Project acronym : CHAIN-REDS
Project full title : Co-ordination & Harmonisation of Advanced and e-INfrastructures for Research Education Data Sharing
Grant agreement : 306819
Start date : December 1, 2012
Duration : 30 months
Programme : 7th Framework Programme (FP7)
Theme : Capacities specific program
Thematic area : Research Infrastructures
Funding scheme : Support action
Call identifier : FP7–INFRASTRUCTURES–2012-1
Project coordinator : Federico Ruggieri (INFN)

D3.3 - Cloud Technical Survey

Deliverable Status : Draft
File Name : CHAIN-REDS-D3.3_V04
Due Date : May 2014 (M18)
Submission Date : May 2014 (M18)
Dissemination Level : Public
Author : CIEMAT (manuel.rodriguez@ciemat.es)

INFN	Istituto Nazionale di Fisica Nucleare - Italy
CIEMAT	Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas - Spain
GRNET	Greek Research and Technology Network S.A. - Greece
CESNET	Zajmove Sdruzeni Pravnickych Osob - Czech Republic
UBUNTUNET	The UbuntuNet Alliance for Research and Education Networking - Malawi
CLARA	Cooperación Latinoamericana de Redes Avanzadas - Uruguay
IHEP	Institute of High Energy Physics Chinese Academy of Sciences - China
ASREN	Arab States Research and Education Network - Jordan
SIGMA	Sigma Orionis – France
C-DAC	Centre for Development of Advanced Computing - India

Disclaimer

Revision Control

Issue	Date	Comment	Author
v01	29/05/2014	First version	Manuel Rodríguez Pascual, Rafael Mayo-García and Ognjen Prnjat. Survey co-design by Christos Kanellopoulos and Kostas Koumantaros. Survey dissemination and regional analysis by Salem Alagtash, Bruce Becker, Subrata Chattopadhyay, Gang Chen, Herbert Hoeger, Luis Nuñez and Sijin Qian
V02	01/06/2014	Major revisions and text contributions	Ognjen Prnjat
V03	02/06/2014	Minor revisions	Roberto Barbera
V04	04/06/2014	Final revision	Federico Ruggieri

Abstract

This deliverable reports on the results of the survey regarding the Research and Education (R&E) clouds in the CHAIN-REDS' target regions as well as in Europe. The survey includes the technical aspects of R&E clouds, as well as issues related to interoperation, interoperability, compatibility, orchestration and federation. The survey has been released in December 2013 and results collected until early February 2104. Taking into account the review recommendation to decrease the surveying efforts, the finalisation of this state-of-the-art analysis was done through direct consultation with regional representatives. Results obtained show a high interest in standards, thus pointing to the fact that the solutions offered by CHAIN-REDS are indeed relevant. Thus, the overall recommendation of this activity is to follow up with promotion of CHAIN-REDS-supported standards in these communities via technical round tables, and specifically the promotion of cloud federations, as specified in DoW of the project.

Table of contents

Purpose

CHAIN-REDS started on December 1st 2012 and aims at promoting and supporting technological and scientific collaboration across different e-Infrastructures established and operated in the various continents. It is a FP7 project co-funded by the European Commission (DG CONNECT) and has as ultimate goal to define a path towards a global e-Infrastructure ecosystem that will allow Virtual Research Communities (VRCs), research groups and even single researchers to access and efficiently use worldwide distributed resources (i.e., computing, storage, data, services, tools, applications).

To do so, the project is structured in several Work Packages that address these different scenarios. Specifically, WP3 "Interoperation and coordination of e-Infrastructures" is devoted to support the stability of existing and emerging Regional Operation Centres (ROCs) by maintaining a set of guidelines for standard grid interfacing customised for the type of region; but also to investigate the emerging cloud solutions and propose the relevant standards in this arena.

Within WP3, task 3.3 is devoted to "Clouds for Research and Education", and its objective is to study cloud compatibility, interoperation, orchestration and federation issues, and provide guidelines for compatibility, and promote interoperation and federability between different clouds initiatives via standards.

In order to understand the state of the art of clouds for R&E in the regions, a survey was conducted between December 2013 and February 2014. This deliverable contains an analysis of the results of the survey and reports on the relevant issues regarding interoperability, orchestration and federation. In conclusion, the core set of cloud standards described in the preceding deliverable D3.2 "Interoperability guidelines and design" and used as proof-of-concept cloud federation exercise, is relevant globally.

It is important to note that main target recipients of the survey were the cloud providers and developers. The only end-users expected to answer it were large institutes, universities, or research groups which use external cloud resources, either public or private. It was not meant for purely research groups experimenting with cloud as a concept, or single end-users.

Glossary

APEM SSM	Accounting Processor for Event Logs - Secure Stomp Messenger
API	Application Programming Interface
CA	Certification Authority
CDMI	Cloud Data Management Interface
CHAIN	Co-ordination and Harmonisation of Advanced e-Infrastructures
CHAIN-REDS	Co-ordination and Harmonisation of Advanced e-Infrastructures for Research Education Data Sharing
CMF	Cloud Management Framework
DR	Data Repository
EC	European Commission
EGEE	Enabling Grids for E-Science in Europe
EGI	European Grid Initiative
GOCDB	Grid Configuration Database
HPC	High Performance Computing
IaaS	Infrastructure as a Service
IGTF	Interoperable Global Trust Federation
LDAP	Lightweight Directory Access Protocol
NGI	National Grid Initiative
NREN	National Research and Education Network
OCCI	Open Cloud Computing Interface
PaaS	Platform as a Service
PKI	Public Key Infrastructure
R&E	Research & Education
RDF	Resource Description Framework
SaaS	Service as a Service
SAM	Service Availability Monitoring
SWIFT	OpenStack Object Storage
SLA	Service Level Agreement
VM	Virtual Machine
VRC	Virtual Research Community

1 Survey Design

The survey was designed with the objective of gathering information from all the CHAIN-REDS' target regions.

The first part of the survey focused to obtain a high-level vision of what the user considers a cloud infrastructure to be. Then, several questions about the status of cloud infrastructures for R&E were asked. Since one of the main objectives of this survey was to discover the public and private institutions running cloud infrastructures for R&E purposes, several questions about their location, quantity and quality of the resources were asked. Several questions were also asked to gather the technical and managerial contacts of the cloud infrastructures.

The core of the survey then consisted of a series of technical questions about the available resources and their management. For this sake, a group of questions focused on the end-user perspective, and the rest on the provider perspective. The questions for cloud providers covered quite many issues for this kind of infrastructure: economic, user management, middleware, size of the managed resources, accounting, inter-cloud orchestration/federation, security and training.

The survey has been distributed to the targeted recipients by different means. First, regional representatives have submitted it to the relevant players in their region, such as infrastructure administrators, project managers, etc. In Europe, it has also been submitted to several relevant distribution lists: Spanish National Supercomputing Network, OpenNebula users list, and EGI Cloud Task Force. In Latin America, it has been sent to RedCLARA's SCALAC, NRENs technical responsibles and core computing centres. In Africa, China and India, it has been sent both to commercial providers with implications in R&E and to public institutions. These institutions were cloud users and /or providers. In these regions the contact has also been direct.

2 Survey results

In this section the results of the survey are analysed.

As a first step of the survey analysis, the results have been clustered according to the different regions of interest for the CHAIN-REDS project, and then interpreted together with regional experts in their specific context.

After this analysis, the results have been presented as a whole, in order to provide a global vision of the status of Cloud Computing for R&E.

Here, both the survey results and the regional analyses are presented.

2.1 Regional analysis

2.1.1 Arab countries

The following section sheds some light on cloud development in the Arab region, based on survey results collected from a number of existing and potential cloud sites representing research and education infrastructures in United Arab Emirates, Egypt, Jordan, Saudi Arabia and Bahrain – these being the countries, to our knowledge, with some cloud computing developments.

2.1.1.1 Summary of cloud providers for R&E in the region

Recently, a growing number of cloud infrastructures have emerged in the Arab World. Cloud computing solutions are still considered as new technologies providing variety of public and private cloud services ranging from hosting applications and services to providing hardware infrastructures. The results of the survey have shown that cloud providers are specifically concentrated in three countries in the Arab Region; United Arab Emirates, Egypt, and Jordan, especially in the telecommunication sector, and this not being directly targeted to the R&E sector.

Telecom operators in these countries are emerging as the main providers of cloud computing and hosting services at enterprise level, covering both small to medium and large companies. This is the case, for example, of Etisalat in the United Arab Emirates (UAE) and Mobily in Saudi Arabia. Cloud services allow a variety of infrastructures to connect to different branches, locally and internationally. The emerging cloud models are: Infrastructure as a Service (IaaS), Platform as a Service (PaaS) and Software as a Service (SaaS). Operators also offer special cloud computing services such as virtual private servers, public clouds or cloud servers, with virtual CPUs, firewall and memory, data transfer or managed storage arrangements. Services also include email hosting solutions, space for information and data storage, co-location of servers, networks and storage equipment, and hosting of research applications.

In contrast, cloud solutions dedicated to R&E are still in their very early stages. The Cloud Computing Lab research initiative at Carnegie Mellon Qatar is an example of those. It aims to advance research and open the cloud infrastructure to local businesses and industries in order to support oil and gas exploration. Ankabut at UAE has begun the negotiation with major cloud providers to setup their own cloud for research and education. IT Synergy has also contributed to setting up cloud infrastructure to support research and education in Egypt. Furthermore, Talal Abu-Ghazaleh has deployed its own private cloud to support universities and research institutions.

2.1.1.2 Summary of cloud users in the region

Cloud computing services are gradually becoming highly regarded in the education and research sector, with few emerging as part of their NREN developments. However, privacy, security, and reliability concerns represent the main barrier to adopting clouds. Clients are concerned about the use and legacy of their data as well as its confidentiality. As an alternative, private clouds are becoming more popular with their model as host providers for internal use and limited external use. Users seem to be concerned with data and infrastructure security, as a result, not much information has been revealed in the survey.

2.1.1.3 Summary of the technical and management solutions in the region

It should be noted that responders to the survey did not show much interest in providing details about their infrastructures and solutions (Hypervisor, security solutions, and management and storage). It is intuitively evidenced that the majority of cloud providers in the region rely on VM ware, Microsoft, and Open Source virtualized infrastructures, but not much information can be collected on the percentages of use of these technologies. In conclusion, cloud services in the Arab region are still in the preliminary stages of implementation and will take some time before they are widely deployed and used in a variety of sectors, including research and education.

