PROCEEDINGS OF SCIENCE

Introduction to the CHAIN-REDS Project: objectives and achievements

F. Ruggieri [1]
INFN Sezione di Roma Tre
Via della Vasca Navale 84, Roma, Italy
E-mail: Federico.Ruggieri@roma3.infn.it

G. Andronico
INFN Catania
Via S. Sofia 64, Catania, Italy
Giuseppe.Andronico@ct.infn.it

R. Barbera
University of Catania and INFN Catania
Viale A. Doria 6, Catania, Italy
Roberto.Barbera@ct.infn.it

L. Matyska
CESNET
Zikova 4, 162 00 Prague, Czech Republic
E-mail: {ljocha,ludek}@ics.muni.cz

R. Mayo Garcia
CIEMAT
Avda. Complutense, 40, 28040 Madrid, Spain
Rafael.Mayo@ciemat.es

O. Prnjat
GRNET
56, Mesogion Avenue, Athens, Greece
Oprnjat@grnet.gr

The CHAIN-REDS project started on the 1st December 2012 for 30 months duration. During the first 11 months of activity has produced good results in relation to the coordination of Distributed Computing Infrastructures, Data and Document repositories, Science Gateways and Federated Identities.

[1] Speaker

PoS(eIeS2013)001

PROCEEDINGS OF SCIENCE

e-Infrastructures for e-Sciences 2013 A CHAIN-REDS Workshop organised under the aegis of the European Commission (eIeS 2013)
October 22, 2013
Beijing, P.R. of China

1. Introduction

The CHAIN-REDS project started on 1st December 2012 for 30 months duration and with a considerable background of knowledge and experience maturated during a number of previous projects dealing with collaboration between European e-Infrastructures and those at other continents.

The consortium has gathered key e-Infrastructure partners, currently 10, from different regions of the world with a strong core of EU organisations already expert in e-Infrastructures and International Cooperation.

CHAIN-REDS vision is to promote and support technological and scientific collaboration across different e-Infrastructures established and operated in various continents, in order to facilitate their uptake and use by established and emerging Virtual Research Communities (VRCs) but also by single researchers, promoting instruments and practices that can facilitate their inclusion in the community of users. The project aims to support and disseminate the best practices currently adopted in Europe and other continents, together with promoting and facilitating interoperability among different regional e-Infrastructures.

The core objective of the CHAIN-REDS project is to promote, coordinate and support the effort of a critical mass of non-European e-Infrastructures for Research and Education (R&E) to collaborate with Europe by addressing interoperability and interoperation of Grids and other Distributed Computing Infrastructures (DCI). From this perspective, CHAIN-REDS will optimise the interoperation of European infrastructures with those present in 6 other regions of the world, both from the development and use points of view, thus catering to different communities.

Overall, CHAIN-REDs will provide inputs for future strategies and the decision-making mechanism, regarding collaboration with other regions on e-Infrastructure deployment and availability of related data. The project will also raise the visibility of e-Infrastructures towards intercontinental audiences, covering most of the world and provide support to establish globally connected and interoperable infrastructures, in particular between the EU and developing regions.

In the following chapters we describe the current state of the art of the activities and the results obtained so far.

2. Distributed Computing Infrastructure & Regional Operation Centres

The activity makes reference to the Work-Package 3 of the CHAIN-REDS project, where an important document has been produced [1] regarding Interoperation model and plan.

In this deliverable an interoperation model between Europe and the regions participating in CHAIN-REDS was presented. The first part of the document analyses the structure and operational model of the European Grid Infrastructure, while, in the second part of the document, a blueprint for the implementation of Regional Operation Centres is presented. Two possible scenarios were identified for interoperation with Europe:

- in the first scenario a Regional Operation Centre can closely collaborate and interoperate with EGI as an “Integrated Resource Infrastructure Provider”, while

- in the second scenario, the ROC can base its interoperation with EGI as a "Peer Resource Infrastructure Provider".

Detailed guidelines are given in the document in terms of functionality, manpower needed, as well as detailed technical recommendations.

The third part of the document presents the current state of the Grid infrastructures in each region. Based on this information, a concrete action plan for each region is drafted, taking into consideration the operational and technological characteristics, along with the level of compatibility.

The CHAIN-REDS project has in parallel strongly supported the existing ROCs, providing guidance, suggesting improvements and best practices to be applied. The effort in this first year has mainly been concentrated on implementations in Latin America (ROC-LA), Africa and Arab Countries (Africa&Arabia ROC) and China (China ROC).

