GridTalk GridBriefings
Grid computing in five minutes

Grids and clouds: the new computing

In areas as diverse as medical research, market analysis, high energy physics and online education, grids, clouds and computer clusters are providing the extra computing power needed for European innovation.

Clouds and grids are two complementary facets of European e-infrastructure that are developing in parallel. An understanding of their relationship, and the ways in which it will continue to change, will help the development and efficient use of existing academic and commercial e-infrastructures.

The changing face of computing: what's in a name?

Grids A grid is a collection of computers, usually owned by multiple parties and in multiple locations, connected together such that users can share access to their combined power. Examples include the Enabling Grids for E-sciencE infrastructure in Europe, or the Open Science Grid in the U.S.

Clouds A cloud is a collection of computers, usually owned by a single party, connected together such that users can lease access to a share of their combined power. Examples include Amazon's Elastic Compute Cloud, Google's App Engine, or IBM's Enterprise Data Center.

Grids and clouds share two main attributes:
1) they provide access to remote computing resources; and
2) they provide a service.

Grid computing is a relatively established form of distributed computing, adopted early by the high energy physics community and increasingly applied to all forms of e-science. In contrast, cloud computing has recently experienced a surge and is a service now offered by agrowing number of IT companies. In general, grids are based in academia and so are more likely to offer free access to shared resources, while clouds are commercial and have a pay-as-you-go approach. Further, grid technology is more focused on sharing different computing resources across a variety of sites, while cloud computing is designed more for on-demand provision of computer power or storage.



Virtual computers doing actual work

Virtualization allows you to create "virtual computers" or computer "images" that include the applications and operating systems of your choice. Any number of virtual machines can run independently on the same physical machine.

Many clouds use virtual machines to schedule tasks or "jobs": users can submit their jobs as virtual images that include all the required software and tools. Starting the virtual machine inside the physical machine provides immediate access to these tools, creating the ideal environment in which to work on the job. This is just one example of an area where cloud technologies can contribute to the development of grids.



Virtualisation allows actual computers to run additional virtual computers; for example, an Apple computer can run its own Mac operating system and a virtual Windows operating system, all from within the same machine.

Cloud computing for science?

Cloud computing has significantly changed the scale and simplicity with which businesses can access generic computing resources. However, the specific IT requirements of research communities have so far justified the existence of tailored grid-style e-infrastructures.

The cost of providing scientific computing and storage resources using cloud services relative to grid

infrastructures has been compared in a number of research papers. These have concluded that commercial clouds are "probably not yet" able to support the complex use scenarios that the research community demands. Equally, questions have been raised about the security aspects of cloud offerings.

It is possible that cloud services offered by industrial players such as Google or Amazon might at some point satisfy many of the advanced requirements of large scientific communities. In this case, publicly funded grids would need to offer services one step ahead of those provided by industry, catering to specific research demands that cannot be met with the same level of efficiency using the generic resource provision of clouds.

Integrating grids and clouds

Increasingly, "proof-of-concept" developments are investigating ways of integrating cloud-based services with existing grid-enabled applications across a number of disciplines, from nuclear physics to gene expression. Response from these projects regarding the benefits of integrations has been almost exclusively positive, but few, if any, have made the transition to everyday production. Potential benefits include advantages for end users (reducing the length and cost of application porting), and resource providers (reducing the operational cost of running grid sites).



Cloudscape: Daniele Lezzi, BSC for OGF-Europe


Cloudscape: Daniele Lezzi, BSC for OGF-Europe

Grid and cloud conclusions

Ultimately, individual projects and agencies need to make their own judgments on the relative cost and merit of cloud-based services as compared to grids. In a White Paper chapter devoted to the study of grids and clouds, the EU's e-Infrastructure Reflection Group concluded that the long-term interests of the research community would be served by access to and use of a mixture of grid- and cloud-based services and technologies.
The e-IRG White Paper recommended the following strategies as being integral to ensuring a healthy future for European e-infrastructures:

- Investigate the integration of cloud-like technologies, especially virtualization, into existing e-infrastructures
- Encourage the deployment of grid services on top of virtual infrastructures
- Encourage the integration of cloud-based services into existing e-infrastructures
- Promote the development of open-source components to build clouds
- Support the development of standards-based grid interfaces to cloud services
- Explore the application of grid technology for the federation of clouds
- Aim for high-level efficiency in both grid and cloud computing



"Even though virtualization requires more work by the user than running on a grid, legacy applications may find it much easier to execute on a cloud than on a grid. However, much still needs to be done to make clouds an execution environment of choice. Tools need to be developed to manage resource provisioning, to manage the environment customization process, and to help users evaluate the costs/benefits of running on a particular cloud, or across them." *Ewa Deelman*, *Information Sciences Institute, University of Southern California*



"Cloud computing doesn't really address fundamental issues around openness, collaboration and sharing. It currently requires that you effectively buy in to a proprietary system. Cloud computing also struggles to cope with the demands of high-end HPC or data processing, but is driving down the cost of supporting more standard computing provision. The basic questions are the same cost-benefit analyses of infrastructure provision, but with some new big "economy of scale" players." *Neil Geddes, UK Science and Technology Facilities Council e-Science Centre*

Left : OFG-Europe's Cloudscape workshop discussed the benefits and challenges of cloud computing.