2.1.2 Africa

This section analyses the results of the survey regarding African cloud infrastructures, and combine them with the consortium's knowledge of the field, in order to provide an overview of the status of cloud infrastructures for research and education in Africa. Please note that the Maghreb Countries have been included in the "Arab Countries" section of the analysis.

In this survey, a significant difference has arisen between South Africa and the rest of the African countries. While in South Africa there is an ongoing process about cloud adoption for R&E, in the rest of countries there is no implementation of this technology at all.

2.1.2.1 Summary of cloud providers for R&E in the region

There are several cloud providers for R&E in South Africa.

CSIR Meraka Institute[1] supports cloud concepts from different points of view: SANREN[2] (the South African Research Network) is mainly devoted to storage and commodities: conference, EduROAM in a box and PaaS services; while SAGrid[3], the South African National Grid, is currently supporting a small number of virtualized resources, and is actively collaborating with CHAIN-REDS project towards an intercontinental cloud federation.

The National Research Foundation provides cloud resources on two of its institutes, the iThemba Laboratory for Accelerator-Based Sciences[4] and the South African Astronomical Observatory[5]. The iThemba Laboratory for Accelerator-Based Sciences, iThemba LABS, is currently providing a small cloud installation. In this case they have two resources, one limited to system administrators and oriented to commodities, and other open and devoted to R&E.

[1] http://www.meraka.org.za/
[2] http://www.sanren.ac.za/
[3] http://www.sagrid.ac.za/
[4] http://www.tlabs.ac.za/
[5] http://www.saao.ac.za/

North West University [6] is currently employing a cloud of about 20 VMs both for commodities and educational purposes.

The situation in the rest of Africa is not so optimistic: until now, to our knowledge, there is neither public nor private provider focused on research or education. There are, however, several private companies devoted to provide on demand resources: Kili[7] and Angani[8] are based in Kenya and CloudAfrica[9] in South Africa.

The following paragraphs explore the issues of cloud provider federations.

Although there is no cloud federation in South Africa, there is one multi-institutional initiative to provide cloud resources to university students and researchers. This is the South African UniCloud, based on Hyper-V, initially run by the University of Cape Town and the University of the Free State. This is a single-domain cloud across geographical and institutional boundaries serving a wide range of applications on demand. Examples of this are e-learning and data storage. The initiative was the result of an agreement between Microsoft Research and the University of Cape Town, through the Association of South African University Directors of IT (ASAUDIT). This is a commercial offering and as such out of the scope of the project.

On the other hand, the Meraka Institute is actively investigating participation to the EGI FedCloud activity, as a stepping stone towards national federation, which is foreseen in South Africa. The provided resources are currently very small, but this institution has successfully participated in the previous CHAIN-REDS cloud interoperability demo, proving that the initiative is viable.

The issue of interoperability between cloud infrastructures has been raised, at least in South Africa, due to the current incompatibility between open-standards-compliant stacks such as OpenNenbula and OpenStack, and the Microsoft Hyper-V implementation. Since there is no existing federation, this is not seen as a priority, but is being actively studied by CHAIN-REDS members in the region.

In other countries, CHAIN-REDS has found neither significant cloud providers locally nor initiatives of cloud federation for research or education. This could, however, change dramatically, as many research centres across the region (SAGrid sites in South Africa, ILRI and iHub in Kenya) are actively investigating the feasibility of virtualising the entire computing infrastructure.

Regarding the user base, no users for R&E besides the communities related to the cloud providers described in the previous section have been identified by the survey.

2.1.2.2 Summary of the technical and management solutions in the region

The technical solutions employed by the cloud providers can be split in two groups. The first one employs a software stack based on OpenNebula or OpenStack with KVM or Xen, providing support to Linux VMs. Some of them (NWU and iThemba) also make use of external resources provided by Amazon to satisfy their peak demands.

[6] http://www.nwu.ac.za/
[7] http://kili.io/
[8] https://angani.co/
[9] www.cloudafrica.net

Universities employing Microsoft resources via UniCloud constitute the second group. In this case the software stack is based on Hyper-V, and Microsoft Azure is used to satisfy peak demands.

Although not directly related to research and education, it is worth mentioning that the private companies contacted (Kili, Angani and CloudAfrica) also employ a software stack based on OpenStack and KVM.

The technical solutions employed by the described providers are what can be expected from mid-size clusters, with 10 to 100 VMs. It is however remarkable that SAGrid, although counting with a limited quantity of resources, provides a state-of-the-art administration. This is due to the previous experience acquired with the administration of their grid site, which has now been extrapolated to the cloud infrastructure.

The adoption of standards is wide, and it is worth mentioning that the willingness to federate is really high among the contacted institutions. Both SANREN and SAGrid (CSIR Meraka Institute) are willing to join the EGI Federated Cloud. Of those, SAGrid is already integrated into the CHAIN-REDS interoperability demo, so this task should be seamless.
It is important to remark that, although none of the private companies has integrated into an international federation, there is a positive attitude towards it. All of them have had conversations with the CHAIN-REDS regional delegate, and although they still have not agreed on federation, they have shown strong interest. Given that all are based on OpenStack, this federation should not be hard from a technological point of view. If carried out, this activity could only be an added value to the project, since the focus of CHAIN-REDS work is certainly not commercial cloud solutions.

As an important drawback, institutions employing the South African UniCloud do not support any kind of standard, and they are in fact comfortable with their current status and not willing to federate.

There is a strong interest on soft activities, both in terms of dissemination and formation (workshops). However, at this moment we are only aware of University of Cape Town providing training.

2.1.3 China

The section is the analysis summary of the survey regarding Chinese cloud computing state of the art.

2.1.3.1 Summary of cloud providers for R&E in the region

Cloud technology is still starting to develop in China, but its growth is really fast due to a generous budget for technology projects. Most of the answers of the survey come from sites with a very small number of resources (10 or less CPUs, a few TBs of storage), which seems to mean that these sites are mostly developed to test this technology and to employ it for commodities and small research, not yet for a full scientific development. There are however notable exceptions in the commercial world, such as AliBaba, with more than 100,000 CPUs and about 100 PB of storage making it the most powerful commercial player in the region.

The main cloud provider for R&E is the ScienceCloud, hosted by the Chinese Academy of Sciences (CAS). Starting in 2013, the infrastructure provides services of Data Cloud and Computing Cloud. Data Cloud consists of 1 major data centre with 2 PB of disks and 5PB of tapes in Beijing, 1 archive centre with 1 PB of disk and 10 PB of tapes at Huairou, and several mid-size data centres, equipped with disks, across the country. Computing Cloud is a cluster of computing facilities fully devoted to cloud services, with one

supercomputing centre in Beijing, 8 regional centres and some computing resources in institutes, with an aggregated computing power of about 200 TFlops. Resource federation has been established on several institutions with the creation of ScienceCloud and some more are looking towards this federation, so although this does not represent a big share right now, this number is expected to grow in the short to middle term.

CNGrid, the China National Grid, is a national HPC computing platform which counts with about 11 HPC centres in the country. CNGrid provides grid computing services as well as cloud computing services to science and industry communities. Some of its centres are starting to offer IaaS services.

The Institute of High Energy Physics, IHEP, has created a small cloud site devoted to run High Energy Physics applications: this site is envisaged to become part of the CHAIN-REDS testbed / global federation.

2.1.3.2 Summary of cloud users in the region

As in the rest of the regions, the usage of Cloud infrastructures in China from the user point of view indicates a mixed research/education demand, with a minor presence of commodities. On the other hand, cloud providers depict a major use for commodities, making it clear that this kind of usage is transparent to the users. Regarding scientific applications, the main areas supported by the providers include climate change, aerospace and aviation, drug discovery, ICT education, genomics and high energy physics.

In the case of ScienceCloud, applications for the following scientific domains are being deployed HEP; Microorganism; Global Change Ecology; Clean Coal Utilization; Astronomy; Stem Cell and Biomedicine; Plateau and Cold Region Environmental Monitoring; Space Science.

In the case of IHEP, their cloud infrastructure is being employed to run high energy physics applications and to test new technologies.

It is important to show that, according to this survey, none of the institutions employs external cloud resources. This may be due to the lack of specific founding, as none of the institutes has official or regular funding to cover the cost of cloud services. Instead, they tend to employ private clouds for R&E.

2.1.3.3 Summary of the technical and management solutions in the region

In China, the hypervisor market is divided between Xen and KVM. It is noteworthy that there is no presence of commercial alternatives, only open software ones.

Regarding the cloud management framework, OpenStack was the only mentioned. Although this is probably not representing the status of all Chinese institutions, it shows a clear market domination of this particular solution.

At this moment, cloud storage is provided by just about a third of the providers. This is probably due to the lack of founding and usage of private clouds, which partially avoid the need.

Regarding user authorization, authentication and support, we found what could be expected from middle size research and educational organizations. There are standard registration procedures - although their credential verification is not very strong in most of the sites - and identity federation can be employed in about half of the cases. There is an interest in allowing social web identification to manage cloud resources, but - probably due to internal policy - only Chinese social network Sina Weibo is currently allowed.

ScienceCloud is being created based on standards, although their implementation is not complete. In particular, OCCI is supported, but CDMI is still not available for data management, employing instead a set of proprietary RESTful APIs. There is a great interest in resource inter-federation and accounting, although APEL is not being considered. And to manage user identification, this project has created its own CA, which employs X.509 certificates to identify users along the different physical resources.

IHEP cloud is also heavily based on standards. In this case both OCCI and CDMI are accepted – as part of the CHAIN-REDS suite of standards. In fact, this cloud resource will soon be integrated in the CHAIN-REDS testbed, serving to demonstrate the benefits of standard adoption for federation and interoperability of resources between China and Europe.

All the providers will consider cloud workshops or awareness seminars of value and importance, and some of them are already providing it.

2.1.4 Europe

This section is devoted to analyse the results of the survey regarding European cloud infrastructures, and combine them with the consortium's knowledge of the field, in order to provide an overview of the status of cloud infrastructures in Europe.

The adoption of new technologies in the fields of research and education tends to be slightly slower than in industries. This may be due to the ownership of infrastructures: as cloud providers for R&E are typically state-managed institutions such as NRENs, it usually takes a bit longer to adopt new technologies and deploy of production-level data centres. Then, after the buzz in the last years regarding cloud computing, it is now getting an increased importance on R&E in the European region.