3. Clouds for Research and Education

The preliminary investigation made by the project has made evident that the European activities in federating the Clouds for Research and Education are among the most advanced in the world. The EGI Federated Cloud Task Force has thus been taken as the starting point for such activity. The general idea has been to extend such experience along two lines:

1. Use OCCI and CDMI standards as recommended by the EGI Task Force extending the number of cloud implementations, including those in other regions of the world addressed by CHAIN-REDS;
2. Add an orchestration system, based on CLEVER, that would allow the deployment of VREs across several Cloud providers.

It was decided to make a demonstration of what is currently feasible at the EGI Technical Forum in Madrid in September 2013. The Cloud implementations were selected by asking resources to partners and external contributors with the only limitation of having an OCCI compliant interface. The resulted list is shown in Table 3.1.

The demonstration was made creating an application accessible by the CHAIN-REDS Science Gateway (see chapter 5 below) and aimed at presenting two use cases:

- Use case #1 (scientist)

  A user can sign in on a Catania Science Gateway using his/her federated credentials, select an application from a menu and seamlessly execute it on either HPC machines, Grids or Clouds.

  The fractions of executions on the three different platforms can be adjusted to simulate the need to "boost" the resources in case of temporary peaks of activity; e.g.during the demo in Madrid, a couple of applications were deployed on different middlewares:

  1) on Grid sites;
  2) on the HPC site of FZK-Juelich in Germany running UNICORE;
  3) on the "local" HPC site of ENEA-Portici in Italy running bare LSF;

4) on VMs that were deployed at some EGI FedCloud sites.

Table 3.1- List of resource providers with Cloud implementations, participating in CHAIN-REDS demo session.

Country	Cloud Middleware	Resource Provider
Czech Republic	OpenNebula	CESNET
Greece	okeanos	grnet
Spain	openstack	CETA Ciemat
Spain	OpenNebula	Ciemat
Italy	OpenNebula	INFN
Italy	openstack	University of Messina
South Africa	OpenNebula	SOUTH AFRICAN NATIONAL GRID
Egypt	openstack	it synergy

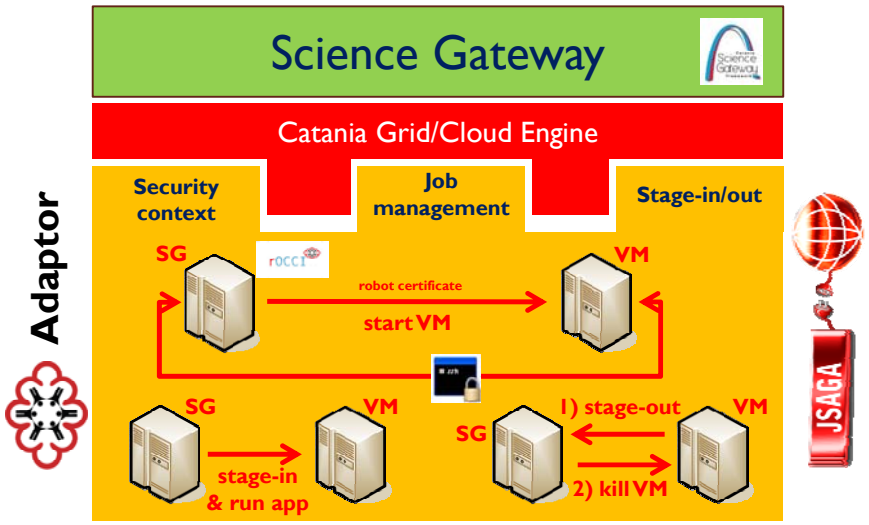

Figure 1 - Use case #1 implementation: JSAGA Adaptor for OCCI

Since there were few cores available on the EGI FedCloud (it's only a Test Facility) the weights of job submission set in the Science Gateway was: 60% on Grid, 20% on UNICORE, 10% on LSF and 10% on OCCI-compliant clouds.

In case of larger cloud-based resources than grid- or other-middleware-based ones, the ratios could have been modified accordingly.

This Use Case required the development of a specific JSAGA adaptor for OCCI as depicted in Figure 1.

