

Title: Strategic Plan for a Scientific Cloud Computing infrastructure for Europe

Authors:

Bob Jones, CERN

Maryline Lengert, ESA

Rupert Lueck, EMBL

Keywords: cloud computing, infrastructure, European Research Area, EIROforum, sustainability

Abstract:

Besides the core processes of R&D agencies and organisations of capturing, processing, analysing and archiving data, a new paradigm of having permanent on-line access to IT resources, information and collaboration tools has become a central aspect of scientific endeavour. Research infrastructures such as the members of EIROForum are beginning to challenge the sustainability of an approach to ICT deployment that has predominated for a quarter of a century but is rapidly being overtaken by events.

In the US, for example, the Federal Cloud Computing Strategy [1] estimates that by moving a quarter of its annual \$80bn ICT expenditure to cloud-based services, it will achieve greater efficiency, agility and innovation in delivery of services – and possibly save 10% of that budget. These efficiencies are primarily brought about by achieving economies of scale, multiple tenancy of irregularly-used resources and more sophisticated approaches to resource management. More fundamentally, making a business case for moving some or all ICT facilities to a cloud-service model at the end of their current life necessitates a reappraisal of the way those facilities are utilised, potentially transforming one of the significant operational cost centres.

The feasibility of such a transformation is, therefore, of great interest to the suppliers of cloud-based services and to the national and European funding agencies as well. The HELIX NEBULA initiative is a preliminary step towards a European cloud-based scientific e-infrastructure: HELIX NEBULA – the Science Cloud. The Science Cloud Strategic Plan [2] was presented and adopted by representatives of all three stakeholder groups at a workshop hosted by ESA/ESRIN in June 2011. The supporters of the HELIX NEBULA initiative include: ATOS, BT Global Services, CAP Gemini, CERN, CloudSigma, CNES, CNR, DLR, EC, EGI.eu, EMBL, ENEA, ESA, Logica, OpenNebula, Orange Business Services, SAP, SIENA, StratusLab, Server Labs, Six², T-Systems International, Terradue srl, Thales and TrustIT.

The HELIX NEBULA initiative will lead and co-ordinate these communities of interest through a two year pilot-phase during which procurement processes and governance issues for a framework of public/private partnership will be appraised. Three flagship use cases from HEP, molecular biology and earth-observation will be used to validate the approach, enable a cost-benefit analysis to be undertaken and the next stage of the Science Cloud Strategic Plan developed and approved. The Science Cloud infrastructure will be initialized with large-scale unique data sets and bring thousands of scientists on the platform.

GEANT, EGI (European Grid Initiative) and PRACE (Partnership for Advanced Computing in Europe) are existing e-infrastructures heavily used by the research community and co-funded by the European Commission. Interoperability with these infrastructures is important to allow the research community to migrate their data and applications between all the infrastructures and combine them so that the most appropriate resource can be used in an efficient manner. Cost-effective and high performance network access to the science cloud across the ERA will be an important aspect of its deployment. GEANT could provide high-performance network connectivity to HELIX NEBULA – the Science Cloud. Terena and the NRENs could play an important role in this initiative by facilitating the connection to the data centres of commercial cloud service providers. Examples of such connections already exist around Europe [3].

References:

- [1] <http://www.cio.gov/documents/Federal-Cloud-Computing-Strategy.pdf>
- [2] <http://cdsweb.cern.ch/record/1374172/files/CERN-OPEN-2011-036.pdf>
- [3] <http://www.ja.net/janetnews/2011/01/13/logicalis-connects-high-density-data-centre-and-hybrid-cloud-into-janet/#more-45>

Biography:

Bob Jones is head of the CERN openlab project (openlab.cern.ch) which facilitates collaboration between CERN and its industrial partners to study and develop data-intensive solutions for scientists working at the next-generation Large Hadron Collider (LHC). His experience in the distributed computing arena includes mandates as the technical director and then project director of the European Commission co-financed EGEE projects (2004-2010 <http://www.eu-egee.org>), which established and operated a production grid facility for e-Science spanning 300 sites across 48 countries for more than 12,000 researchers.

Rupert Lueck, Head of IT Services at EMBL headquarters in Heidelberg. Rupert is responsible for the laboratory's IT, including operations and development of the high performance computing infrastructures and the provision of first class IT services to EMBL's science communities. With 18 years in professional IT and lead positions during the past 13 years in international life science IT environments, his roles included global responsibility for the Professional Services business of a dual-listed bioinformatics company targeting global pharmaceutical industry and biotech firms after having led the company's global IT infrastructure and services unit. Rupert holds a degree in Biology and a PhD in Bioinformatics.

Maryline Lengert, Senior Advisor in the IT department of ESA, started this Cloud Computing Strategic Plan initiative end of 2009 to catalyse the creation of a European owned public cloud, serving initially the Science Research Area. Maryline has 20+ years of experience in the IT domain within an international and distributed environment. Her previous positions include Technical Manager and Operations Director for Hospital Integrated Information System in S.A.I.C., Quality Manager and Head of the Account Management and Requirement Analysis Division in the ESA IT Department. Maryline holds a PhD in Physics and a Master in International and European Studies.