

Grids and Clouds Integration and Interoperability: an overview

Alberto Di Meglio¹

*CERN – European Organization for Nuclear Research
1211 Geneva, Switzerland
E-mail: alberto.di.meglio@cern.ch*

Morris Riedel

*Forschungszentrum Jülich
Jülich, Germany
E-mail: m.riedel@fz-juelich.de*

Shahbaz Memon

*Forschungszentrum Jülich
Jülich, Germany
E-mail: m.memon@fz-juelich.de*

Cal Loomis

*Laboratoire de l'Accélérateur Linéaire
Orsay, France
E-mail: loomis@lal.in2p3.fr*

Davide Salomoni

*INFN-CNAF
Bologna, Italy
E-mail: davide.salomoni@cnafl.infn.it*

Are grids and clouds different solutions to the same problems? Or are they simply different aspects of the same solution? Or maybe different solutions to different problems? Are they independent or complementary? Can they and should they be used together or is one a replacement for the other? Starting from the most accepted definitions of grids and clouds, this presentation describes the main differences and commonalities between the two models and the typical scenarios where grids and clouds can be used together or even merged into a common set of technologies and services. The talk gives an overview of the work being done in various contexts to make grids and clouds interoperable or integrable. Technological and operational aspects like virtualization, security, dynamic provisioning and standardization are briefly assessed. Finally the current work and future directions on cloud and grid

¹ Speaker

integration explored by a number of projects like EGI-InSPIRE, EMI and StratusLab in the context of European Research Infrastructures are introduced.

The International Symposium on Grids and Clouds and the Open Grid Forum
Academia Sinica, Taipei, Taiwan
March 19 - 25, 2011

1. Introduction

During the past decade, a number of important projects and initiatives have been run in Europe, the USA, Asia and other parts of the world to build collaborative infrastructures of shared resources to support intensive data analysis based on the concept of grid [1]. In Europe projects like the European Data Grid (EDG) [2], the Enabling Grids for E-science (EGEE) [3] and the KnowARC [4] projects have focused on High Throughput Computing, while DEISA and other projects have focused on integrating an infrastructure of High Performance Computing (HPC) resources to support large-scale computing simulations.

The provision of a European e-Infrastructure has been driven by the needs of user communities to access large scale computing and data analysis infrastructure to support their research needs as part of their pan-European research collaborations. Over the last decade, the European e-Infrastructure has benefited greatly from the growing maturity of the available open-source software solutions. Middleware consortia such as gLite, UNICORE, ARC and Globus, and specialised technology providers like dCache have developed new or improved existing solutions, in order to meet the needs of its user communities.

More recently the pressure on decreasing staffing costs, using green energy, enforcing economies of scale have promoted the consolidation of data centres and the wide-scale adoption of virtualisation in the commercial sector. The academic and research sectors are of course under similar constraints and look at new technologies with growing interest. Many campuses are encouraging the move of departmental or group level computing resources into central locations where they can be managed and supported by dedicated staff.

The trend is foreseen to continue over the next years, forcing a greater integration between existing production grid infrastructures and the ability to provision resources ‘on-demand’ that characterises the emerging cloud models [5]. The use of virtualisation to deliver ‘Infrastructure as a Service’ (IaaS), hosted environments to provide a ‘Platform as a Service’ (PaaS) and hosted applications to access ‘Software as a Service’ (SaaS) promise to facilitate access to the research infrastructure to growing numbers of researchers and lower the deployment and maintenance costs for resource and middleware providers.

2. Grid and Cloud Comparison

In order to understand how grid and cloud can interoperate and at what level, it is necessary to understand what their commonalities and differences are. Both grid and cloud are often referred to as “computing models” although this definition is not entirely correct and doesn’t take into account many of the more operational aspects of grid and cloud infrastructures. We need therefore to start with basic, but generally agreed definitions of grid and cloud and move from them to understand what their specific characteristics are, where they overlap and what distinguishes them.

