Multi-Cloud Earth Observation Data Processing with Commercial and Institutional Clouds

Alastair PIDGEON, Martin LANGE, Marc-Elion BEGIN, Lukasz DUTKA, Michal ORZECHOWSKI, Matthew VILJOEN, Yannick LEGRE, Roberta PISCITELLI, Emmanuel MONDON, *Cedric SEYNAT

1. RHEA Group, Avenue Pasteur 23, 1300 Wavre, Belgium; a.pidgeon@rheagroup.com
2. EOProc, Maronstrasse 11, 81373 Munich, Germany; mlange@eoproc.com
3. SixSq, Rue du Bois-du-Lan 8, 1217 Meyrin, Switzerland; meb@sixsq.com
4. Cyfronet, Akademickie Centrum Komputerowe, ul. Nawojki 11 30-950, Kraków 23, Poland; lukasz.dutka@cyfronet.pl, michal.orzechowski@cyfronet.pl
5. The EGI Foundation, Science Park 140, 1098 XG Amsterdam, The Netherlands; matthew.viljoen@egi.eu, yannick.legre@egi.eu, roberta.piscitelli@egi.eu
6. AdviceGEO, KönigsdorferStrasse 15, 81371 Munich, Germany; emondon@advicegeo.com
7. RHEA Group, 6700 Cote-de-Liesse, suite 102, Montreal, H4T2B5, Canada; c.seynat@rheagroup.com

* Corresponding Author

ABSTRACT

The Earth Observation (EO) industry is now rapidly changing, both in terms of market and technology. A sustained growth of the EO market worldwide is ongoing and expected to continue. Fueled by publicly funded programs and commercial initiatives, EO data supply will dramatically increase in the coming years. Access to EO data at affordable prices, open data policies, and the advent of cloud-based platforms supporting the provision of EO value-added services, are opening huge opportunities to democratize the use of EO. Recently the concept of Information as a Service (InfoaaS) has emerged to describe an EO service delivery model encompassing the on-demand delivery of Information Technology infrastructure (storage, processing), of EO data and value-added products, and of software applications, delivered in an open ecosystem where suppliers and users of services can engage.

However, many stakeholders in the EO value-added chain are concerned by the complexities associated with the discovery and management of multiple EO data sources, and by the risk of being locked-in to proprietary solutions from data suppliers or cloud infrastructure vendors. To address this issue, we have developed a proof-of-concept for an open platform delivering a multi-cloud EO data processing service, with the objective to prevent cloud vendor lock-in situations while addressing end-to-end processing of EO data in multiple cloud environments. Our vision is to offer a generic and open service, intended to work with any EO data located anywhere and with any EO processor, that can be rapidly deployed with flexibility, scalability, at low cost, and which provides services accessed via public and standardised APIs.

Our multi-cloud EO data processing platform has been demonstrated with the EGI Federated Cloud, Innovation Platform Testbed (IPT) Poland, and the Amazon Web Services (AWS) cloud. It presents the following capabilities:

- **Multi-cloud data discovery**: a data discovery service allows to discover and select EO scenes relevant to an EO application. It uses a metadata catalogue currently tracking nine satellites including Landsat-8, Sentinel-1 and 2, and currently able to hold up to 10+ millions of scenes.

- **Multi-cloud data management and access**: a data management service provides a data location API that allows to expose and stage EO data on the cloud infrastructure on which the EO application is deployed and executed. It uses a global high-performance data management system providing access to distributed storage resources.

- **Multi-cloud application deployment**: an application deployment service allows to select the cloud infrastructure on which to deploy the EO application, and provides automated deployment, execution and monitoring. It uses a multi-cloud application management service automating the full application management lifecycle, and providing connectors to multiple clouds, public and private, leveraging both open source and proprietary cloud APIs.

This presentation outlines the concept of InfoaaS, addresses the underlying technologies supporting our end-to-end multi-cloud EO data processing platform, and presents concrete examples of its use, showing how Sentinel-1 data was retrieved and processed to produce change detection maps of a glacier in Greenland and of shipping activity in the Panama canal. Another use case is also shown, that is based on the use of a Sentinel-2 data processor to generate true colour 10m resolution images. The applicability of such multi-cloud EO data processing services to the Canadian context is also presented.