e science TALK
e-ScienceBriefings
Talking about e-science

Cloud computing: What's on the horizon?

Cloud computing is making the revolutionary collaboration models of today's European e-Infrastructure more broadly accessible and applicable. Clouds can allow businesses, governments and educational institutes to access services and data on demand, and pay for what they need at the point of use. For those who need large amounts of computing power for short periods of time, the cloud appears to be the perfect solution, as well as a useful complement to existing e-Infrastructures.

A European vision

Neelie Kroes, Vice-President of the European Commission responsible for the Digital Agenda - *"Normally I prefer clearly defined concepts. But when it comes to cloud computing I have understood that we cannot wait for a universally agreed definition. We have to act."*

Clouds for storage and computing are changing the way that businesses, governments and academia deal with computing services. As such, Europe is keen to be at the forefront of developing this technology. According to Neelie Kroes, Vice-President of the European Commission responsible for the Digital Agenda, Europe should not just be 'cloud friendly' but 'cloud active'. At an address given to the World Economic Forum at Davos in January 2011, she outlined plans for an EU-wide cloud computing strategy. This will aim to look at the legal framework for cloud computing, resolve technical issues and establish a more open and competitive market for IT services.

A report from the EC entitled 'The Future of Cloud Computing' recommends Europe should exploit the available expertise and results from areas such as grid and other e-infrastructures to help realise the next generation of services on cloud systems. The EC intends to hold a series of discussions with both cloud providers and users with the aim of developing a plan for future actions by 2012.

The many faces of cloud computing

Cloud computing is still an ill-defined term and can be used to describe any of the following:

- **Infrastructure as a Service** (IaaS) - buying access to raw computing hardware over the internet, such as servers or storage. Also known as utility computing. IaaS often employs virtualisation in which users can create their own 'virtual computer', with specified applications, software and operating machine to deploy in the cloud. Advantages are that users don't need to worry about what software is being run where but disadvantages include that this is difficult to do.
- **Platform as a Service** (PaaS) - developing applications using web-based tools so they run on systems software and hardware provided by another company. The Google App Engine is an example of PaaS.
- **Software as a Service** (SaaS) - using a complete application running on someone else's system. Web based email is an example of this.

Clouds, grids and bioinformatics

StratusLab is applying cloud technologies to grid sites. By taking existing grid sites and running them on top of a cloud, StratusLab is simplifying the management of grids.

Grid sites at academic institutes can benefit from StratusLab, but these are by no means their only customers. The cloud distribution platform developed by the project is open for use by anybody who wants to outsource IT services. The StratusLab toolkit can provide increased flexibility thanks to virtualisation. Users often need applications with specific software requirements for their experiments.

StratusLab's ability to pre-define very specific and well suited virtual machines and to make them available in central repositories ready for deployment on academic clouds, would satisfy even the most demanding user. For example, StratusLab has already had interest from the French bioinformatics network ReNaBi.

Clouds for science

Clouds are increasingly being used by businesses, governments and scientists in areas from astronomy to zoology. Clouds can give scientists added flexibility for both storing and addressing data which can be easily accessed regardless of time or location. They can prove particularly useful for small research groups who require additional IT services.

Guy Coates, Wellcome Trust Sanger Institute - *"The life sciences community is reeling under the 'data deluge' generated by high throughput DNA sequencing. Sequencing technologies and data-production rates are rapidly evolving, placing IT organisations under increasing strain. The large data volumes produced require sophisticated data-tracking and large scale storage technologies, whilst analysis of the data needs access to HPC resources. On-demand cloud based IaaS and SaaS services from both academic and commercial providers can meet these data management and HPC requirements, and can react quickly to the rapidly changing needs of next-generation sequencing."*

A number of European Commission funded projects are demonstrating how academic and research communities can best profit from cloud services. For example the VENUS-C project brings together industrial partners and scientific user communities to develop, test and deploy an industry-quality cloud computing service for Europe.

The RESERVOIR project is also enabling massive scale deployment and management of complex IT services. The OpenNebula open source toolkit, a key outcome of this project, is being used as an open reference stack for cloud computing by several new European projects including StratusLab, BonFIRE and 4Caast.

Complementary technologies

Clouds are just the latest in a range of computing technologies such as grids and supercomputing which are available to researchers.

While clouds are sometimes seen as the successor to grids, the two technologies can complement each other as they offer users different benefits. For a full discussion of the differences between grids and clouds see the GridBriefing 'Grids and clouds: The new computing' (January 2009).