2.1 Definitions

A comprehensive and generally agreed definition of grid can be found one an early short paper by Ian Foster, where a checklist is presented to determine whether a system is a Grid:

“A Grid is a system that coordinates resources that are not subject to centralized control [...] using standard, open, general-purpose protocols and interfaces [...] to deliver nontrivial qualities of service” [6]

Although a similarly comprehensive and generally agreed definition of what cloud is currently more difficult to find, the definition given by NIST¹ was recently accepted as the one of the most complete at the recent CloudScape conference in Brussels in March 2011²:

“Cloud computing is a model for enabling convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction” [7]

From the previous basic-principle definitions it is already possible to draw an important conclusion. A major difference between grid and cloud is essentially in the focus and scope. While grids focus on the coordination and federation of resources to provide a set of computing and data services to users, a cloud is essentially a resource provisioning model, which can be used by users to deploy and run their services. A second important conclusion is that there is no significant incompatibility between the definitions, which on the contrary hint at the fact that grid and cloud can be quite complementary if the former is seen as a service using (federated) resources and the latter as resources on which to provide a service. This implies of course that we primarily focus on the infrastructural aspects of cloud, what is generally referred to as IaaS (Infrastructure as a Service).

If we extend and go beyond the basic definitions of grid and cloud, of course more commonalities and differences can be found. A practical although not exhaustive list is show in Table 1. The case for interoperability between grid and cloud rests mainly on the differences highlighted in the table and indeed provide reasons for considering synergies between the federative aspects of grid and the rapid provisioning aspects of cloud, which are at the very base of its attractive business model.

¹ National Institute of Standards and Technology

² <http://www.sienainitiative.eu/StaticPage/Cloudscape.aspx>

	grid	cloud
Users own resources	No	No
Resources are in a known physical place	No	No
Resources are used and discarded when not needed anymore	Yes	Yes
API-based access	Yes	Yes
Federation or coordination of resources across different domains	Yes	No (not yet?)
Computational/storage abstraction	Job/file	Machine/application/service/disk
Fast VM provisioning/de-provisioning (using APIs)	No	Yes
Interactive	No	Yes
Customized environment	No	Yes
Cost model	Investment+maintenance	Pay-per-use

Table 1: Grid and cloud properties comparison:

3. The Use Case for Cloud and Grid Interoperability

During 2010 six projects³ were partially funded by the European Commission's 7th Framework Programme in the area of Distributed Computing Infrastructures. These projects cover different aspects of the research infrastructure service provision activities from resource management and operations to distributed computing, cloud and volunteer computing middleware. Their goal is to provide a pan-European production infrastructure built from federated distributed resources, ensure the continued support, maintenance and development of the middleware (gLite, ARC, UNICORE and Globus) that are in common use in Europe, explore how grid sites and different applications can be hosted sustainably in commercial, public, publicly procured and private 'cloud computing' environments, and provide desktop resources to the European research community.