- Use case #2 (cloud tenant)

  The cloud tenant of a real or virtual organisation can sign in on a Catania Science Gateway using his/her federated credentials, select virtual machine(s) from a geographically shared repository and deploy/move/copy it/them across his/her personal cloud.

  The graphical user interface is very intuitive and including point & click and drag & drop functionalities.

  The virtual machine(s) belong to the same domain name (chain-project.eu in the particular case) independently of the site where it/they will be instantiated and of the underlying Cloud middleware stack.

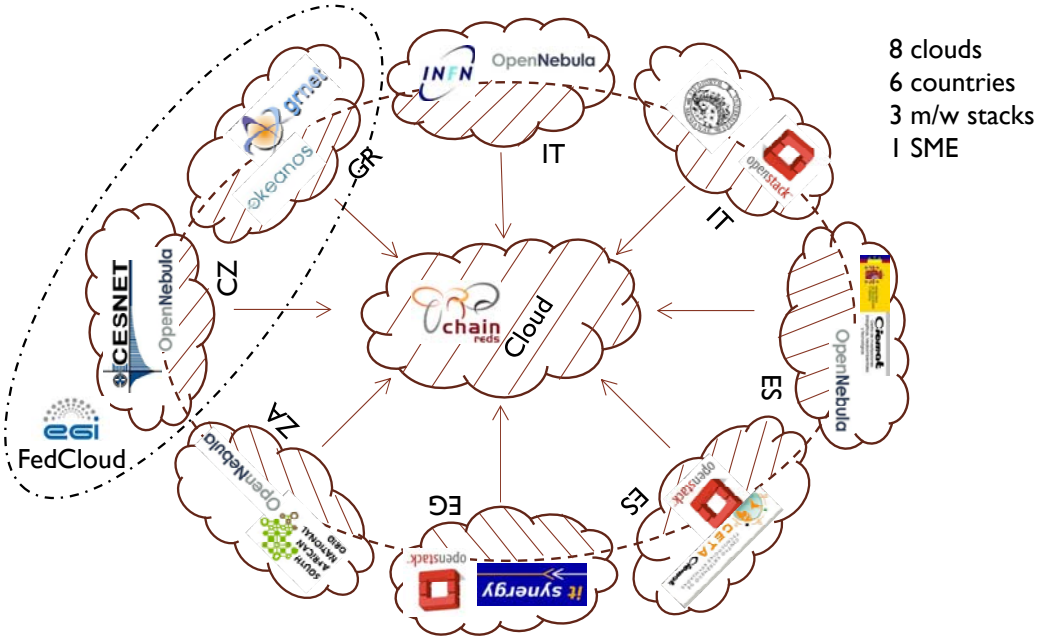

Figure 2 - Actual test-bed configuration for use case #2

The Use Case 2 required the deployment and customisation of the CLEVER Middleware and the setting up of a graphical interface to control and manage the Cloud. The applet used has been named "MyCloud" and the current architectural implementation is shown in Figure 3.

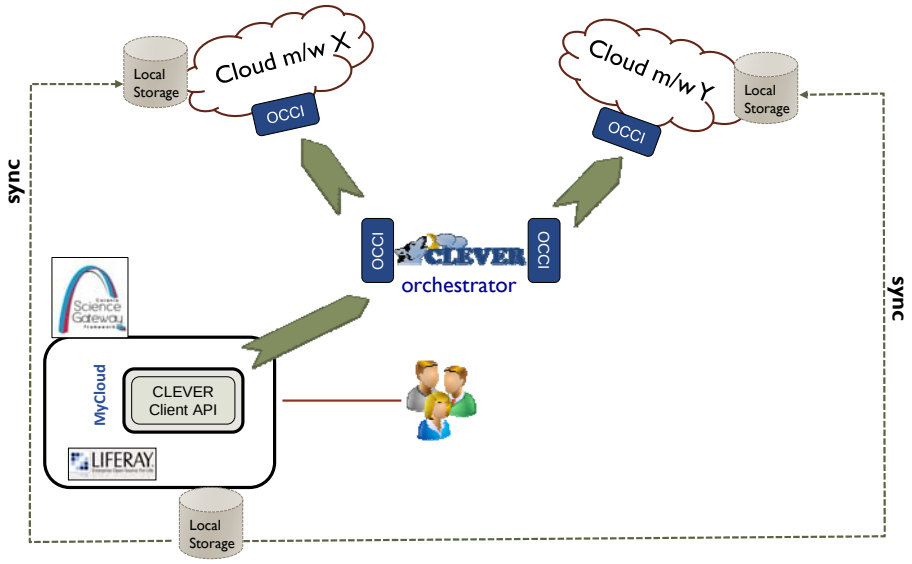

Figure 3 - MyCloud Architecture: current implementation

CLEVER orchestrates the Computing managing stateless Virtual Machines.