Clouds can provide significant added value to the existing range of computing resources. External parties can gain access to data grids and supercomputers through secure and paid access services via the cloud model. Integrating clouds with other forms of computing is an opportunity for development of both technologies. For example StratusLab aims to develop and deploy cloud technologies with the aim of simplifying and optimising the use and operation of distributed computing infrastructures such as the European Grid Infrastructure (EGI).

Charles Loomis, Linear Accelerator Laboratory - *"Cloud is a very selfish technology. Individuals acquire resources, use them for their work and get rid of them when they're finished. Grids are more about collaboration – sharing data and experiences. One of the challenges for cloud is how to enable federation, for example, how to move around data in an academic environment."*

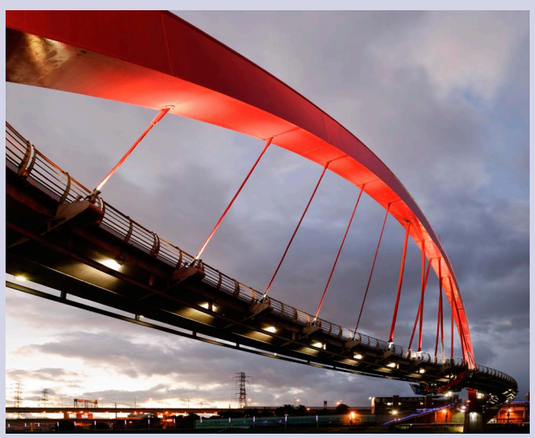
Moving components onto third-party clouds can be useful for a number of reasons including added resilience against events such as local power loss at sites. In the UK, GridPP, the grid for particle physics, has demonstrated that some grid components that are usually run locally can be successfully moved onto third-party clouds.

Building with clouds

The VENUS-C project is providing a platform for scientific applications, striving towards interoperable services which avoid vendor lock-in. VENUS-C supports a host of research including drug discovery, marine biodiversity and even predicting the risk of wildfires.

One of VENUS-C's most promising developments is a platform that can perform static and dynamic analysis of building structures using the cloud. This will involve running multiple simulations of buildings under a variety of conditions – be it different earthquake loads or building materials – leading to a reduction in construction costs and response time.

By working with VENUS-C, users have the opportunity to be the first to experiment with a cloud platform at industry-level quality. They gain direct access to developers, infrastructure experts and the computing platform until June 2013. They can test and improve their applications and business models and be prepared for a sustainable exploitation after the project ends. VENUS-C is an important opportunity to learn, test, share knowledge and gain visibility for research in the cloud.



Working together

Standards will play an increasingly important role in the drive towards an interoperable system of e-Infrastructures to ensure Europe retains its world-class position and plays a leading role in tackling global challenges.

The SIENA initiative is focused on accelerating and coordinating the evolution of interoperable distributed computing infrastructures by fostering an open and constructive dialogue with industry, standards bodies and major stakeholders. SIENA will build consensus on the best practices for cloud and grid technologies which are driving the development of standards. With the advent of cloud computing in both commercial and scientific settings, interoperability is becoming key to success, while the credit-squeezed economy puts additional pressures on eGovernment and research communities to ensure value for money from the public purse.

Silvana Muscella, Technical Director of SIENA - *"In collaboration with the European Commission, SIENA is working with major stakeholder communities on a roadmap focusing on interoperability and standards. The roadmap will define scenarios, identify trends and investigate the innovation and impact sparked by cloud and grid computing, delivering insights into how standards and the policy framework are shaping current and future development and deployment in Europe and globally. The development of the roadmap is timely in support of the aims of the Digital Agenda for Europe, which reinforces the need for effective interoperability between IT products and services to build a truly digital society by 2020."*

Addressing user concerns

The fast pace of cloud development has raised a number of technical and commercial questions, such as security, access and availability of cloud services. Unlike collaborative

Christine Morin, Project Coordinator Contrail - *"Contrail aims at removing some constraints that are currently present in cloud systems. It will federate clouds allowing companies and research organisations to easily switch from cloud providers. Contrail will develop a complete, secure cloud computing stack, including Platform-as-a-Service and Infrastructure-as-a-Service."*

infrastructures such as grids most cloud providers have unique and proprietary application programming interfaces (APIs). However to gain the best deals, users would like to be able to move their data around from provider to provider. Provider lock-in is a major concern.

The Contrail and mOSAIC projects are both exploring ways of making choosing a particular cloud easier. The mOSAIC project, especially, hopes to tackle the issue of provider lock-in. It will develop a platform to demonstrate how users can determine and access the cloud provider that offers the best services for them.

The overall quality of service such a multi-cloud environment offers can be hard to predict. With single clouds risks can be mitigated through Service Level Agreements (SLAs) between users and providers. However, managing SLAs between multiple parties requires applying lessons already

Improving drug research

Contrail is investigating how cloud computing could help cut costs and increase the impact of drug research.