³ EGI-InSPIRE, EMI, IGE, EDGI, StratusLab, VENUS-C

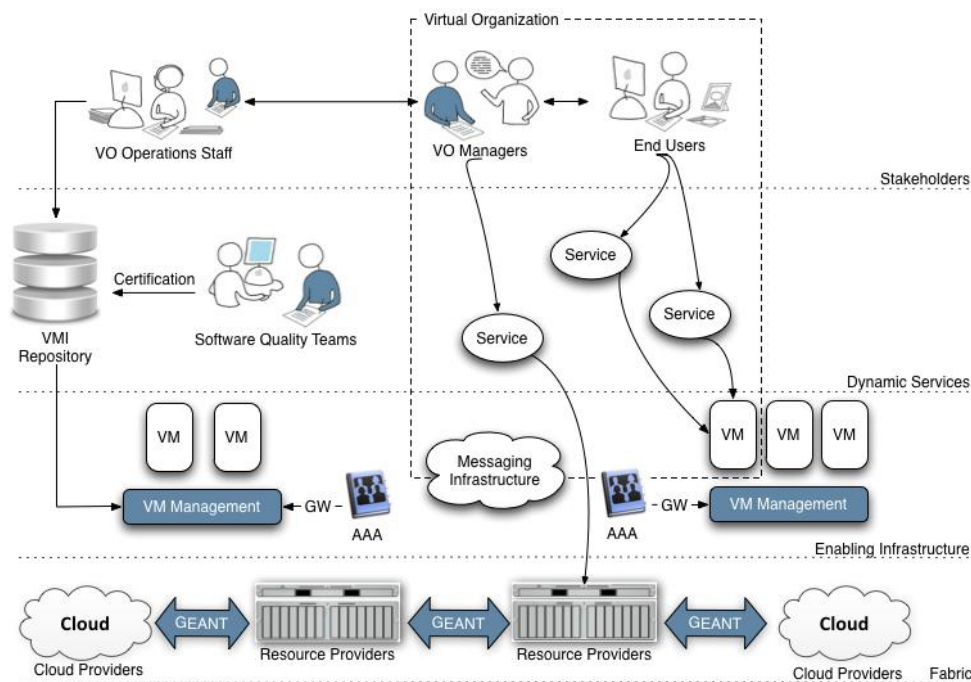


Figure 1: The European DCI projects service provision architecture

The six projects have designed together a vision of how distributed computing services provision can evolve in the next 2 to 4 year to build on existing mature technologies and add emerging new resource provision models based on the recent cloud business approach. This vision has been presented in September 2010 during the 1st EGI Technical Forum in Amsterdam [8]. It foresees the capability of the European research infrastructures to provision resources ‘on-demand’ to meet the needs of particular research collaboration by seamlessly integrating public resource providers alongside the academic resource providers offering virtualised compute resources.

The envisioned service provision architecture is depicted in Figure 1 and represents the main use case for the analysis of the grid and cloud interoperability scenarios.

4. Implementation Scenarios

Based on the defined use cases, four scenarios of increasing complexity have been investigated. One of the main criteria that have been considered is that the implementation of the scenarios should be possible without disrupting the infrastructure and the regular operations of site administrators and users. The four scenarios start therefore from the typical resource sharing architecture used by existing grid sites. The assumption is that the current resource sharing model is still a valid model for large scientific collaborations and that VO-based communities are still the default method of defining authorization and authentication policies, resource sharing policies, administrative responsibilities, and so on. The proposed scenarios are therefore practical attempts at integrating the grid model with technologies typical of the cloud

model. The main goal is to understand how technologies like virtualization and dynamic service provision can be used to simplify the administration of grid sites and at the same time how consolidated grid technologies like the authorization and authentication infrastructure can be used to extend the standard cloud architectures.

4.1 Scenario 1: Cloud resources behind grid services

This scenario consists in the use of virtualized resources as worker nodes in a standard grid infrastructure. The worker nodes are instantiated as needed by a virtualization manager (VM) that is seen by the Compute Elements (CEs) as a standard LRMS (Local Resource Management System). This is a scenario that in one form or another is already tested by many sites. The Infrastructure part is relatively simple to implement and requires very little modifications on the grid services with the possible exception of the adaption of the CEs to see the VMs as an (LRMS). The deployment is transparent for existing grid users and the impact on site administrators is limited to the need of deploying a VM in addition the the existitg LRMS. Standardization plays a limited role in this scenario, since it concerns the interface between CEs and VMs, where a standard API would allow to separate CE and VM implementations.

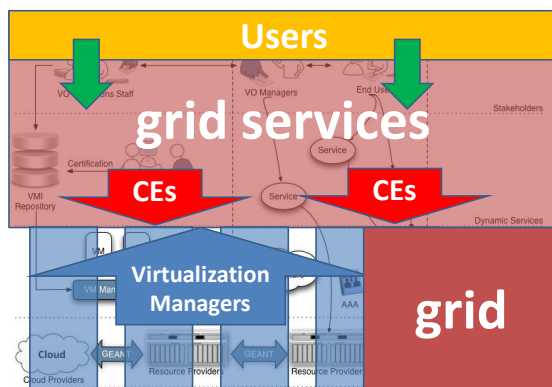


Figure 2: Cloud resources behind grid services

4.2 Scenario 2: Shared grid and cloud access

In this scenario part of the local resources are directly exposed to users via typical cloud interfaces. Access to grid services is unchanged, but users get access to local private clouds or to remote private clouds. Remote access can be based on existing Virtual Organization principles and make use of existing Authentication, Authorization and Accounting (AAA) technology already extensively used by grid services, like VOMS [9], Argus [10] or APEL [11]. The Infrastructure part is relatively simple to implement, but the access model requires extension of existing grid authentication and authorization policies and technologies, accounting and brokering mechanisms to the cloud systems.

