

Abstract -- EarthCube All-hands Meeting 2018

Contribution ID : 118

Type : Poster Abstract

Title : Growing the foundation for Pangeo: Bringing together MetPy and XArray

Author(s) : May, Ryan Michael; Leeman, John

MetPy's is a Python toolkit for meteorology, encompassing tools for reading data, performing calculations, and making plots. As part of the Pangeo project, whose goal is to provide a framework for analyzing earth system model output that scales to petabyte-scale datasets, MetPy serves as a set of domain-specific functionality that rests on a foundation based on other scientific Python libraries, such as numpy and matplotlib.

In order to scale to the needs of large datasets, Pangeo has identified the need to leverage the XArray and Dask libraries as part of this foundation. XArray provides a standard data model for n-dimensional gridded data based on the netCDF data model, similar to the Common Data Model used within the netCDF-Java library. Dask provides a framework for distributed computation that greatly simplifies the task of doing out of core computation, necessary to work with petabyte-scale datasets.

This work discusses the experience of integrating MetPy with this broader foundation, including concrete examples of the user-facing benefits that have been achieved, as well as the challenges encountered, such as the integration of MetPy's support for physical units with XArray.