Towards the end of Year 1, the project has formulated a very detailed cloud survey, which will be circulated amongst the world regions in year 2. On the basis of the results, state of the art will be analysed and guidelines and recommendations produced, complementing the above cloud demo and setting the path for long-term harmonisation.

More detailed information can be found in [2].

4. Data Infrastructures and Repositories

The extension of the CHAIN-REDS Knowledge Base (KB) has been one of the most relevant activities during this first year. The KB is one of the largest existing e-Infrastructure-related digital information systems. It currently contains information, gathered both from dedicated surveys and other web and documental sources, for more than half of the countries in the world.

It can be accessed from the project's web site (www.chain-project.eu) and presents the information to visitors through geographic maps and tables.

Besides e-Infrastructure sites, services and applications, the CHAIN-REDS KB publishes information about Open Access Document Repositories and Data Repositories.

CHAIN-REDS consortium has also decided to semantically enrich OADRs and DRs and build a search engine on the related linked data. The CHAIN-REDS Semantic Search Engine is also accessible from the project's web site. Using it, visitors can either enter a keyword and submit a SPARQL query to the Virtuoso triple store or select a language and get, on the left side of the page, the list of subjects available in that language with the indication, between parentheses, of the number of records available for that particular subject.

The results of the query are listed in a summary view and for each record found, the title, the author(s) and a short description of the corresponding resource are provided. In the "Dataset information" panel users get the link to the open access document and, if existing, to the

corresponding dataset. Clicking on the "Graphs" tab, which appears at the top of the summary view, users can select one or more of the resources found and get a graphic view of the semantic connections among Authors, Subjects and Publishers.

Figure 4 - Open Access Document Repositories. Red markers are info from DRIVER, OpenAIRE, and OpenDOAR Yellow markers are new added by the project (e.g. La Referencia in Latin America)

The CHAIN-REDS search engine actually performs a single search on the project repositories and also a parallel one taking into account the ENGAGE repositories too. Deeper explanation about this and other characteristics can be found in [3] and [4].

5. Science Gateways

The CHAIN-REDS project has continued to promote the use of Science Gateways (SG) and has further developed the CHAIN SG based on the implementation made by the INFN Catania Group.

The new developments have been concentrated to widen the possible DCIs that could be used and specifically the Cloud implementations.

As already mentioned the project has deployed the CHAIN-REDS Science Gateway as a portal to interface different types of Distributed Computing Infrastructures (DCIs): local cluster, High Performance Computing (HPC), cloud and grid.

The portal, built with the Catania Science Gateway Framework, was presented at the EGI Technical Forum (16-20 September 2013) in Madrid, Spain, by Roberto Barbera and Giuseppe La Rocca (INFN, Catania) in a demo entitled "Managing and using interoperable DCIs through a standard-based Science Gateway". The demonstration showed the possibility of

interoperability, at application level, among different DCIs, including Clouds, using OCCI and SAGA as standard interfaces, and the CHAIN-REDS Science Gateway as Virtual Research Environment (see details in previous chapter 5).