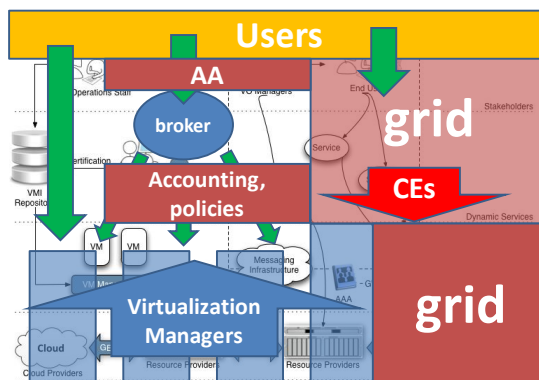


Figure 3: Shared grid and cloud access

Very little modifications on the grid services is necessary (the same as in Scenario 1). Standardization plays an important role in in the AAA services and in the Data Access services in case data sharing between the grid and the cloud services required. This can be done via mounted file systems on the virtualized nodes or via one or more of the existing and proposed data standards like SRM or CDMI. In the first case the mounted file systems would expose the same data to different nodes using a standard POSIX file system. In the second case the implementation of SRM or CDMI in existing grid and cloud data management service would allow transparent access to data irrespective of whether the requests come from traditional grid worker nodes and clients or on-demand virtual appliances instantiated by the users.

4.3 Scenario 3: Shared access with dynamic brokering

In this scenario the cloud brokering mechanism becomes part of the infrastructure services and its transparent to cloud users. Resource allocation is performed locally whenever possible, but resources are dynamically provisioned form collaborating private clouds or commercial public clouds as necessary to compensate for peak requests. The infrastructure part of this scenario is more complex to implement than previous cases. The access model requires extensions to the cloud virtualization managers of the authentication and authorization policies and technologies and of the accounting and brokering mechanisms already used in the grid. In addition potentially complex accouting and billing systems have to be put in place to interact with commercial public providers. Very few modifications are needed to the grid services as before. Most of the effort concerns site admins, who have to set up and manage the private and public cloud access services. In this scenario standardization also plays an important role. The implementation of common authentication, authorization and accounting methods in the Data Access services is necessary to enable data sharing between the grid and the cloud services. However, an additional complexity in this scenario is due to the need to interact with commercial cloud data service, which may not adhere to open standards. This scenario exploits the full benefits of using a cloud IaaS together with standard grid services and allows users to use custom execution environments, persistent and shared storage spaces, commercial cloud storage systems.

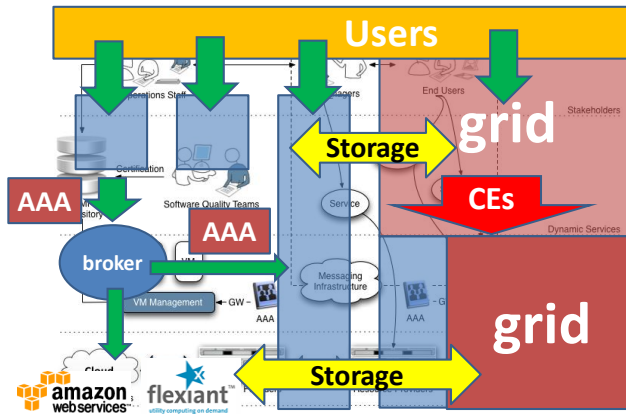


Figure 4: Shared access with dynamic brokering

4.4 Scenario 4: Dynamic grid services

In the previous scenarios grid and cloud interoperate at a very thin level, but are essentially separate infrastructures with data sharing needs. However, the full potential of cloud can be exploited by realizing that grid services are indeed just services that can be provisioned according to the standard cloud approach. This model is based on the assumption that grid services are specialized services needed by specific communities, but do not differ from a “business” point of view from other services required by other communities and provisioned using the dynamic capabilities of clouds in their PaaS or SaaS declinations. The grid services, or suitable subsets of the current grid services, can therefore be instantiated on demand using pre-packaged appliances or by deploying and configuring the services on base virtual machines according to specific user communities requirements and then disposed of when not needed anymore.