2.1.4.1 Summary of cloud providers for R&E in the region

Cloud Federations

One of the leading projects is the EGI-Inspire project, which has devoted a task group to the creation, support and maintenance of the EGI Federated Cloud. This project aims to build a grid of clouds, identifying itself as 'a seamless grid of academic private clouds and virtualised resources, built around open standards and focusing on the requirements of the scientific community'. Here, the federated operation services that EGEE and EGI projects have encouraged to build over the last years devoted to grid computing are now employed to interoperate cloud services. It is strongly based on standards, definitely acting as a cornerstone for its adoption in the region.

The EU-funded Helix Nebula project aims to 'pave the way for the development and exploitation of a Cloud Computing Infrastructure, initially based on the needs of European IT-intense scientific research organisations, while also allowing the inclusion of other stakeholders' needs'. These research organizations are CERN, EMBL and ESA, and are serving to validate the approach and to enable a cost-benefit analysis that can serve as a reference to future business and scientific research. After a successful proof of concept integrating resources from different providers with a standard-oriented software stack to solve scientific problems, it is ready to act as a resource provider for the aforementioned organizations and more to come.

CHAIN-REDS project includes a cloud orchestration integrated into its Science Gateway. Its architecture is depicted in deliverables 3.2 and 5.2. It is fully based on standards, and devoted to encourage the adoption of this technology through a practical point of view. Here, institutions from different regions (currently Europe, Africa and Arab region) share

their resources in a federated way, allowing to seamlessly access various e-Infrastructures in a way that is transparent for end users. Although it partially overlaps with EGI Federated Cloud, by sharing some of the members, their different objectives make both initiatives complementary.

Cloud Providers

Regarding EGI, 17 institutions from different countries have already joined the initiative and are already providing resources to research and education. Two more institutions are being integrated, and over 20 more have manifested their interest and are working on the integration with different degrees of success.

EGI-Inspire heavily relies on National Grid Initiatives (NGIs) as resource providers. This way, the creation of EGI Federated Cloud has led to national cloud initiatives, with individual resource providers federating their resources at a national level.

These national initiatives are providing resources to research, education and management, mainly to local universities and research centres, although users can be aligned with international projects. There are also small regional initiatives all around Europe arising as resource providers and consumers for local user communities, either of general purpose or very specific initiatives. The provided resources greatly vary in terms of CPUs and storage, from small test sites with 4 or 8 cores to GRNET's Okeanos, which currently hosts about 8.000 running VMs and several hundreds of TB of storage.

Within Helix Nebula, after a series of successful pilot with 15 resource providers and four research institutions as consumers, at this moment more than 35 public and private institutions have joined the consortium either as resource providers or consumers.

Five out of the seven cloud providers integrated in the CHAIN-REDS testbed / global cloud federation, belong to European institutions. Three of those are also affiliated to the EGI Federated Cloud, and two are not.

Beyond all these public providers, private ones are also employed on R&E. Among them, Amazon is by far the most used one. It is universally known and widely employed for commodities, CPU-intensive applications and storage.

2.1.4.2 Summary of cloud users in the region

The range of cloud users in the region is large, as is the range of possibilities that this new technology enables. There are anyway three separate tendencies: that is, education, commodities and CPU-intensive applications for research.

Many research and educational centres employ virtualized infrastructures for commodities. In this case, the IaaS model is massively adopted, benefiting from the advantages of virtualization (robustness, flexibility, security) for non CPU-intensive applications where the performance loss would have an impact. SaaS is also increasingly employed, especially for email. Institutions with this model tend to employ their own resources, although external cloud providers – and among them, Amazon - are sometimes employed to satisfy peak demands.

Cloud computing is also being adopted for educational purposes, both in educational institutions and research centres. This is especially visible in computer science colleges, where the employment of virtual machines with different configurations over a small set of physical resources is becoming the standard on practical courses. In this case, the resources come from the own institution, probably due to the low requirements in terms of CPU and data storage.

And last, cloud is employed for different research communities as a resource provider. Our survey shows a special interest by the bioinformatics community – DNA sequencing, data storage- and the physics one, although there are many more. As a reference, EGI Cloud Task Force is supporting nine communities from different areas, from language analysis to data visualization.

2.1.4.3 Technical and management solutions in the region

Regarding cloud management frameworks (CMFs), there is an overwhelming presence of free software solutions. Market is somehow divided between OpenNebula and OpenStack, with Synnefo, Emmotive and StratusLab having a minor presence. The apparent tendency is to employ OpenNebula in smaller sites and OpenStack in the bigger ones, although there is not a precise statistical result on this. OpenStack is definitely winning the battle over OpenNebula with research institutions, although OpenNebula has a wider presence in Spanish institutions, probably influenced by the fact that it is being developed by a Spanish University.

Again, Amazon is here the main resource provider for outsourced resources. It probably helps that both OpenNebula and OpenStack count with a full integration, so its adoption is straightforward.

Going down to the hypervisors, cloud provided in our survey shows a major presence of KVM, followed by Xen, with VMWare being employed on a small number of sites. As with CMFs, there is a strong tendency on the usage of free, Open Source solutions.

Most projects provide some kind of cloud storage. It is important to regard that there is a very strong interest on OCCI for public management of virtual machines and CDMI for data management, so we will probably see a nearly universal adoption in the short to middle term – which strongly supports the CHAIN-REDS choice of standards. At this moment, OCCI is supported by two thirds of the providers and CDMI by about half of them, but nearly 100% of the survey answers show an interest for its adoption.

About half of the resource providers surveyed employ some tools for accounting. APEL SSM 2.0 is still not fully extended, but it is definitely being considered. In fact, two survey answers talked about building a bridge from their local solution to APEL, what shows an interest on this technology.

The employment of monitoring tools is much bigger, as virtually all our survey responses for Europe show a positive answer in the area. This percentage falls to about 50% when coming to EGI SAM monitoring system, so there is still work to do on the area.

Regarding user authorization, authentication and support, we found what could be expected from middle size research and educational organizations. There are standard registration procedures -although their credential verification is not very strong in most of the sites-, and identity federation can be employed in most of the cases. Their adoption is not complete though, and more sophisticated tools such as group based policies have still minor presence.

It is important to point-out that, although the usage of the resources through social web (Facebook, Twitter, etc.) identities is still irrelevant, there have been many answers showing their interest in its adoption.

This survey shows that many of the resources are part of inter-cloud federations, and most of the ones that are not have shown interest on it. Here, the interest on standards (GLUE 2.0) is approximately the same that with CMDI, OCCI and the rest.

These computational resources are being managed with the usual security measures: the role of security officer is present and well defined; there are Acceptable User Policies, well documented security patching mechanisms and instruction detection.

Summing up, this technical section of the survey has clearly shown two factors: an ongoing but incomplete adoption of standards and opening measures, but a strong intention on doing it, at least for most the providers.

The last section of the survey was devoted to gather for interest in cloud training and awareness seminars of importance. The answers show that both users and providers are interested in training, as cloud is still a new technology and there is a significant user base to be taught. Heavily related to that, there is the need of disseminating the possibilities that cloud offers to the scientists, educators and system administrators.

There are however many initiatives being carried out, such as the dissemination efforts by CHAIN-REDS, devoted sessions in many computing-related meetings, conferences by OpenNebula and OpenStack and university courses in many computing science universities all around the continent.

2.1.5 India

This overview of the current status of Cloud infrastructures in the academic and research organizations in India was gathered over last few months through the cloud survey, personal interactions and web search. Since the area of research is growing and evolving rapidly, the current update is in no way comprehensive.

2.1.5.1 Summary of cloud providers for R&E in the region

Cloud Federations

Although there is no cloud federation in India, there are several multi-institutional initiatives worth mentioning and their existence taken into account, considering them as a helpful support on the area.

The Department of Electronics and IT (DeitY) in India has launched Meghraj (GI Cloud[10]), a national cloud initiative to help the government leverage cloud computing for effective delivery of e-services. The roadmap declares that cloud computing environments will be established at the National and State levels, starting with one National Cloud, using new and existing data centres.

The Cloud Computing Innovation Council would be driven as an open forum with subject matter experts participating as volunteers in working groups delivering innovation proposals and outcomes such as: Standards recommendations, interoperability testbeds, certifications & benchmarks; White papers and reference architectures; Pilot projects to develop prototypes; Research projects resulting in "eminent" publications and Intellectual Property; Policy and regulatory recommendations; Incubation of innovation promoting entities – Innovation Sandbox, Innovation Clusters etc.; Capability building through conferences, workshops, eLearning platforms, design challenges, awards, participation in global forums etc.

R&E clouds

CDAC has deployed primarily two cloud stacks, namely megadoot and sumegha. While the first one is for general purpose cloud environment, the Sumegha focus only on scientific applications.

[10] http://deity.gov.in/content/gi-cloud-initiative-meghraj

Baadal, from the Indian Institute of Technology (IIT) Delhi, is a Cloud orchestration and virtualization management software. It is being employed at IITD to manage 48 blade servers with about 500 cores and 50 TB of virtualized storage and currently has over 60 virtual machines commissioned for high performance computing.

eGovernment clouds
The Department of Electronics and IT (DeitY) in India has launched Meghraj (GI Cloud[11]), a national cloud initiative to help the government leverage cloud computing for effective delivery of e-services. The project, initiated to support the implementation of the National e-Governance Plan of India (NeGP), aims to create a private cloud environment for the use of government departments and agencies at the Centre and State levels that will accelerate the delivery of e-services to citizens and support other objectives such as increasing standardization, interoperability and integration, pooling of scarce and underused resources and the spread of best practices. The National Cloud will be able to provide services such as computing, storage and network infrastructure, backup and recovery and application development supported by State Clouds.

National Informatics Centre (NIC) cloud computing approach provides government partners with an opportunity to cost effectively manage technology resources which accelerates speed to market and delivers resources on a scalable as-needed basis. The NIC private cloud offers federal, state and local government partners a state-of-the-art, high performing and fully secure hosting operation that support secure transaction processing worth several billion dollars per year.

Private company clouds
One of the largest telecommunication providers in India, Bharat Sanchar Nigam Limited, provides various cloud service namely Co-location and Hosting services. It also provides compute as service and Managed services like standalone database, Clustering, Load balancing, and Application accelerated service.

In addition to this, a lot of small players both public and private organizations, are offering cloud services and consultancies all over India.

2.1.5.2 Summary of cloud users in the region

The range of cloud users in the region is large mostly in education and research applications. SaaS type of resources is utilized in managing ERP, email and eb hosting environment.

Cloud computing is also being adopted for educational purposes, both in educational institutions and research centres. Lot of engineering colleges from premier institutes like IISC and IITS where the employment of virtual machines with different configurations over a small closed set of physical resources is done, are starting to test clouds. They mostly build small private cloud using their own resource specially targeting their research activities.

