Flocks of a feather dock together: Using Docker and HTCondor to link high-throughput computing across the USGS

USGS scientists often face computationally intensive tasks. Examples include calibrating groundwater models, monitoring invasive species using genetic methods, and processing geospatial data. Several USGS centers have developed high-throughput computing facilities that address these needs and most of these facilities use HTCondor to run their computational pools. Our project will help centers by documenting how to connect HTCondor pools by “flocking” within the USGS. Additionally, we will help use HTCondor by developing tutorials on how to “sandbox” code using Docker within USGS for use with high-throughput computing. To date, we have posted our tutorials on the CDI BitBucket page (https://my.usgs.gov/bitbucket/projects/CDI/repos/hunting_invasive_species_with_htcondor/) and have flocked three USGS Centers.

Science Support Framework Category: Communities of Practice

Authors: Richard Erickson <rerickson@usgs.gov> (USGS UMESC), Sunnie McCalla <smccalla@usgs.gov> (USGS UMESC), Michael Fienen <mnfienen@usgs.gov> (WI WSC)