Grids and clouds, side-by-side

	Grids (e.g. EGEE)	Clouds (e.g. Amazon)
What?	Grids enable access to shared computing power and storage capacity from your desktop.	Clouds enable access to leased computing power and storage capacity from your desktop.
Who provides the service?	Research institutes and universities federate their services around the world.	Large individual companies.
Who uses the service?	• Research collaborations. • "Virtual organizations," comprising researchers located around the world.	• Small to medium commercial businesses. • Researchers with generic IT needs.
Who pays for the service?	Governments: providers and users are usually publically funded research organizations.	The cloud provider pays for the computing resources; the user pays to lease them.
Where are the computing resources?	In computing centers distributed across different sites, countries and continents.	In the cloud provider's private data centers, which are often centralized.
Why use them?	• You don't need to buy or maintain your own personal computer center. • You can complete more work and tackle more difficult problems. • You can share data with your distributed team.	• You don't need to buy or maintain your own personal computer center. • You can quickly access extra resources during peak work periods.
What are they useful for?	Grids were designed to handle large sets of limited duration jobs that produce or use huge quantities of data.	Clouds best support long-term services and longer-running jobs.
How do they work?	Grids are an open source technology. Resource users and providers alike can understand and contribute to the management of their grid.	Clouds are a proprietary technology. Only the resource provider knows exactly how their cloud manages data, job queues, security requirements and so on.
Benefits	• **Collaboration:** grids offer a federated platform for distributed and collaborative work. • **Ownership:** resource providers maintain ownership of the resources they contribute to a grid. • **Transparency:** the technologies used are open source, encouraging trust and transparency. • **Resilience:** grids are located at multiple sites, reducing the risk in case of failure at one site.	• **Flexibility:** users can quickly outsource peaks of activity without long-term commitment. • **Reliability:** provider has financial incentive to guarantee service availability (Amazon, for example, can provide user rebates if availability drops below 99.9%). • **Ease of use:** relatively quick and easy for non-expert users to get started.
Drawbacks	• **Reliability:** grids rely on distributed services maintained by distributed staff, often resulting in inconsistency in reliability. • **Complexity:** grids are complicated to build and to use, and users require some level of expertise.	• **Generality:** clouds do not offer many of the specific high-level services currently provided by grid technology. • **Security:** users with sensitive data may be reluctant to entrust it to external providers. • **Opacity:** the technologies used to guarantee reliability and safety of cloud operations are not made public. • **Rigidity:** the cloud is generally located at a single site, which increases risk of complete cloud failure. • **Oligopolism:** there is a risk of being locked in to services provided by a very small group of suppliers.
When?	The concept of grids was proposed in 1997. The EGEE project began in 2004.	The term cloud computing debuted more recently, gaining prominence in 2007.



"While cloud computing is the latest buzzword in distributed computing, the grid concept is still largely unrealized and yet needed. As we move forward, clouds should be used to expose local resources, while grids can federate these resources into wider ubiquitous infrastructures. For this vision to materialize, we need lightweight and simple open source cloud solutions, and to depart from the current heavy grid interfaces and solutions. Only with this shift will the grid and cloud paths converge." *Marc-Elian Begin*, *SixSq.com*



"While grids are about interoperability, clouds are about on-demand provisioning of virtualized resources. These complementary technologies will coexist and cooperate at different levels of abstraction in future e-infrastructures. Clouds could help to overcome many of the barriers to grid adoption, and should also take advantage of research and development conducted by the grid community, for example in the adoption of standard grid interfaces and functionality for federation." *Ignacio M. Llorente, Complutense University of Madrid*



"Grids are a result of the research and academic community catering for complex scientific tasks and workflows requiring data sharing and compute processing. In contrast, clouds are industry's response to providing easy-to-use on-demand computing, storage and other specialised software resources. A mix of the best of these technologies can create a new breed that will benefit all users, including e-science and e-business. We should all be making an effort towards this direction." *Fotis Karayannis, Athens University of Economy and Business*



"For enterprises embracing a grid model, the emergence of large-scale commercial cloud providers offers at the very least another deployment option. An engrossing aspect of cloud computing is the potential emergence of a new type of business which exists wholly in the cloud, exploiting the rapid scaling and pay-per-use features of cloud computing to provide services to individuals and enterprises alike. How this new marketplace might develop remains to be seen. One thing is certain, cloud computing has caught the business world's attention and is now undergoing a level of debate and scrutiny that it (and grid computing) can only benefit from." Graeme Kerr, *Oracle EMEA External R&D*



"There is a particular benefit that cloud computing provides European Independent Software Vendors (ISVs). Innovation is one of the European software industry's fortes. Yet many companies have trouble growing in the European market and beyond because of the very complex procedures needed for an ISV to access a sizeable market. Cloud computing provides ISVs with access to a global marketplace." *Siada El Ramly*, European Software Association





"Cloud and grid are based on a similar vision of shared IT resources enabling organizations to solve problems otherwise out of reach through traditional technologies and IT delivery models. Cloud goes beyond grids, enabling a whole new range of applications for enterprises and consumers, deployed dynamically as a service. Many of the challenges facing cloud are similar to those facing grid, including the need to restructure applications, as well as security and reliability." *Eva Beck,* Hewlett-Packard

For more information:

e-Infrastructure Reflection Group (e-IRG)
www.e-irg.eu/

e-IRG White Papers
tr.im/c5fv

Enabling Grids for E-sciencE (EGEE)
www.eu-egee.org

Amazon Elastic Compute Cloud (EC2)
aws.amazon.com/ec2/

An EGEE Comparative Study:
Grids and clouds: evolution or revolution
tr.im/c5fz

OGF-Europe's "Cloudscape"
tr.im/c5gr

International Science Grid This Week
www.isgtw.org

GridTalk
www.gridtalk-project.eu

GridTalk is co-funded by the EC under FP7