, etc.), weather and climate simulation, high energy physics are the current active domains among APGI partners. APGI has demonstrated a feasible model of e-Science in Asia Pacific to have coherent regional level collaboration with more flexibility on the national level. Without a doubt,

the international drivers such as EGI and CHAIN are instrumental in APGI development. [short—long, CHAIN-REDS WP5, APGI and involved VRCs]

2. **Interconnect middleware.** Some middleware usage is compatible with EGI (gLite and EMI), some is compatible with Globus and OSG – (*D4.1, pg. 36*). Interoperability should be secured via Science Gateways,, preferably by setting up a regional instance (ASGC is an obvious candidate to set it up). Possibilities of deploying these instruments would be useful to investigate with the increasing number of users. [short/medium, CHAIN-REDS WP5 and ASGC]
3. **Setup infrastructure of certification authorities.** Only when there are enough real users, the need of a local certification authority (CA) emerges. However, due to the existence of APGridPMA establishing national CAs is a relatively straightforward process (comparing to other target regions). APGI is not just extending the coverage of collaboration but also working closely with APGridPMA to encourage partners to establish their CA infrastructure. Additionally, building identity federations can be considered. [medium, APGI]

7.4.4. User Communities

1. **Continue dissemination activities.** Attracting more users and communities will secure the funding scheme as well as the sustainability. The APGI dissemination activities such as communication with communities and their support are in good course and should continue in this area. ISGC 2012 revealed two new portals for drug discovery and weather simulation [short—long, APGI and CHAIN-REDS WP2]

7.4.5. Organization

1. **Keep the current organization drive.** APGI seems to be a good model of functioning for the region. However it should encourage developing national level organizations for the purpose of stimulating participation in the collaboration from users from more disciplines and maintaining the certification authorities to support increasing number of users (Section 7.4.3). Before the establishment of APGI formal body, the enhancement of regional collaborations in virtual model will keep going. [medium, CHAIN-REDS WP2 and APGI]

7.4.6. Funding

1. **Support the funding appeals.** There should be a clear strategy on sustainable funding. According to SWOT analysis in D2.3 there is neither a regional funding model nor an assurance of national level funding support. To show the importance of sustainable funding to governments, more resource providers and user communities should be attracted to join APGI. Some support can also come from international cooperation so attracting more VRCs could help. The clear strategy on sustainable funding should be developed in the short term, before the current funding exhausts. [short, APGI]
2. **Strengthen APGI.** Every organization is stronger with more partners involved. Attracting more resource providers and user communities and involving more partners in APGI is recommended in medium term. [medium, APGI]

7.5. China

7.5.1. Networking

1. **Better interconnection of CSTNET and CERNET.** According to D2.3 pg. 23, there is an expectation to have 5Gbps connection to Europe till 2014 and 10Gbps till 2017. However in China there are two separate academic network providers CSTNET (China Science and Technology NETwork) and CERNET (China Education and Research

NETwork) which need more properly interconnect each other. For better cooperation it is necessary to raise in short term the current connection of 1Gbps to at least 5Gbps and in medium term to 10Gbps.
Connect all sites to backbone via at least 100Mbps links in short term. In medium term upgrade to gigabit links. (*D4.1 pg. 37*). [short/medium, CSTNET and CERNET]

2. **Develop better IPv6 connectivity.** (*D2.3, pg 104, point 3*) While there is a lack of IPv4 addresses, it would be very useful to start providing IPv6 connectivity. Strategy formulation from D2.3 recommends using IPv6 as the alternative backbone. [medium/long, CSTNET and CERNET]

7.5.2. Hardware Resources

1. **New resources.** Lot of funding and effort was put in the HW infrastructure in past years. Now, according to D2.3 pg. 99 and 100, China has a 12th national 5-year plan till 2015 to "*focus of the new-generation information technology industry will be on next generation telecommunication networks, next generation internet technologies, internet of things, triple network convergence (telecom, computer and cable TV networks), cloud computing, integrated circuits, new generation displays, high end software, high end servers and information services*" which is translated as "*support HPC and cloud computing*". So the hardware resources will still grow, although this way of growing may influence the usage of grids.