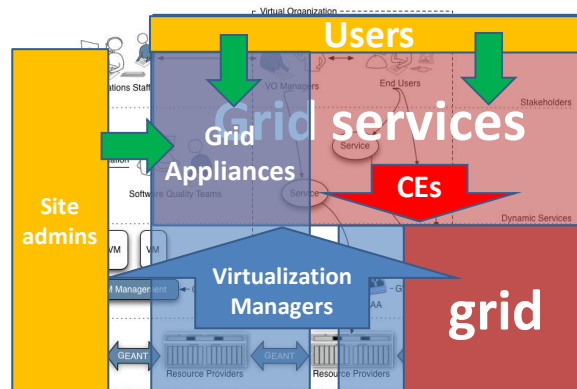


Figure 5: Dynamic grid services

The infrastructure and the service deployment aspects of this scenario are of course more complex to implement and, although the AAA requirements do not differ from those in scenario 2 or 3, potentially non-trivial modifications to grid services must be foreseen to make

them run as dynamically instantiated appliances or dynamically deployed and configured services. Grid access is still transparent for users, but self-consistent or self-contained grids are potentially easier to set up for site administrators or virtual organization managers. Standardization plays an additional important role in service discovery and registration, dynamic configuration and information systems.

5. Project and Activities

The definition, analysis and investigation of the four scenarios described above has been done in collaboration by EMI, StratusLab and WNoDes in the context of the DCI Roadmap implementation coordinated by EGI. In this section the project and activities are briefly described.

5.1 EGI

The European Grid Infrastructure (EGI)⁴ is a federation of over 40 national resource providers that together provide resources within Europe for over 13,000 users and their international collaborators. Over the last decade a production quality infrastructure has been established that provides secure access to federated computational and storage resources. Many of the current users and potential new users of this infrastructure are interested in having ‘cloud-like’ interfaces that will enable them to have greater control over the configuration of these resources and more flexibility in the software environments they deploy to support their work.

EGI.eu coordinates the resources across Europe on behalf of its stakeholders through the support of the EGI-InSPIRE project and is consulting with users and resource providers as to the cloud capabilities required by the user communities that the resource providers are able to deliver through the available technologies. As a result of this consultation a roadmap will be developed based on the technologies coming from outside the project identifying the policy, technical and operational work needed to fully integrate these new resource types into the production infrastructure. The scenarios described in this paper represent an input into this road mapping process.

5.2 EMI

The European Middleware Initiative (EMI)⁵ is a collaboration among the four major European providers of distributed computing middleware for reasearch infrastructures, ARC, dCache, gLite and UNICORE. The main goals of EMI are to reduce the current complexity of deploying and operating grid infrastructures, make the middleware more standard and accessible and fully adopt well established open source policies for packaging, release, distribution and support. At the same time EMI focuses on advancing the middleware functionalities to support the computing and data management requirements of growing scientific communities.

⁴ European Grid Infrastructure (EGI): <http://www.egi.eu>

⁵ European Middleware Initiative (EMI): <http://www.eu-emi.eu>

One of the critical requirements is to enable the grid services to address the emergence of on-demand computing, dynamic runtime environments, and a deployment over different types of virtual infrastructures. Within EMI, a Virtualization and Cloud Computing task force has been set up to investigate the possible options to operate grid services deployed over cloud computing infrastructures.

5.3 StratusLab

StratusLab⁶ is developing a complete, open-source cloud distribution that allows grid and non-grid resource centres to offer and to exploit an “Infrastructure as a Service” (IaaS) cloud. It is particularly focused on enhancing distributed computing infrastructures such as the European Grid Infrastructure (EGI).

By deploying and operating a production grid site for EGI on top of a StratusLab cloud, the project has demonstrated that its distribution and cloud technologies in general are sufficiently mature for use on production e-infrastructures. Feedback on operating grid services over a cloud will be used to improve the cloud distribution and to understand better how grid and cloud technologies can be brought together.

