



Virtual Geophysics Laboratory: Exploiting the Cloud and Empowering Geophysicists

Ryan Fraser (1), Josh Vote (1), Richard Goh (2), and Simon Cox (1)

(1) CSIRO, Australia, (2) Curtin University, Western Australia

Over the last five decades geoscientists from Australian state and federal agencies have collected and assembled around 3 Petabytes of geoscience data sets under public funding. As a consequence of technological progress, data is now being acquired at exponential rates and in higher resolution than ever before. Effective use of these big data sets challenges the storage and computational infrastructure of most organizations. The Virtual Geophysics Laboratory (VGL) is a scientific workflow portal addresses some of the resulting issues by providing Australian geophysicists with access to a Web 2.0 or Rich Internet Application (RIA) based integrated environment that exploits eResearch tools and Cloud computing technology, and promotes collaboration between the user community.

VGL simplifies and automates large portions of what were previously manually intensive scientific workflow processes, allowing scientists to focus on the natural science problems, rather than computer science and IT. A number of geophysical processing codes are incorporated to support multiple workflows. For example a gravity inversion can be performed by combining the Escript/Finley codes (from the University of Queensland) with the gravity data registered in VGL. Likewise, tectonic processes can also be modeled by combining the Underworld code (from Monash University) with one of the various 3D models available to VGL. Cloud services provide scalable and cost effective compute resources. VGL is built on top of mature standards-compliant information services, many deployed using the Spatial Information Services Stack (SISS), which provides direct access to geophysical data. A large number of data sets from Geoscience Australia assist users in data discovery. GeoNetwork provides a metadata catalog to store workflow results for future use, discovery and provenance tracking.

VGL has been developed in collaboration with the research community using incremental software development practices and open source tools. While developed to provide the geophysics research community with a sustainable platform and scalable infrastructure; VGL has also developed a number of concepts, patterns and generic components of which have been reused for cases beyond geophysics, including natural hazards, satellite processing and other areas requiring spatial data discovery and processing.

Future plans for VGL include a number of improvements in both functional and non-functional areas in response to its user community needs and advancement in information technologies. In particular, research is underway in the following areas (a) distributed and parallel workflow processing in the cloud, (b) seamless integration with various cloud providers, and (c) integration with virtual laboratories representing other science domains.

Acknowledgements: VGL was developed by CSIRO in collaboration with Geoscience Australia, National Computational Infrastructure, Australia National University, Monash University and University of Queensland, and has been supported by the Australian Government's Education Investment Funds through NeCTAR.