2.1.5.3 Summary of technical and management solutions in the region

Regarding cloud management frameworks (CMFs), there is an overwhelming presence of free software solutions. Market is somehow divided between Nimbus, Eucalyptus and OpenStack having citrix, reseviors, etc. small presences. OpenStack has lot of backing from providers who are either on the verge of offering commercial service or actually does. Nimbus has its presence especially in Research organization as it is has more support for creating MPI based cluster.

[11] http://deity.gov.in/content/gi-cloud-initiative-meghraj

Going down to the hypervisors, a major presence of KVM as it come along with REDHAT as a native Hypervisor, followed by VMWare and Xen being deployed on a small number of sites. There is strong inclination to use only Free and Open source based solutions.

Most of the providers have their own cloud storage through self-managed data centres. Some cases like CDAC also provide storage on external resources such as CloudVault.

About half of the resource providers employ some customisation using open source tools or developed in house for accounting and monitoring. Regarding the user authorization, authentication and support most of the people use database based server authentication. Some of them use LDAP of PKI certificate although the numbers are very small. Their verification process is not strong and still based on human to human interactions through other communicating channels.

These computational resources are being managed with the usual security measures: the role of security officer is present and well defined, there are Acceptable User Policies, well documented security patching mechanisms and intrusion detection.

The last section of the survey was devoted to gather for interest in cloud training and awareness seminars of importance. The answers show that both users and providers are interested in training, as cloud is still a new technology and there is a significant user base to be taught. Heavily related to that, there is the need of disseminating the possibilities that cloud offers to the scientists, educators and system administrators.

2.1.6 Latin America

This section is devoted to the analysis of the results of the survey regarding Latin America (LA) cloud infrastructures.

The adoption of new technologies in the fields of research and education is mostly tedious in LA, mainly due to the lack of support from these same research & education institutions and local/national governments. After the buzz in the last years regarding cloud computing, cloud computing is still incipient in the LA region. Brazil however is a notable exception.

2.1.6.1 Summary of cloud providers for R&E in the region

Cloud Federations

There has been an early successful effort of Brazilian research teams to cooperate with the EGI Federated Cloud Task Force. This is the case of the EUBrazilOpenBio[12] that aims at deploying an e-Infrastructure of open access resources (data, tools and services) to support the biodiversity scientific community. This data e-Infrastructure will be the federation and integration of existing EU and Brazilian developed infrastructures and resources, namely through Catalogue of Life[13], D4Science-II[14], openModeller[15] and Venus-C[16], becoming one of the earliest use cases along with examples from biomedical research, linguistics, literature studies and physics. Now a new EU-Brazil project is pushing this effort further. In the EUBrazil Cloud Connect[17] project, applications and services will provide a framework for solutions for societal challenges in fields like health, biodiversity and climate change. The cooperation between Brazilian and European

12 http://www.eubrazilopenbio.eu
13 http://www.catalogueoflife.org/
14 http://www.d4science.eu/
15 http://openmodeller.sourceforge.net/
16 http://www.venus-c.eu/
17 http://www.eubrazilcloudconnect.eu/

centres, using data and computing services under a federated environment is essential to this project. On the Brazilian side there is a strong expertise in biodiversity modelling, like openModeller, and data integration/handling, such as speciesLink[18], both CRIA[19] projects, which are available through the EUBrazilOpenBio project.

The SCALAC [20] initiative, supported by RedCLARA, involves research and education institutions from Ibero-America aiming at providing advanced computing service for Latin America and the Caribbean. In this respect, UNAM is leading the work to set up grids of clouds to be accessed via the SCALAC Science Gateway. It is now being setup initially with resources from two institutions and after the test phase is completed, other institutions will add resources.

Cloud Providers

Among cloud providers in LA, we can mention:

- RADEI, in Costa Rica, is in a very early stage of a commodity virtual infrastructure with plans to expand it up to a point where it can provide a real cloud service to institutions in Costa Rica.
- UniAndes, Colombia, uses Microsoft cloud services with Amazon EC2/S3 technology.
- RedCLARA is exploring the possibility to offer some cloud services tied to university applications.

Several alternatives have been evaluated. So far there are no national working cloud initiatives providing resources to research and education in LA, however private ones are being used even though there are complains that they are expensive.

2.1.6.2 Summary of cloud users for the region

The range of cloud users in the region is limited and Clouds are mostly used for education purposes based on IaaS, PaaS and SaaS models. Most institutions/organisations are cloud consumers but there are also some cloud providers and in an e-Infrastructure building stage.

In Brazil, EUBrazilOpenBio is providing support to biology users. EuBrazil Cloud Connect is currently working on three use cases from different areas (health, biodiversity and climate change), aiming to establish a stable collaboration between Brazil and Europe in the provision of data, services and expertise.

2.1.6.3 Summary of technical and management solutions in the region

Regarding cloud management frameworks (CMFs), OpenNebula and OpenStack are used, being OpenStack the preferred middleware.

VMWare is the only hypervisor reported to be employed.

Storage is provided and there is support for Cloud Data Management Interface.

It is important to remark that Universidade Federal de Campina Grande (UFCG) is building FogBow, an opportunistic private cloud platform that will be integrated within EUBrazil Cloud Connect eInfrastructure. According to its specifications[21] it is fully open and based on standards (OCCI, CDMI).

[18] http://splink.cria.org.br/: a distributed information system that integrates primary data from biological collections

[19] http://www.cria.org.br/: Centro de Referência em Informação Ambiental (Reference Center on Environmental Information)

[20] https://comunidades.redclara.net/wiki/scalac

[21] http://www.fogbowcloud.org/architecture

Open Cloud Computing Interface standard is being used for public management of virtual machines, support APEL SSM 2.0 is being considered for inter-cloud accounting aggregation, and monitoring for cloud infrastructure are used. There is familiarity with GLUE 2.0 schema.

In particular, the EUBrazil Cloud Connect project as well as the work being developed at Universidade Federal de Campina Grande (UFCG) [22], fogbow [23]: the federated, opportunistic and green cloud, are adopting OCCI.

EuBrazilOpenBio supports the Open Access Movement, promoting the concept of openness for scientific research. This has repercussions all along the project and, among it, in a full support of standards for cloud implementation, operation and support.

In these institutions, user registration service, helpdesk, and Service Level Agreement/monitoring are in place. Also, there are user registration procedures, affiliation of the users to the academic institution is checked, x509 certificates for user access are utilized, rely on CAs, there is support for group based policies and it is being considered to dynamically publish information about cloud infrastructures so that services are discoverable within an Inter-Cloud Federation context.
In Latin America, mainly in the academic community, incorporating Cloud Services strategies is recent and it is expanding. Thus the dissemination, development and awareness activities are relevant to promote the use of these services.

Courses on the creation of Identity Federations and eduroam federations for Academic Networks have been provided. By making the decision makers aware of the advantages of identity and resource federation and inter-compatibility, they are expected to have an impact on the federation of cloud infrastructures as soon as these infrastructures are settled.

Finally, regarding other surveys: there is a high interest in knowing the different aspects of cloud usage, and several surveys have been lately conducted by different organizations. In particular, OpenStack and OpenNebula, the two main open source cloud orchestrators in the market, have been collecting information about their users and making it public. Although their objectives are not identical to the ones in this deliverable, we consider that the provided results can be of interest, helping to determine the quality of our results and where our user base is located comparing with one of the leading industry standards.

Results on OpenStack survey[24] among their users show that the Academic and Research community is the second biggest on cloud usage with a 10%, with Information Technology being the first with about 65%. The weight of this community, although not being determinant is big enough to take them into account in the decision making process. The size of the deployed clouds is slightly bigger in OpenStack survey than in ours. This is probably due to the novelty of the technology: as has been depicted, public institutions tend to be slightly slower in the adoption of new technologies, so we are now in an earlier development phase. Regarding hypervisors, KVM and Xen are by large the most used ones, just as in our survey. Summing up, the similarity among the results on both surveys help to validate them and confirm the extracted solutions.

OpenNebula development team C12G conducted one survey among their users on the second half of 2012[25]. Although it is a bit older, results are also interesting. First, as our

[22] http://www.ufcg.edu.br/
[23] http://fogbowcloud.org/
[24] http://www.openstack.org/blog/2013/11/openstack-user-survey-october-2013/
[25] http://c12g.com/resources/survey/

survey shows, implantation of OpenNebula in Europe is bigger that in the rest of regions. This is opposite to Open Stack, which counted with a bigger use base in the USA. These users employ Opennebula to manage mainly mid-size clusters, with 10 to 100 physical nodes, with small presence in the <10 nodes, and significant ones in 100 to 500 and >500 nodes clusters. Again, this number is slightly bigger than in our survey, but consistent with it. Regarding hypervisors, it is important to remark that the presence of VMWare is much higher than in our survey and in OpenStack ones, while KVM is smaller. This is probably due to the evolution of the hypervisor market in the last year, showing a higher presence of Open Source solutions.

Overall, the results of this survey are consistent with ours.

2.2 Global Analysis of survey results

2.2.1 Summary of cloud providers for R&E

Cloud Providers

In the fields of research and education, the presence in Europe is most noted: currently, there are more than 20 participants in the EGI Federated Cloud, more than 35 (including private companies) in Helix-Nebula, and 4 European providers in CHAIN-REDS global cloud test-bed. In China, CAS, IHEP and CNGrid are providing cloud computation and storage. In India, several institutions are deploying clouds for R&E, the most important being CDAC. In South Africa, apart from the commercial presence of Microsoft through the UniCloud service, some institutions are slowly building cloud implantations. The rest of Africa and Latin America (apart from Brazil) are the ones with a smaller implantation, although different initiatives are appearing and will hopefully become providers in the mid-term. In the Arab Countries the implantation is also being slow, although Cloud computing services are gradually becoming highly regarded in the education and research sectors.

eGovernment and commercial clouds were not the targets of this survey, however some results were collected. In the case of eGovernment, India and Europe have performed significant movements, with public providers offering tools for this purpose. There is however no concrete evidence of clouds for eGovernment in the rest of the regions from the survey. The presence of private companies helps to serve as a reference of the interest on the technology. As it has been detailed in the previous sections, all regions count with telecom operators starting to provide resources on demand, being mostly employed for web hosting, application development and commodities.