7.5.3. Software and Services

1. **Interconnect EMI and GOS.** The users should profit from using the sites running EMI as well as sites running GOS. There are both systems available in China (*D4.1 pg. 36 - 37*). This task can be accomplished by using of scientific gateways supporting both systems. The first kick could be the CHAIN scientific gateway[5]. Implementation and start of using of scientific gateways supporting Chinese and European systems together should be the short term task, including setup of a regional Science Gateway instance (IHEP may be a good candidate). [short/medium, CHAIN-REDS WP5 and IHEP].
2. **Establish national CA.** There are some independent certification authorities, according to SWOT analysis in D2.3 pg. 103, but there is no national level certification authority yet. The situation should be clarified, and either a single CA or a system of CAs approved by APGridPMA should emerge. [short/medium, CHAIN-REDS WP2, CSTNET and CERNET]
3. **Leverage cloud infrastructure.** There is a clear emphasis on clouds in the Chinese government plans, and significant funding will go in this direction. Therefore the integration of the grid infrastructure with the cloud resources is of particular importance in China (*D2.3 pg. 104*). Clear interoperability plan, using also the Science Gateway paradigm, should be developed. [short, CHAIN-REDS WP3/5]

7.5.4. User Communities

1. **Develop national and international cooperation.** It seems that the cooperation among user communities on national level needs to be improved. This is visible for example via lack of national CA and NGI yet. International collaboration is also a challenge according to SWOT analysis in D2.3. There should be some national (to support user communities across more grid sites) and international (to support EGI VOs) cooperation model (with science gateways support) in middle term. (*D4.1 – pg. 37*) [medium, CHAIN-REDS WP2]

[5] http://science-gateway.chain-project.eu/

2. **Support existing collaborations.** Within the framework of the INDICATE project a pilot portal to access a digitalized tiny fraction of the Chinese cultural heritage was provided (http://indicate-gw.consorzio-cometa.it/chinarelics-dr). Provision of this service should be maintained and extended. [long, CERNET]

7.5.5. Organization

1. **Evaluate the need for NGI.** The technical connection between CNGrid and China ROC is in the test stage. However, CNGrid and China ROC services should be interconnected more thoroughly also at the organizational level, emerging into formal establishment of NGI-like body eventually, if the application demand would increase. [medium, CNGrid and China ROC]
2. **Define SLA/OLAs.** The D4.1 survey analysis raw data shows that there is no regional SLA for sites. Establishing of NGI is a good opportunity to do that. Establish and monitor regional SLA/OLA if there is no such mechanism yet. [medium, CNGrid and China ROC]

7.5.6. Funding

1. **Find sustainability model.** As mentioned in previous Sections, a clear vision of grid funding and sustainability is needed due to the governmental focus on HPCs and clouds. D2.3 pg. 104 targets this issue and promotes establishing the NGI and through this and other activities, such as conferences and workshops, to escalate the awareness of various officials on grid computing. CHAIN-REDS WP2 will help in this task so the middle term is the right scope. [short, China ROC]

7.6. India

7.6.1. Networking

1. **Continue in NKN project.** The National Knowledge Network (NKN) project is aimed at establishing a strong and robust internal Indian network which will be capable of providing secure and reliable connectivity. Using NKN, all vibrant institutions, with vision and passion, will be able to transcend space and time limitations in accessing information and knowledge. They will, thus, derive the associated benefits for themselves and for the society. Establishing the NKN is a significant step towards ushering in a knowledge revolution in the country with connectivity to 1500+ institutions. The NKN is intended to connect all the knowledge and research institutions in the country using high bandwidth/low latency network. The NKN was approved in March, 2010, by the Government of India [short—long, NKN]

7.6.2. Hardware Resources

1. **Develop technologies upon GPGPU/FPGA.** (*D2.3 pg. 94*) There are quite good hardware resources so far so there is time to think about extend traditional HW by some corresponding to current trends. Purchasing of GPGU/FPGA resources will bring new computing possibilities to user communities and raise their expertise. A data centre is already being planned in the later part of the 12th five year plan period (2012-2017) of the Government of India. [short—long, NKN]

7.6.3. Software and Services

1. **Develop connection to other computing infrastructures.** There is no documented evidence of coordination of cloud and perhaps voluntary computing with grid computing. D2.3 pg. 94 recommends provision of cloud interface to grids; the same should be developed for voluntary computing, if it is present in the region. [medium, NKN and CHAIN-REDS WP3]