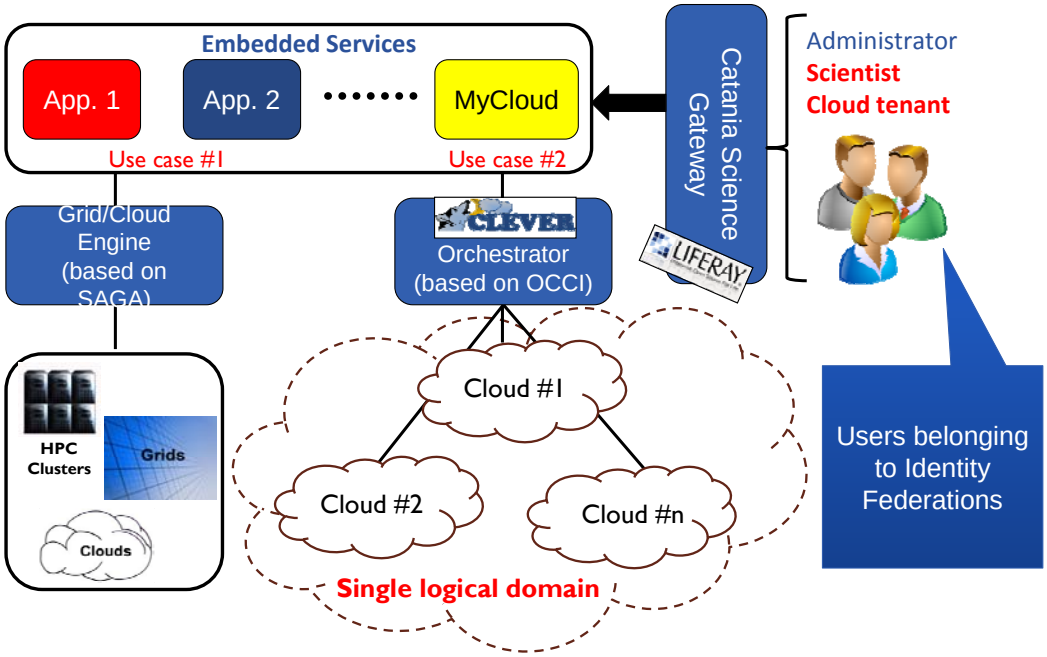

Figure 5 - Science Gateway schematic diagram

6. Federations of Identity Providers

Promoting Federations of Identity Providers (IdP) is a key action of the CHAIN-REDS project and other activities, such as Science Gateways, are based on users' authentication by means of IdPs.

In the Kick-Off Meeting in Dubai (December 2012), detailing the work-plan, it was decided to organise an event at the TERENA Conference in 2013 with the aims of:

- Promoting the deployment of IdPs and Federated IdPs;
- Review the status of the regions addressed by CHAIN-REDS concerning Federated Identities initiative and programs.

The project thus organized a Birds-of-a-Feather (BoF) session during the TERENA Conference 2013 (TNC2013) in Maastricht, the Netherlands on June 4, 2013, with the title: "Federated Identities in Global context".

Major part of the BoF consisted of presentations given by speakers from several world regions (China, Africa, Arab and Mediterranean countries, Latin America), providing up to date status and experiences with deploying distributed authentication approaches.

There was also a presentation of Perun system, developed by CESNET, which aims at helping with establishing Identity Federations (IF). The system provides tools to manage user's

enrolments, group management and resource management which are essential for IFs and/or Certification Authorities (CA), helps with user management and simplifies authorization setup.

The main conclusions from a discussion that followed BoF presentations can be summarised as follows:

- participants to the BoF underlined the importance of coping with the likely development of several IFs in China, and possibly other regions;
- the top-down approach presented by Siju Mammen was considered as being very promising for national IFs;
- it was agreed that the best choice between the two types of IFs (hub & spoke/centralised and meshed) depends on specific local technical and legal conditions;
- it was acknowledged that, when establishing a federation, eduGAIN templates for policy documents can be directly used;
- the Eduroam service should be promoted, as it is a step towards web based IF. During the BoF a promise to help with Eduroam setup has been offered from Terena.

The input from this BoF has been used in the preparation of the major AA-related deliverable [5] that has been published at the end of the first year of the project.

7. Conclusions

The CHAIN-REDS project has completed the first year of activity with several achievements. The work in Clouds and DCIs has led to the presentation of a demonstration; while the Grid operations are supported by the ROCs.

The coordinated activities of promoting Federated Identities have involved other projects and institutions with concrete results in Africa.

The focus on data and document repositories has produced the extension of the CHAIN-REDS Knowledge Base and a new search engine.

The project has successfully presented its results in three Workshops and in other events in many regions addressed by CHAIN-REDS.

References

All the following documents can be downloaded from: http://www.chain-project.eu/deliverables.

[1] CHAIN-REDS Deliverable D3.1 – “Interoperation model and plan”

[2] CHAIN-REDS Deliverable D3.2 – “Interoperability guidelines and design“

[3] CHAIN-REDS Deliverable D4.1 – “Transcontinental Data Infrastructures and Data repositories”.

[4] CHAIN-REDS Deliverable D4.2 – “Analysis of Data Infrastructure and Data repositories”.

[5] CHAIN-REDS Deliverable D5.1 - “Federations and other alternate AA mechanisms”