The answers to the survey can be clustered in three groups according to the number of resources. About a 40% of them have very humble infrastructures, with less than 10 CPUs, not much storage and 2-4 GB memory. These are probably used for testing purposes, or being employed in production on small environments. Then, a second group of about 40-50% has a wider set of resources, with 10 to 100 CPUS, and more storage and memory. These surely correspond to production environments, where some of the commodities have been migrated to the cloud and resources are being employed by scientists or educators too. The last group of providers, about 10%, are heavy ones. They have environments with thousands of CPUs, storage on different means (disk, tapes) and a lot of memory. These are possibly cloud providers for third parties or administrators of large institutions where a significant part of their infrastructure has been migrated.

Besides questions about their own infrastructures, the survey recipients were asked about other cloud initiatives they are aware of. Results show that, beyond non-commercial R&E clouds, there is a growing interest in the area both by the governments and private companies in all regions. Annex 1 includes an extensive list of initiatives, combining the previously known ones and those identified by this survey, for future reference.

Cloud Federations

Following the activities on Grid computing, there is an increasing effort towards the federation/orchestration of cloud infrastructures. There are however significant differences between the various continents. The European efforts are evolving towards a stable cloud federation, while in Latin America - namely Brazil – same efforts in similar direction are made. In the case of India and China, country-wide institutions (namely CAS and CDAC) are deploying distributed resources, allowing the researchers to access them with standard procedures. Although they are not strictly federations, they surely are steps in the correct direction.

2.2.2 Summary of cloud user aspects

The survey showed three different usages of cloud infrastructures for R&E, each representing roughly a third of the users: education, research and commodities.

A group of questions was specifically devoted to determine the friendliness of the solution from the user's point of view. They have indicated that a high knowledge on IT is required, but when you have it, dealing with clouds is not particularly complex, and the existing tools and interfaces do a fairly good job. However, it would be desirable to reduce the need for this technical knowledge.

National data privacy and protection laws have been considered by more than half of the users. A strong relationship between this parameter and the externalization of services is noticeable: the vast majority of the users of external providers has considered data management issues, while people employing in-house solutions do not put so much effort on this, probably because of their trust on their system administrators.

It is important to mention that the employment of outsourced public clouds is still reduced, being usually employed to satisfy peak demands and not very often. That may be probably due to the limited budget on most institutions for cloud infrastructures, and to the fact that both Education and Research - that represent 2/3 of the whole usage - are usually predictable and flexible on their computational demands.

The survey has shown the areas of interest in cloud infrastructures for research. Results show that main interest relies on the areas of bioinformatics, physics, chemistry and engineering. This is not surprising, as those communities are heavy users of distributed infrastructures and have previous experience on Grid computing.

2.2.3 Summary of technical solutions

Next two questions are among the most interesting ones in the survey, as they provide a true insight of the users. The users were asked, both as cloud producers and consumers, about the problems when dealing with cloud infrastructures. Responses have been abundant and sincere, and truly point at the real deals.

The economic problems are identical in both groups. Funding is of course the main issue, both to hire resources and to pay for the hardware and energy bills. Producers are worried about the sustainability of the cloud model, probably because of the large investment that they have to make in order to be competitive. A particular issue arises: while producers complain about the lack of long-term commitment of users, they complain about the need for long-term contracts.

Other common problems are the ones related to hardware. While there are some specific to the groups (location and maintenance for providers, opacity for users) the most important one is the network. Of course, it worries both groups: the providers must guarantee a high performance network and fluent communication with the servers, while the user is afraid of that not happening, and how to achieve a fair QoS.

When dealing with software, the differences arise. Providers complain about the lack of a clear defined suit of technologies related to building the cloud, fast changing standards, task automation, software deployment and accounting. Users are worried about privacy, security, confidentiality, data movement and vendor lock-in.

These problems are somehow related to standards. There is the need of a clear, universal and well known stack of standards, so the different pieces composing the cloud architecture are interchangeable and stop being a problem. Although these standards do exist, there is still a lack of information that should be addressed in order to speed up the

adoption of the cloud paradigm and dismiss both users and providers fears and technological problems.

Last but not least, there were a set of problems related to the human behaviour. Providers pointed at the lack of qualified manpower to build and operate the infrastructures; migration of legacy environments to cloud ones was also –and correctly- pointed as problematic, although this problem will probably be dismissed in the mid-term, as the oldest and more problematic infrastructures are migrated; identity management was seen as a big deal, directly hitting one of the issues that CHAIN-REDS deals with and proposes solutions to, via its identity federation work; and documentation both for users and administrators was considered a problem too. Also, reticence to change was spotted as a problem by providers too.

Regarding the employed hypervisors, the huge majority of the answers pointed towards Open Source solutions, namely KVM and Xen. A small number of VMWare and hyperV users are also present, showing that there is still room for different paradigms and competitors on the field.

2.2.4 Summary of management and operational solutions

Next group of questions were related to standards, and answers clearly reflect two tendencies among providers. About two thirds of them are familiar with OCCI protocol, consider it important and support it in their infrastructures – or are in the process of supporting it. The other third is not really interested in any kind of cloud federation, do not support OCCI and in principle won't do it although all the possible issues are solved. These correspond to private clouds with no interest in scaling or collaborating, and to some public providers in China and India regions which are partially using their own implementations to handle these issues.

The result of Data Management section shows that about half of the providers that answered the survey provide cloud storage, most of them including CDMI or SWIFT. Other configurations include ownCloud or proprietary systems.

Moving to the operation section, nearly a 90% of the cloud providers employ some kind of monitoring, with about half of them having some kind of accounting mechanism. Regarding the monitoring, the vast majority uses standard existing tools (nagios, Ganglia, OpenNebula monitoring, etc.) although, in the case of accounting, most solutions come from in-house developments.

Regarding standards, most providers are open to the inter-cloud accounting aggregation via APEL SSM 2.0. There is a strong correlation among this parameter and the percentage of accounting usage. There should however be investigated why the rest of providers are reticent to this.

Only about a third of the providers include a SLA monitoring tool. This probably has to do with the previously proposed usage of the infrastructures, where a 10% of the survey answers were by heavy providers and 30-40% mid-size ones. In this scenario, keeping a SLA is probably out of the scope of most providers.

Authentication and authorization is probably one of the most important sections of the survey. As it has already been stated, dealing with authentication and authorization has been detected as a problem both by cloud users and providers. Thus, knowing what and how the providers are doing is a key for finding a suitable solution.

Three quarters of the providers – probably all but the small test installations - use a user registration procedure. In this registration they ask for different authentication ways,

depending on the service provider and desired level of security. Standard ones are the name, mail, affiliation and some kind of ID, together with payment information – in case of commercial providers.

At this moment, only about a quarter of the providers are providing some kind of identity federation such as national initiatives or eduGAIN. At last, social web identities (Facebook, Twitter, etc.) are only supported on 20% of the sites. It is important to remark here that several sites reported interest on this or are already testing it, so there is plenty of space to work on this area.

X.509 certificates are supported by half of the providers. Most of them do not implement their own certification authority. Instead, they rely on IGTF CAs or commercial ones. Again, although these stats show a positive tendency, the project should encourage the adoption of these standard certificates and CAs.

The section of Cloud Federation and Information Discovery shows some interesting results. The first noticeable thing is the very positive expectations of the cloud providers towards an inter-cloud federation: Nearly 75% consider publishing their infrastructure information in an inter-cloud context. Slightly over half the providers are already part of some kind of inter-cloud federation, and an additional third in the process of federating or very interested in doing it. Only a 10% of the providers are not interested at all in this possibility. The reasons that they claim for this federation cover pretty much all the well-known advantages of the process: performance, reliability, scalability. Surprisingly, virtually all the federated providers distribute the working load manually among the different infrastructures, with no one using a devoted commercial or open software tool. CHAIN-REDS publishes widely its information on standards adopted and the federation process and is thus indirectly stimulating this group of providers.

A similar situation occurs when dealing with GLUE 2.0 and GOCDB. About half the providers know these technologies or are already using it, an additional quarter is interested in them, and the remaining quarter is not interested or does not know the technology. A greater dissemination effort must be made by the affected parts in order to fully reach the community and explain the advantages that these standards provide.

2.2.5 Summary of soft activities

The survey finished with a glance on the soft actions being and to be performed.

The first question showed that only half of the organizations are providing some kind of cloud training. From CHAIN-REDS point of view this is a surprisingly low quantity, especially given that the technical complexity was raised as one of the main problems of these infrastructures.

The second one demonstrated that there is a high interest (85%) in making people aware of the cloud capabilities, both by workshops or awareness seminars. Among the actions being taken or suggested, a special interest was in taking contact with the educational communities by university workshop, academic clouds or portfolios of applications.

3 Conclusions and guidelines

In this survey, a thorough analysis of the cloud status for R&E at a global level has been performed.

The results show that, even though there is a very strong interest in cloud computing by the scientific and educational communities, this technology has still not been fully adopted. Based on the resource size and status of the infrastructures, it seems that the infrastructure providers are moving from testbeds to small production environments, slowly solving the administrative and technological issues that the adoption of this new paradigm is arising.

The users were asked, both as cloud producers and consumers, about the problems when dealing with cloud infrastructures. Responses have been abundant and sincere, and truly point at the real issues. This conclusion section thus outlines CHAIN-REDS added value regarding these issues and guidelines for the way forward.

Standards: Most of the fears and issues that keep both users and provides away from adoption and use of cloud infrastructures can be solved by a correct deployment of the technology and adoption of standards.

The objective of Task 3.2 of CHAIN-REDS is 'Promoting and fostering the adoption of standards in R&E eInfrastructures', so we consider that in this respect the current project activity is highly relevant. If the issues worrying the users are corrected through a wise selection of components implementing standards, it will foster cloud adoption, interoperability and cooperation among regions. For this sake, it is important to keep disseminating the importance of standards, together with the creation of a set of guidelines for standards adoption, to orient the infrastructure providers. CHAIN-REDS has provided a set of recommendations for cloud standard usage in Deliverable D3.1. These include:

- If a region/organisation plans to create a public cloud for research and education and they want to share its computing and storage resources with other clouds available in other parts of the world, for the benefit of the VRCs they support, it is suggested to choose a cloud middleware already supporting the CDMI and OCCI standards;

- If a region/organisation already owns and operates a public cloud for research and education and chosen middleware is not compliant with CDMI and OCCI standards, it is suggested to create the needed services and endpoints to support those standards.

Specifically related to the first point, technical guidelines about how to configure sites in order to be part of the CHAIN-REDS cloud testbed, which is interoperable with the EGI Federated Cloud, are provided in Annexes I, II and III of Deliverable D3.2.

CHAIN-REDS will keep these technical guidelines publicly accessible, while it will also proactively promote their use through cloud technical round tables, which already took place related to regions of China and Europe, with upcoming technical round table targeted at the Arab region, also with participation of Indian experts.

