EGU2020-13133
https://doi.org/10.5194/egusphere-egu2020-13133
EGU General Assembly 2020

Distributed Earth-Observation satellite data processing with Pytroll/Satpy

Salomon Eliasson, Martin Raspaud, and Adam Dybbroe
Swedish Meteorological and Hydrological Institute, Atmospheric Remote Sensing Unit, Norrköping, Sweden
(salomon.eliasson@smhi.se)

As earth-observing (EO) satellite data volume is growing, servers struggle to keep up with the computational load needed to process even single segments of satellite data with reasonable performance. Pytroll is a suite of free and open-source python tools from which the Satpy package is made to easily and efficiently read, process, write EO satellite data. To obtain computational efficiency, Pytroll approaches the performance problem from multiple angles through optimized data processing, built-in support for out-of-memory computations (using the underlying Dask python library), and allowing distributed processing (using the Dask Distributed tools). In this work, we will show how large volumes of satellite data can be read, processed, resampled, and written swiftly and easily with the Pytroll/Satpy package, in a cluster environment. Specifically, examples of efficiently processing Sentinel-1, Sentinel-2, and Himawari/AHI data will be shown, along with performance figures.