The EPOS Vision for the Open Science Cloud

Keith Jeffery (1), Matt Harrison (1), and Massimo Cocco (2)

(1) British Geological Survey, Keyworth Nottingham, UK, (2) Istituto Nazionale di Geofisica e Vulcanologia, Rome, Italy

The EPOS Vision for the Open Science Cloud
Keith Jeffery (1), Matt Harrison (1) Massimo Cocco (2), and the EPOS Consortium (3)
(1) British Geological Survey, Keyworth, UK (kjeff@bgs.ac.uk, mharr@bgs.ac.uk)
(2) Istituto Nazionale di Geofisica e Vulcanologia, Rome, Italy (massimo.cocco@ingv.it)
(3) EPOS www.epos-eu.org

Cloud computing offers dynamic elastic scalability for data processing on demand. For much research activity, demand for computing is uneven over time and so CLOUD computing offers both cost-effectiveness and capacity advantages. However, as reported repeatedly by the EC Cloud Expert Group, there are barriers to the uptake of Cloud Computing: (1) security and privacy; (2) interoperability (avoidance of lock-in); (3) lack of appropriate systems development environments for application programmers to characterise their applications to allow CLOUD middleware to optimize their deployment and execution.

From CERN, the Helix-Nebula group has proposed the architecture for the European Open Science Cloud. They are discussing with other e-Infrastructure groups such as EGI (GRIDs), EUDAT (data curation), AARC (network authentication and authorisation) and also with the EIROFORUM group of 'international treaty' RIs (Research Infrastructures) and the ESFRI (European Strategic Forum for Research Infrastructures) RIs including EPOS. Many of these RIs are either e-RIs (electronic-RIs) or have an e-RI interface for access and use.

The EPOS architecture is centred on a portal: ICS (Integrated Core Services). The architectural design already allows for access to e-RIs (which may include any or all of data, software, users and resources such as computers or instruments). Those within any one domain (subject area) of EPOS are considered within the TCS (Thematic Core Services). Those outside, or available across multiple domains of EPOS, are ICS-d (Integrated Core Services-Distributed) since the intention is that they will be used by any or all of the TCS via the ICS. Another such service type is CES (Computational Earth Science); effectively an ICS-d specializing in high performance computation, analytics, simulation or visualization offered by a TCS for others to use. Already discussions are underway between EPOS and EGI, EUDAT, AARC and Helix-Nebula for those offerings to be considered as ICS-ds by EPOS..

Provision of access to ICS-Ds from ICS-C concerns several aspects:
(a) Technical : it may be more or less difficult to connect and pass from ICS-C to the ICS-d/ CES the 'package' (probably a virtual machine) of data and software;
(b) Security/privacy : including passing personal information e.g. related to AAAI (Authentication, authorization, accounting Infrastructure);
(c) financial and legal : such as payment, licence conditions;
Appropriate interfaces from ICS-C to ICS-d are being designed to accommodate these aspects. The Open Science Cloud is timely because it provides a framework to discuss governance and sustainability for computational resource provision as well as an effective interpretation of federated approach to HPC(High Performance Computing) –HTC (High Throughput Computing). It will be a unique opportunity to share and adopt procurement policies to provide access to computational resources for RIs.

The current state of discussions and expected roadmap for the EPOS-Open Science Cloud relationship are presented.