As the last point regarding standards, it is important to put a special attention on not duplicating efforts, both inside CHAIN-REDS project and among it and other EU-funded ones. Examining the worries of users (privacy, security, confidentiality, data movement and vendor lock-in) and providers (user identification, monitoring, accounting), the work

performed within EGI Cloud Task Force is of high relevance for CHAIN-REDS, both in terms of documentation and interoperability results – both effectively adopting a same set of standards.

Global interoperable cloud test-bed: it is important to deploy standards-based, proof-of-principle intercontinental cloud federations, to be able to demonstrate the feasibility of such cloud federations on a global scale.

CHAIN-REDS will continue its efforts on building an intercontinental interoperable cloud, fully based on standards, as proof-of-principle of cloud orchestration and federation. Currently, the CHAIN-REDS test-bed consists of a number of European sites, an Egyptian site from the Arab region, and, following up on survey results one site from South Africa has been also successfully integrated after being provided with CHAIN-REDS standards and guidelines. Efforts will be invested in further dissemination of CHAIN-REDS adopted standards, targeting primarily China, India and Latin America, to ensure proof-of-principle global coverage. These sites pinpointed by the survey and the related analysis presented here, and not yet fully members of the demo will be contacted on technical and managerial levels, so collaboration can be established and some resources potentially integrated. This way, these flagship institutions will serve as an entry point in their communities.

Ease of access (authentication and authorisation): this is one of the crucial factors and as such should be a priority in stand-alone as well as federated cloud environments.

Regarding this aspect, CHAIN-REDS approach to Identity-Federation based access is directly relevant, with eduroam being adopted both on the European level, while the rest of the world regions involved in CHAIN-REDS are being exposed these concepts by the project. This approach should continue being promoted: at the moment only about a quarter of the providers are providing some kind of identity federation such as national initiatives or eduGAIN. On the other hand, X509 certificates are supported by half of the providers; most of them do not implement their own Certification Authority and instead rely on IGTF CAs or commercial ones. Again, although these stats show a positive trend, the project should also encourage the adoption of this standard certificates and CAs.

Documentation and training: documentation and training on cloud computing is of high importance to ensure the take-up of the technology and overcome the adoption barrier.

The problems related to user and administrator formation, as well as the technology dissemination, are necessary to be tackled with. For this sake, valuable documentation has to be created or located, and offered on a single entry point –such as the project web- so it can be employed at different levels: administration reference, academic formation and so. Members of CHAIN-REDS project, fully aware of this problem, are working together with the EGI Federated Cloud on creating documentation regarding these scenarios, so it can be used as a reference in both projects. Also, as stated, one of the main uses of the cloud infrastructures is the academic one. This has to be surely seen as an opportunity to promote the use of standard technologies and identification. System administrators and decision takers should be made aware of importance of the use of standard technologies as a way of offering the widest possible set of resources with the minimum effort. Moreover, making the pupils aware of these standards –especially in computing courses- maximizes the possibility of employing them when creating their own cloud infrastructures or obliging their providers to have them. To contribute on this process, CHAIN-REDS will provide information about the outcomes of the ASREN Cloud event[26] (both tutorial itself from Jun 24th to 26th and EUMEDCONNECT workshop on 27th where a CHAIN-REDS session will be allocated) in future WP3 reports.

[26] http://asrenorg.net/article/25553/Cloud-Computing

Overall, if standards are promoted, proof-of-principle deployments created, security and ease of access enhanced, identification simplified and documentation improved, users will stop seeing the cloud as a problem and start seeing it as a solution.

Certain considerations must however be taken into account. First, it is important to consider that Cloud is not Grid. Even in scientific environments, a lot of usage comes from non-cpu intensive applications, such as commodities. This won't easily lead to a federation such as the EGI or WLCG. In this heavily commodity-oriented environment, private clouds make more sense (except on peak demands, and in this case, commercial providers can also be used and provide interesting capabilities, among them, lack of need to share your resources). Clouds are extremely useful for a set of purposes, but trying to make them a universal solution is not the silver bullet. The niche for this technology has to be clearly defined and limited, so time or economical efforts are not wasted by the users, scientists, administrators or course teachers.

4 Annex 1

4.1 Survey results

The following section presents the detailed results from the survey. They have been partially processed in order to simplify their readability.

Please regard that only responses that are self-interpretable (i.e., have no dependences with previous or next ones) are included here. Complete survey answers have been interpreted in the regional analysis with the core body of the deliverable.

4.1.1 Knowledge and Adoption of Cloud Technologies

What do you consider as a cloud? Please provide us with your definition.
Responses to this question are in general correct, although having different technical levels.

Which of the following cloud-related technologies do you use? IaaS, PaaS, Saas, Amazon EC2/S3, EGI federated cloud, Hypervisor, Cloud Resource Manager, OCCI, CDMI, Identity Federation, etc.

Most people were familiar with IaaS, PaaS and SaaS. Amazon technologies were also quite extended. EGI Federated Cloud was slightly less known but also popular. Some people also referred to technologies such as Helix Nebula, standards like OCCI and CDMI, and some more.

Are you aware of any de facto standards for cloud infrastructure, resources, and services?	
No	45%
Yes (please comment)	55%

Which of the following cloud-related technologies do you use? IaaS, PaaS, Saas, Amazon EC2/S3, EGI federated cloud, Hypervisor, Cloud Resource Manager, OCCI, CDMI, Identity Federation, etc.

IaaS, PaaS, SaaS and amazon services arose as the most employed services by far. Identity federations were also popular.

Regarding clouds, are you a resource consumer, a resource provider, or both? Please note that for this survey, "Resource Consumer" means that you are either paying for external cloud resources (Amazon, ESB, etc) or are using another R&E cloud for the purposes of your institution.	
Cloud Consumer	25%
Cloud Provider	30%
Both	40%

Are you aware of any government support to cloud initiatives in your country?	
Yes	46%
No	54%

Are you aware of any user communities in your region using cloud resources?	
Yes	58%
No	42%

4.1.2 Your details as a cloud consumer

Which of the following represent the use of your cloud infrastructure?	
Educational	32%
Commodities	18%
Scientific Applications	35%
Other	15%

Please provide an estimate of the monthly cost in USD
552 ± 469

Please indicate the user-friendliness of the solution from 1-5, 5 being most friendly	
1	4%
2	4%
3	35%
4	40%
5	17%

Please indicate the level of IT know-how required by staff using these external resources, 5 being highest	
1	9%
2	17%
3	35%
4	23%
5	17%

Have you considered national data privacy and protection laws?	
Yes	60%
No	40%

Is the external provider employed on a regular basis?	
Yes	35%
No	65%

4.1.3 Your details as a cloud provider

4.1.3.1 General

Do you use your cloud infrastructure for	
Educational Purposes	40%
Commodities (Web servers...)	33%

For the resources you currently have internally – please indicate the % of the cost associated	
OPEX cost: manpower, electricity, networking, consumables (indicate total OPEX and sub-categories also if possible)	40% ± 20%
CAPEX cost: (compute hardware, data hardware, data centre, network infrastructure, software licenses)	54 ± 26%

Sometimes it can be difficult to build from scratch, operate and employ a cloud infrastructure. For your local resources and infrastructure, what would you describe – in order of magnitude – as your three main barriers to entry: [For you as a resource provider]
Economic issues: - Funding - Energy consumption - Hardware budget - Sustainable business forecast
Hardware issues: - Network capacity - Resource allocation and scheduling at Cloud data centers
Software issues: - A clear defined suit of technologies related to building the cloud - Application software deployment - Accounting - Fast changing standards - Change of software management stacks - Security in terms of both authentication and authorization - Scalable, flexible federation of distributed Cloud resources - Automation of cloud management and integration as part of the normal service delivery process.
Human issues: - Lack of long term commitment of customers - Resistance to change - Creating confidence, awareness to user community - Migration from legacy environment to cloud - Identity management - Documentation for end-users - Personnel skills on Clouds

The employment of externally purchased cloud-based resources or services can be difficult or even impossible for different reasons (economical, technical, managerial…). For cloud-based resources and services, what would you describe – in order of magnitude – as the three main inhibitors/barriers to entry for [You as a customer]
Economic issues: - Budget cap - Cost-effectiveness - Long term contracts
Hardware issues: - Network performance - QoS - Transparency in resource usage
Software issues: - Privacy, Confidentiality, Trustworthiness - Security - Data movement - Vendor lock-in
Human issues: - Lack of flexibility required for education and learning environment - user friendly interfaces and access for resources

4.1.3.2 *VM Resource provisioning and management*

Do you find issues of technologies for federation, such as OCCI, as important?	
Yes	69%
No	31%

Would you be ready to federate your cloud into a larger international cloud, taking into account that security and reliability issues are ok?	
Yes	65%
No	35%

Do you support Cloud Data Management Interface (CDMI, SWIFT)?	
Yes	41%
No	59%

Does your cloud implementation support Open Cloud Computing Interface standard (OCCI, http://occi-wg.org/) for public management of virtual machines?	
Yes	65%
No	35%

4.1.3.3 *Data management*

Do you provide any cloud storage?	
Yes	47%
No	53%

# of CPU (Sockets and Cores) per host:
45% of the sites have 1 to 10 CPUs 45% of the sites have 10 to 100 CPUs 10% of the sites have over 1,000 CPUs
Online Storage size (disks, Terabytes)
40% of the sites have less than 5 TB. 25% of the sites have 5 to 50 TB. 25% of the sites have 50 to 500 TB 10% of the sites have over 5,000 TB
Nearline storage space (Tape systems, Terabytes)
45% of the sites don't have any 25% of the sites have less than 10 TB 7% of the sites have 10 to 100 TB 7% of the sites have 100 to 1,000 TB 15% of the sites have over 1,000 TB
Memory per CPU (Gigabytes)
40% of the sites have 4 GB or less 15% of the sites have 8 or 16 GB 15% of the sites have 32 or 64 23% of the sites have 128 GB 7% of the sites have 512 GB

Does your organization run a user registration service?	
Yes	50%
No	50%

Is there a distinct data interface/protocol available and if so what is it?
ownCloud
Proprietary
RestFUL API and SDK

4.1.3.4 Operations

Do you employ any accounting mechanism for CPU usage, VMs, storage, network or applications?	
Yes (please comment)	56%
No	44%

Would you be considering to support APEL SSM 2.0 (https://wiki.egi.eu/wiki/Fedcloud-tf:WorkGroups:Scenario4) for inter-cloud accounting aggregation?	
Yes	57%
No	43%

