



A Cloud-based Science Gateway for the Geoscience Community with End-to-end Workflows on the Jetstream Cloud System

Mohan Ramamurthy and Juilien Chastang

University Corporation for Atmospheric Research, Unidata, Boulder, United States (mohan@ucar.edu)

Unidata is a diverse community of education and research institutions with the common goal of sharing geoscience data and the tools to access and visualize that data. For more than 30 years, Unidata has been providing data, software tools, and support to enhance Earth-system education and research. Funded primarily by the National Science Foundation (NSF), Unidata is one of the University Corporation for Atmospheric Research (UCAR)'s Community Programs (UCP).

The advent and maturity of cloud computing technologies and tools have opened new avenues for addressing both Big Data and Open Science challenges to accelerate scientific discoveries. There is broad consensus that as data volumes grow rapidly, it is particularly important to reduce data movement and bring processing and computations to the data. Data providers also need to give scientists an ecosystem that includes data, tools, workflows and other end-to-end applications and services needed to perform analysis, integration, interpretation, and synthesis - all in the same environment or platform. Instead of moving data to processing systems near users, as is the tradition, one will need to bring processing, computing, analysis and visualization to data – so called data proximate workbench capabilities, also known as server-side processing.

As part of its current five-year plan "Unidata 2018: Transforming Geoscience through Innovative Data Service", Unidata is developing data-driven scientific workflows using cloud computing technologies for accessing, analyzing, and visualizing geoscience data. The Unidata Science Gateway has implemented Unidata-related technologies hosted on Jetstream, a cloud facility funded by the U. S. National Science Foundation. Through the Science Gateway, researchers can make use of well-integrated resources either directly in their browser or using one of the client applications. Our hope is that by combining the resources available in this gateway – and perhaps coupling them with XSEDE HPC resources – community members can create end-to-end scientific computing workflows in the Jetstream Cloud.

In this talk, we report on our ongoing efforts, experiences, lessons learned, and offer a vision for future scientific workflows using data-proximate tools and services to advance research and education in the geosciences.