The massive explosion in the volume of data generated through genomic research, pharmacological sources and clinical trials has greatly increased the number of potential drug compounds. However finding promising candidates using only traditional computing technology is hugely challenging. The cloud makes it possible to use more computing and data storage power at the same cost.

Accessing and analysing the data through the cloud, and drawing on the unused computing resources of other companies or organisations, could potentially lower the cost of commercial electronic drug discovery services. It could also enable SMEs to play a competitive role in the pharmaceutical industry.

learned from other e-Infrastructures. For example the gSLM project analyses grid environments to come up with recipes and best practices that address this challenge.

Bob Jones, Chair of the European e-Infrastructure Forum - *"For scientific research communities, the question is not whether they will use cloud computing but rather how quickly and under what conditions."*

Legal and security minefields

By putting data into the cloud, users can lose control of where it is stored. Cloud providers operate across or in many jurisdictions - data centres span countries, and even continents. In these cases determining which country's laws apply to stored data is not straightforward.

For scientists working with sensitive data such as health records, a private or community cloud may be a better solution to commercial cloud providers as users can be sure of where data is being stored. Otherwise legal issues such as data protection, privacy and user's rights will need to be factored into cloud provisioning.

Gudmund Høst, e-IRG Chair - *"The e-IRG has had a focus on cloud computing since 2008, where the e-IRG White paper 2009 made a first comparison of clouds and grids. The e-IRG roadmap also gave recommendations on commodity and cloud computing 'recognising the importance of transparency and compatibility of the organisational and financial models as key factors in maximising the benefits of broad commodisation of computing services for all scientific users.' The e-IRG Blue paper 2010 to ESFRI touched on the issues of grid, cloud and virtualisation proposing 'collaboration among grid and cloud infrastructure providers and users to raise awareness of the range of available technologies and how to best use them'. The current e-IRG White Paper 2011 is also considering cloud computing as a natural element within the new service-oriented e-Infrastructure."*

Groups such as the e-IRG (e-Infrastructure Reflection Group) have been looking into both legal and governance issues for e-Infrastructures. Regarding security, the Cloud Security Alliance works to promote the use of best practices for providing security assurance within cloud computing. Events such as Cloudscape also provide an opportunity for experts, developers and end users to have meaningful discussions on the subject

Governing from clouds

A number of countries are employing the cloud model for use in government. The US has employed clouds for transparent government, while the EU uses the technology for public procurement. In Japan clouds are being deployed to aid collaboration across ministries. This shift to indirect governance of data through clouds leads to a number of specific challenges.

Governmental data could be subject to the laws and regulations of other countries if moved beyond national borders. National laws and regulations of the Member States of the European Union currently impose some restrictions on the movement of data outside national territory. However due to proprietary systems of cloud providers this may be hard to monitor and enforce.

The European Network and Information Security Agency (ENISA) suggests that there are a number of factors that will need to be taken into account in this situation. The ENISA report 'Security & Resilience in Governmental Clouds' proposes that governments assess whether current legal frameworks can be changed to facilitate the communication, treatment and storage of data outside their national territory, without exposing the security and privacy of citizens or national security and the economy to unacceptable risks.

Chrysanthi Papoutsi, Oxford Internet Institute - *"e-Infrastructure governance becomes increasingly important in the effort towards sustainability and cost-effectiveness. Long-term experience drawn from governance in other IT related areas can be used productively to achieve strategic alignment. In the face of the opportunities and challenges presented by technologies such as cloud computing, careful consideration of the legal issues involved relies partly on governance structures and mechanisms, which enable different decision-making processes in different e-Infrastructures."*

For more information:

Contrail: ***www.contrail-project.eu***

e-IRG: ***www.e-irg.eu***

ENISA: ***www.enisa.europa.eu***

Grids and Clouds GridBriefing: ***www.e-sciencetalk.org/briefings/GridBriefing_Grids_and_clouds.pdf***

HPC in the Cloud: ***www.hpcinthecloud.com***

Mosaic: ***www.mosaic-cloud.eu***

SIENA: ***www.sienainitiative.eu***

StratusLab: ***www.stratuslab.eu***

The Future of Cloud Computing report: ***http://cordis.europa.eu/fp7/ict/ssai/docs/cloud-report-final.pdf***

Venus-C: ***www.venus-c.eu***

EGI: ***www.egi.eu***

iSGTW: ***www.isgtw.org***

e-ScienceTalk: ***www.e-sciencetalk.org***

Scan this QR code into your smart phone for more on this e-ScienceBriefing

e-ScienceTalk is co-funded by the EC under FP7