Do you employ any monitoring tool for your cloud infrastructure - for CPU usage, VMs, storage, network or applications?	
Yes (please comment)	88 %
No	12 %

Does your organization as a cloud provider have a Service Level Agreement/ monitoring in place?	
Yes	37%
No	63%

4.1.3.5 Authentication and authorization

Do you have a user registration procedure?	
Yes	75%
No	25%

Are you formally checking affiliation of the users to the academic institution?	
Yes	73%
No	27%

Are you providing your cloud as a service in some of Identity Federation (national federation or EduGAIN)?	
Yes (please comment)	25%
No	75%

Can users use their federated identity to access your resources?	
Yes	50%
No	50%

Can the users use their identities for the open social web (e.g. Google, Facebook, Twitter, Github, etc) to access your resources?	
Yes	20%
No	80%

Do you use x509 certificates for user Access? (Either Access to web interfaces or API Access)	
Yes	50%
No	50%

Do you rely on IGTF CAs, have you rolled out your own CA or you are using commercial CAs?	
Yes	35%
No	65%

Do you support Group based policies on your infrastructure? (e.g. users X1 & X2 belong to group X which has a specific quota on the available resources)	
Yes	47%
No	53%

4.1.3.6 Cloud Federation and Information Discovery

Would you be considering to dynamically publish information about your cloud infrastructure so that your services are discoverable within an Inter-cloud Federation context?	
Yes	73%
No	27%

Is your cloud infrastructure part of an Inter-cloud Federation? If not, would you be considering federating your cloud resources, within your country borders but also beyond?	
Yes	54%
Not yet, but I'm interested	36%
No	10%
If so, is the federation established among commercial providers, private resources, or a mix of both?	
Both	
Private	
Between other research institutions	

What would be your objective in this context (reliability, performance, economic matters…)?
Performance, reliability, scalability
Research
Software development

Are you familiar with the GLUE 2.0 schema?	
Yes	50%
No	50%

Are you familiar with the GOCDB (http://goc.egi.eu/)? Would you be considering registering you cloud infrastructure in the GOCDB?	
Already registered	15%
Yes	46%
Maybe	15%
No	24%

4.1.3.7 Security

Is there a person/team with the role of the Security/Officer in your Infrastructure?	
Yes	83%
No	17%

Do you have an AUP (Acceptable use policy)?	
Yes	90%
No	10%

Do you have a procedure for applying Security Patching? Is this documented?	
Yes	44%
No	56%

Do you have a mechanism for Intrusion Detection?	
Yes	67%
No	33%

4.1.3.8 Soft Actions

Does your organization receive or provide cloud training?	
Yes	52%
No	48%

Would you consider cloud workshops and or awareness seminars of value and importance?	
Yes	85%
No	15%

5 Annex 2

5.1 Cloud initiatives and infrastructures for R&E

Besides questions about their own infrastructures, the survey recipients were asked about other cloud initiatives they are aware of. Results show that, beyond non-commercial R&E clouds, there is a growing interest in the area both by the governments and private companies in all regions. Annex 1 includes an extensive list of initiatives, combining the previously known ones and the identified on this survey, for future reference.

Europe			
Inititative	**Services**	**Funding/Access**	**Country**
EGI federated cloud	IaaS	public access for EGI members	EGI federated cloud
Fundacion COMPUTAEX	IaaS	Research contract	Spain
GARR	IaaS, Storage	Services accessible to GARR users	Italy
GWDG Compute and Storage Clouds	IaaS, Storage	Restricted to Max-Plank-Society and University Goettingen	Germany
INFN Cloud infrastructure.	IaaS	Public	Italy
Meghacloud project - (RedIRIS	IaaS	Academic & research institutions affiliated to RedIRIS	Spain
Okeanos, powered by GRNET.	IaaS		Greece
University of Thessaloniki.			Greece
Marche Cloud.			Italy
PRISMA			Italy
Open City Platform			Italy

India			
Inititative	**Services**	**Funding/Access**	**Country**
Meghdoot	IaaS, PaaS		India
Sumegha, powered by Centre for Development of Advanced Computing C-DAC	IaaS, PaaS	Private. Funded by Centre for Development of Advanced Computing C-DAC	India
PanCloud/Scientific Cloud CDAC.	IaaS	Private. Funded by Centre for Development of Advanced Computing C-DAC	India
Meghraj.	IaaS,SaaS	Private. Public funds	India
National Knowledge Network (NKN). Private.	IaaS	Private. Public funds	India
Private companies			

Baadal Sanchar Nigam Limited (BSNL) Cloud	IaaS, PaaS, SaaS	Public. Private funds	India
IIHT Cloud Solutions Pvt. Ltd.	IaaS, PaaS	Public. Private funds	India
TCS iON	SaaS	Public. Private funds	India

China			
Inititative	**Services**	**Funding/Access**	**Country**
Distributed cloud at IHEP	Used by BESIII experiment	Private. Public funds	China
CNGRid	IaaS	Private. Public funds	China
ScienceCloud	IaaS	Private. Funded by Chinese Academy of Sciences	China
Private companies			
Alibaba Cloud Computing	IaaS and PaaS	Public. Private funds	China
Baidu Cloud	IaaS and SaaS	Public. Private funds	China
Tencent Cloud	IaaS, PaaS and SaaS	Public. Private funds	China

Latin America			
Inititative	**Services**	**Funding/Access**	**Country**
Cloud Academic. University of Costa Rica	IaaS, PaaS and SaaS	Private	Costa Rica
Cloud storage Provided by CNIC.	Storage	Accessible by the research institutes of CAS	Chile
The ELCIRA Project	Private	Private	
Tienda Cloud			Colombia

Africa			
Inititative	**Services**	**Funding/Access**	**Country**
iThemba LABS	IaaS	Private., Ffunded by National Research Foundation	South Africa
SAGrid	IaaS	Private cloud. Funded by CSIR Meraka Institute	South Africa
SANREN	IaaS, PaaS	Private cloud. Funded by CSIR Meraka Institute	South Africa
South African Astronomical Observatory	IaaS	Private. Funded by National Research Foundation	South Africa
Unicloud	IaaS, PaaS	Private cloud. Founded by Microsoft	South Africa
North West University	IaaS	Private. Public funds	South Africa

Private Companies			
Kili	IaaS	Public. Private funds	Kenya
Angani	IaaS	Public. Private funds	Kenya
CloudAfrica	IaaS	Public. Private funds	South Africa

Arab Countries			
Inititative	**Services**	**Funding/Access**	**Country**
Electronics Research Institute, Giza and Information Research Institute at Borg El Arab, Alexandria.		Private, public funds	Egypt
Talal Abu-Ghazaleh	IaaS, SaaS	Private. Funded by ASREN	
Arowaad Group	SaaS ERP for Schools	Private	Saudi Arabia
Cloud Computing Lab	IaaS	Private. Funded by Carnegie Mellon University	Qatar
Private Companies			
Orange Jordan	IaaS, PaaS, SaaS	Public	Jordan
Umnia Jordan	IaaS, PaaS, SaaS.	Public	Jordan
TAG Cloud	IaaS, PaaS, SaaS.	Public	Jordan
Etisalat	IaaS, PaaS, SaaS	Public	United Arab Emirates
Mobily	IaaS, PaaS, SaaS	Public	Saudi Arabia

5.2 Contacts

The following contacts have been identified in the survey and have been added to CHAIN-REDS contact list.

Technical Contacts	
INSITUTION	**NAME**
Alibaba Cloud Computing	Alibaba Cloud Computing
Beihang University	Zhongzhi Luan
Centre for Development of Advanced Computing	Kailash.S.
Chinese Academy of Sciences	Kejun Dong
Cloud Computing Innovation Council of India	Pamela Kumar
CNRS	Mohamad Awad
CSC	Schelleken
Electronice Research Institute Egypt	Salwa Nassar
Fundacion COMPUTAEX	Soporte Cenits/COMPUTAEX
GARR	Fulvio Galeazzi
Greek Research and Technology Network (GRNET)	Kostas Koumantaros
GWDG Göttingen	Piotr Kaspzrak/Philipp Wieder
I represent a few.	Ricus Ellis
IIHT Cloud Solutions Pvt. Ltd.	Felix Arun Kumar
INFN	Davide Salomoni

Institute of High Energy Physics, Chinese Academy of Sciences	Xiaofei Yan
iThemba LABS	Sean Murray
Khalifa University	Jacob Kuriyan
Ministry of Communication and IT - Afghanistan	Danyal
Mustansaryiah University IRAQ MOHESR	Bashar
North-West University, South Africa	Wilhelm van Belkum
ESKADENIA Software	Suhail Otoum
RADEI	Reiner Campillo
REDCLARA	Gustavo Garcia
SANREN (CSIR Meraka Institute) and SAGrid (CSIR Meraka Institute)	Uli Horn
Talal Abu-Ghazaleh Organization	Majd Farahat
The Italian National Institute for Nuclear Physics (INFN).	Davide Salomoni
University of Cape Town	Heine de Jager
University of Los Andes - Colombia	Andrés Holguín

Managerial Contacts	
INSITUTION	**NAME**
Alibaba Cloud Computing	Alibaba Cloud Computing
Beihang University	Zhongzhi Luan
Beihang University	Zhongzhi Luan
Centre for Development of Advanced Computing	Kailash.S.
Chinese Academy of Sciences	Kai Nan
Cloud Computing Innovation Council of India	Pamela Kumar
CNRS	Mohamad Awad
Electronice Research Institute Egypt	PROF Hesham El Deeb
Fundacion COMPUTAEX	Fundacion COMPUTAEX director
GARR	Federico Ruggieri
Greek Research and Technology Network (GRNET)	Kostas Koumantaros
GWDG	Dr. Philipp Wieder
GWDG Göttingen	Ramin Yahyapou
IIHT Cloud Solutions Pvt. Ltd.	Shivajee.R.Sharma
INFN	Davide Salomoni
Institute of High Energy Physics	Jingyan Shi
Institute of High Energy Physics, Chinese Academy of Sciences	Shi Jingyan
iThemba LABS	John Pilcher
Khalifa University	Dr. Arif AlHammadi
Ministry of Communication and IT - Afghanistan	Aimal Marjan
Mustansaryiah University IRAQ MOHESR	bashar
North-West University, South Africa	Wilhelm van Belkum
ESKADENIA Software	Suhail Otoum
RADEI	Jose Pedro Diaz Gonzalez
REDCLARA	Florencio Utreras
SANREN (CSIR Meraka Institute) and SAGrid (CSIR Meraka Institute)	Leon Staphorst
Talal Abu-Ghazaleh Organization	Ala Zaroure
The Italian National Institute for Nuclear Physics (INFN).	Davide Salomoni
Universidad Complutense de Madrid	Ignacio M. Llorente

6 Annex 3

6.1 Survey

This annex shows the whole survey as it was presented in the project web and circulated as a .doc document.