2. **Interoperability with WLCG sites.** Historically the Indian part of WLCG and the GARUDA grid evolved independently, using also incompatible middleware. The resources should be interoperable, so the adoption of some solution is recommended. This technology can be built with use of the CHAIN science gateway[6] in short term. The interoperability of the GARUDA grid with the WLCG sites has been already established, and the activity should continue to larger extend [medium, NKN]

7.6.4. User Communities

1. **Continue dissemination.** Resource sharing and expertise exchange is being regularly done through Workshops, Conferences and Seminars on the NKN and the GARUDA grid. Keep this drive for long term. [short--long, NKN]
2. **Support existing communities.** The NKN has the capability to handle high bandwidth with low latency and provision to overlay grid computing. Some of the grid based applications are climate change / global warming, science projects like Large Hadron Collider (LHC) and ITER. The NKN can be the platform to realize many such innovative applications. The Garuda Grid has enhanced its power and stability by migrating to NKN. Needs of these established communities should be monitored and reflected whenever possible. [medium—long, NKN]
3. **Specific support to climate modeling and agriculture.** These disciplines have specific position in the region, due to the high political and economic impact, while the need for extensive computing and data storage resources is apparent. There already exist contacts between NKN and the communities. Those should be extended and strategic plans defined to allow long-term presence of these applications on the infrastructure. [short, NKN]

7.6.5. Organization

1. **Coordinate GARUDA/NKN Grid and WLCG.** According to D2.2 pg. 30, there are two concurrent grid infrastructures in India. Both are cooperating with CHAIN and with NKN but there is no evidence of cooperating mutually. At least some MoU about joint grid activities should be signed and implemented. [short, NKN, WLCG and CHAIN-REDS WP2]
2. **Make strategy on clouds and voluntary computing.** A strategy on clouds and voluntary computing will be planned in the future. As mentioned in Software and Services Section, there is no evidence in CHAIN deliverables about GARUDA dealing with cloud and voluntary computing. Except the developing of corresponding SW, there should be also some strategy. (*D4.1 – pg. 37, D2.3 pg. 93 - SWOT*) [short, NKN and CHAIN-REDS WP3]
3. **Establish SLAs.** Without SLA/OLAs it is difficult to evaluate and enforce the reliability of individual grid sites and therefore worse possibility of make the infrastructure better usable for users. Regional SLA/OLA among NKN and participating grid sites should be established. [medium, NKN and sites]

7.6.6. Funding

1. **Long term funding plans.** The NKN and the GARUDA grid project are being regularly funded by the Government of India through its 5 year plan allocations. The current 5 year plan period has commenced on the 1st of April, 2012 and will end of the 31st of March, 2017.This is pretty good situation with respect to many other regions, nonetheless long term funding sources should be pursued either. [long, NKN]

[6] http://science-gateway.chain-project.eu/)

8. Conclusions

This deliverable summarizes and concludes activities performed by CHAIN project partners in a form of a Roadmap, a document with clear and concise recommendations for individual grid and research communities working to build and operate interoperable grid infrastructure. In its four main Sections, a description of necessary and optional follow up activities covering all important aspects of managing and extending grid infrastructures is provided. To make the recommendations easy to follow, and also to guarantee that activities common to all world regions are taken care of in a similar way, the first three Sections deal with common recommendations on organizational (non-technical) and technical issues and also deal with the proper interaction with research communities worldwide. Only the last of the four Sections is dedicated to more specific recommendations targeting individual regions or even countries, taking into account their maturity, cultural and technological background while still keeping in focus the goal of setting and operation a worldwide interoperable distributed computing infrastructure.

The Roadmap also provides hints for the CHAIN-REDS project, especially in areas that need further focus and are strictly related to the CHAIN-REDS description of work (big data, cloud infrastructures, novel authentication approaches like identity federations, and the particular role of Science gateways).

As such, the Roadmap should serve as a guide to all involved actors interested (and active) in the setup and operation of world-wide distributed computing and data infrastructure able to seamlessly support large-scale scientific collaboration spanning countries and regions.