5.4 WNoDeS

The WNoDeS (Worker Nodes on Demand Service)⁷ is a software framework created by INFN to uniformly create, manage and access virtualized resources. WNoDeS is specifically targeted at Grid/Cloud integration and provides interfaces for Grid, Cloud and local access to a common set of compute and storage resources; these may be accessed through direct invocation of middleware services like European Middleware Initiative’s gLite Job Management products, Cloud APIs like the Open Cloud Computing Interface⁸, direct job submissions by local users, of via or a Web portal. WNoDeS has been written with scalability in mind and with the explicit purpose to be deployed into large data centers, where it fully and transparently integrates with existing scheduling and access policies. WNoDeS is in production since November 2009 at the INFN Tier-1 National Computing Center in Bologna, where it currently manages about 2000 dynamically created Virtual Machines, serving 20 international scientific collaborations.

6. Conclusions

Grid and cloud can be successfully integrated to provide even greater functionality and accessibility to researchers in Europe and elsewhere. Clear definitions of what aspects of cloud can be beneficial to grid and a specific focus on practical use cases are important to deliver value to existing users in reasonably short times.

A number of possible models have been explored and presented. Some of them are technically feasible today, but have to be implemented and deployed step by step with the goal

⁶ StratusLab: <http://stratuslab.eu>

⁷ WNoDeS: <http://web.infn.it/wnodes>

⁸ Open Cloud Computing Interface (OCCI): <http://occi-wg.org>

of minimizing the disruption on existing production infrastructures serving today thousands of users across hundreds of resource sites.

Cloud principles and technologies can be very beneficial to users and admins of grid, but many grid services are mature and standard and provide a solid backbone of functionality that should be extended and reused in a balanced compromise between stability of the infrastructures and the need to innovate and grow.

7. Acknowledgements

This work has been partially funded by the European Commission as part of the EMI (Grant Agreement INFISO-RI-261611) and StratusLab (Grant Agreement INFISO-RI-261552) projects.

References

- [1] I. Foster and C. Kesselman (Eds), *The Grid: Blueprint for a New Computing Infrastructure*, 2nd Edition, Morgan-Kaufmann, 2004
- [2] F. Gagliardi, *The EU DataGrid Project and other grid projects in Europe*, in Proceedings of ISCSE'02, Tokyo, March 5-6, 2002
- [3] E. Laure, R. Jones, *Enabling Grids for e-Science: The EGEE Project*, in Grid Computing: Infrastructure, Service, and Application, CRC Press, 2008
- [4] M. Ellert et al., *Advanced Resource Connector middleware for lightweight computational Grids*, Future Generation Computer Systems **23** (2): 219–240. (doi:10.1016/j.future.2006.05.008)
- [5] M. Armbrust, A. Fox, R. Griffith, A. Joseph, et al., *Above the clouds: A Berkeley view of cloud computing*, Electrical Engineering and Computer Sciences, University of California at Berkeley, TR No. UCB/EECS-2009-28, 2009 (<http://www.eecs.berkeley.edu/Pubs/TechRpts/2009/EECS-2009-28.html>)
- [6] I. Foster, *What is the Grid? A Three Point Checklist*, GRIDToday, July 20, 2002
- [7] P. Mell, T. Grance, *The NIST Definition of Cloud Computing, Recommendations of the National Institute of Standards and Technology*, NIST Special Publication 800-145 (Draft), January 2011
- [8] S. Newhouse, A. Di Meglio, A. Papaspyrou, et al., *Distributed Computing Infrastructure (DCI) Collaborative Roadmap*, EGI-InSPIRE Deliverable D2.4, October 2010
- [9] Y. Demchenko, C. de Laat, V. Ciaschini, *VO-based Dynamic Security Associations in Collaborative Grid Environment*, in Proceedings of the International Symposium on Collaborative Technologies and Systems, 38-43, 2006 (10.1109/CTS.2006.71)
- [10] V. Tschopp, *Argus, the EMI Authorization Service*, 1st EMI Technical Conference, Vilnius, April 2011, (<https://www.egi.eu/indico/getFile.py/access?contribId=145&sessionId=20&resId=0&materialId=slices&confId=207>)
- [11] M. Jiang, C. Del Cano Novales, et al., *An open and flexible distributed CPU usage accounting infrastructure for the WLCG and EGEE Grids*, 5th IEEE International Conference on e-Science, December 9-11, 2009 (oai:epubs.cclrc.ac.uk:work/52291)