METHODOLOGY FOR THE SURVEY: This survey is to be circulated in the regions by the regional representatives – either online version or the paper version.
TARGET RECIPIENTS TO BE SURVEYED: Main target recipients are the cloud providers/developers. If the questionnaire is answered by the end-user, then this end-user should be a large institute, university, or a research group which uses external cloud resources, either public or private. The questionnaire is NOT meant purely research groups experimenting with cloud as a concept, or single end-users.

SECTION 1: GENERAL
Cloud is in this context referring to Virtualized resources (computing, storage, networks) that are provided within a data center or over a number of data centers, allowing the user access to configurable infrastructure, platform or software resources. The users in this context are the Research and Education community and our focus is on Infrastructure-as-a-Service clouds.

Q1.1 What do you consider as a cloud? Please provide us with your definition.

Q1.2 Are you aware of any *de facto* standards for cloud infrastructure, resources, and services

Yes (please specify) No

Q1.3 Which of the following cloud-related concepts and technologies are you familiar with on a theoretical level.

IaaS, PaaS, Saas, Amazon EC2/S3, EGI federated cloud, Hypervisor, Cloud Resource Manager, OCCI, CDMI, Identity Federation, etc.

Q1.4 Which of the following cloud-related technologies do you use?

IaaS, PaaS, Saas, Amazon EC2/S3, EGI federated cloud, Hypervisor, Cloud Resource Manager, OCCI, CDMI, Identity Federation, etc.

SECTION 2: CLOUD LANDSCAPE in your country/region
Q2.1 Are you aware of any clouds available to R&E in your region? If so please list them. For each please indicate

Q2.1.1 Is it public or private?

Please choose one of the following:

Public Private

Q2.1.2 Which kind of cloud can it be categorized as?

Please choose one of the following:

IaaS, PaaS, SaaS

Q2.1.3 What are the access policies to this cloud?

Q2.2 Which are the flagship cloud projects/initiatives in your region? Please list these if possible with pointers to the website/contact.

Q2.2.1 Are you aware, for each one of these, what is the funding body for these?

Q2.3 Are you aware of any government support to cloud initiatives in your country

Yes (please specify) No

Q2.4 Are you aware of any user communities in your region using cloud resources?

Q2.5: Are you aware of what applications and or data resources are mostly used on the cloud in your region?

SECTION 3: YOUR DETAILS AS A CLOUD CONSUMER:
complete only if you are either paying for external cloud resources (Amazon, ESB, etc) or are using another R&E cloud for the purposes of your institution.

Q3.1 Which organization do you represent?

Organization Name:

Q3.2 What are the name and coordinates of your technical contact?

Name:

Contact details:

Q3.3 What are the name and coordinates of your management contact?

Name:

Contact details:

Q3.4 Please provide the details of your cloud provider

Cloud provider name:

Cloud type:

Deployment method:

Q3.5 Which of the following represent the use your cloud infrastructure?

Educational / Commodities / Scientific applications / Others (please specify)

Q3.6 Please provide an estimate of monthly consumption of your institution (number of VMs, storage)

# of Virtual machines:

Size of storage:

Q3.7 Please provide an estimate of the monthly cost in USD

Q3.8 Please roughly indicate how is this cost covered (university budget, other..)

Q3.9 Please indicate the user-friendliness of the solution from 1-5, 5 being most friendly

Q3.10 Please indicate the level of IT know-how required by staff using these external resources, 5 being highest

Q3.11 Have you considered national data privacy and protection laws?

Q3.12 Is the external provider employed on a regular basis?

Q3.13 Is the external provider your main resource provider or do you only use it to satisfy peak demands?

Q3.14 Overall, are you satisfied for your use case by your provider ?

NOTE: FOLLOWING SECTIONS FOR CLOUD PROVIDERS ONLY

SECTION 4: YOUR DETAILS AS A CLOUD PROVIDER: GENERAL

Q4.1 Which organization do you represent?

Name:

Contact details:

Q4.2 What are the name and coordinates of your technical contact?

Name:

Contact details:

Q4.3 What are the name and coordinates of your management contact?

Name:

Contact details:

Q4.4 Do you use your cloud infrastructure for: Educational purposes / Commodities (Web servers….) / Scientific applications / Other

Q4.5 Briefly elaborate further for which scientific and educational topics are you providing virtualized resources?

Q4.6 For the resources you currently have internally – please indicate the % of the cost associated:

CAPEX cost: (compute hardware, data hardware, data centre, network infrastructure,

software licenses)

OPEX cost: manpower, electricity, networking, consumables (indicate total OPEX and sub-categories also if possible)

Q4.6 Sometimes it can be difficult to build from scratch, operate and employ a cloud infrastructure. For your local resources and infrastructure, what would you describe – in order of magnitude – as your three main barriers to entry:

for you as a resource provider

for your users

Q4.7 The employment of externally purchased cloud-based resources or services can be difficult or even impossible for different reasons (economical, technical, managerial…). For cloud-based resources and services, what would you describe – in order of magnitude – as the three main inhibitors/barriers to entry – for

you as a customer

your users

SECTION 5: YOUR DETAILS AS A CLOUD PROVIDER: VM Resource provisioning and management

Q5.1 Which hypervisor do you currently employ?

Q5.2 How many virtualized resources have you got, in terms of CPUs, memory and storage? (quantitative)

# of CPU (Sockets and Cores) per host:

Online Storage size (disks, Terabytes)

Nearline storage space (Tape systems, Terabytes)

Memory per CPU

Q5.3 Could you describe your virtual resources? (qualitative)

Q5.4 Do you employ any orchestrator / Cloud Resource Manager (e.g. StratusLab, OpenNebula, OpenStack, CloudStack, Synnefo etc.)?

Q5.5 Does your cloud implementation support Open Cloud Computing Interface standard (OCCI, http://occi-wg.org/) for public management of virtual machines?

Q5.6 Do you find issues of technologies for federation, such as OCCI, as important?

Q5.7 Would you be ready to federate your cloud into a larger international cloud, taking into account that security and reliability issues are ok?

SECTION 6: YOUR DETAILS AS A CLOUD PROVIDER: Data management

Q6.1 Do you provide any cloud storage?

Q6.2 What is the overall size of the resource?

Q6.3 Is there a distinct data interface/protocol available and if so what is it?

Q6.4 Do you support Cloud Data Management Interface (CDMI, SWIFT)?

SECTION 7: YOUR DETAILS AS A CLOUD PROVIDER: Operations

Q7.1 Do you employ any accounting mechanism for CPU usage, VMs, storage, network or applications?

Q7.2 If so, which one (for each of those)

Q7.3 Would you be considering to support APEL SSM 2.0 (https://wiki.egi.eu/wiki/Fedcloud-tf:WorkGroups:Scenario4) for inter-cloud accounting aggregation?

Q7.4 Do you employ any monitoring tool for your cloud infrastructure - for CPU usage, VMs, storage, network or applications?

Q7.5 If so, which one (for each of those)

Q7.6 Have you got any experience with EGI SAM monitoring system? If so, have you employed SAM to monitor a Grid site or cloud infrastructure before?

Q7.7 Does your organization run a user registration service?

Yes (Please specify) No

Q7.8 Does your organizations run a helpdesk? *

Yes (Please specify) No

Q7.9 Does your organization as a cloud provider have a Service Level Agreement/ monitoring in place?

Yes (Please specify) No

SECTION 8: YOUR DETAILS AS A CLOUD PROVIDER: Authentication and authorization

Q8.1 Do you have a user registration procedure?

Q8.2 Are you formally checking affiliation of the users to the academic institution?

Q8.3 What kind of data do you require from the users (name, organization, contacts)?

Q8.4 Are you providing your cloud as a service in some of Identity Federation (national federation or EduGAIN)?

Q8.5 Can users use their federated identity to access your resources?

Q8.6 Can the users use their identities for the open social web (e.g. Google, Facebook, Twitter, Github, etc) to access your resources?

Q8.7 Do you use x509 certificates for user Access? (Either Access to web interfaces or API Access) If so,

Q8.8 Do you rely on IGTF CAs, have you rolled out your own CA or you are using commercial CAs?

Q8.9 Do you support Group based policies on your infrastructure? (e.g. users X1 & X2 belong to group X which has a specific quota on the available resources)

SECTION 9: YOUR DETAILS AS A CLOUD PROVIDER: Cloud Federation and Information Discovery

Q9.1 Is your cloud infrastructure part of an Inter-cloud Federation? If not, would you be considering federating your cloud resources, within your country borders but also beyond?

Q9.2 If so, is the federation established among commercial providers, private resources, or a mix of both?

Q9.3 What would be your objective in this context (reliability, performance, economic matters…)?

Q9.4 If your infrastructure is already part of an inter-cloud federation, do you manually distribute the workload among the available resources (on your local physical resource or a federation), or do you employ any distribution manager? If so, which one?

Q9.5 Would you be considering to dynamically publish information about your cloud infrastructure so that your services are discoverable within an Inter-cloud Federation context?

Q9.6 Are you familiar with the GOCDB (http://goc.egi.eu/)? Would you be considering registering you cloud infrastructure in the GOCDB?

Q9.7 Are you familiar with the GLUE 2.0 schema?

SECTION 10: YOUR DETAILS AS A CLOUD PROVIDER: Security

Q10.1 Is there a person/team with the role of the Security/Officer in your Infrastructure?

Q10.2 Do you have an AUP (Acceptable use policy)?

Q10.3 Do you have a procedure for applying Security Patching? Is this documented?

Q10.4 How do you manage vulnerabilities in your Infrastructure?

Q10.5 Do you have a mechanism for Intrusion Detection?

Q10.6 How do you regulate access on your systems?

Q10.7 How do you manage contact information? (Users, Service Providers, Resource

Providers)?

Q10.8 How do you manage Incident Responses?

Q10.9 How do you manage Traceability? (User Identification, Production & Retention of logs)

Q10.10 How do you handle Personal Data (Account, Registration, Monitoring, Logging, and Scientific Results)?

SECTION 11: SOFT ACTIONS

Q11.1 Does your organization receive or provide cloud training?

Yes (please specify) No

Q11.2 Would you consider cloud workshops and or awareness seminars of value and importance

Yes (please specify) No