Mar 23, 2022 | Pangeo + Globus Labs Meeting

Attendees: Charles Stern Rachana Ananthakrishnan Ryan Abernathey Ian Foster Rick Wagner Julius Busecke

- Something came up - I won't make it - Max

Agenda

I. Quick introduction to all participants (< 5 min)
II. Presentation from Globus Labs (10 min)
    A. Four core capabilities
        1. Data transfer
        2. Data sharing
        3. Research data search
        4. Research task orchestration: managed service
    B. Hosted service and platform: APIs (https://docs.globus.org/api/), Web UI, python SDK, federated authentication
    C. Globus connect: an agent (must be installed at facilities)
        1. https://www.globus.org/globus-connect
    D. Data access protocols:
        1. Gridftp for high-performance bulk transfers
        2. HTTPS
    E. Modern data research platform (https://mrdp.globus.org/)
    F. Including Foundry
III. Presentation from Pangeo (Ryan) (10 min)
    A. Including Pangeo Forge
IV. Open discussion of collaboration points
    A. CMIP6
    B. How can Pangeo Forge use globus to transfer data into the cloud?
        1. Trying to follow these instructions; stuck on the step "1. Find a Globus Connect Server that supports the Google Cloud Storage connector"
    C. How can Pangeo Forge access data stored on university HPC systems via Globus?
        1. Recommended solution is here; requires GCSv5 public collections
        2. That appears to be cost-prohibitive for many institutions (e.g UC-Irvine) https://github.com/pangeo-forge/staged-recipes/issues/100#issuecomment-1073772869
    D. Synergies btw. Pangeo Forge and Foundry?

V. [YOUR IDEA GOES HERE]
VI.
VII.

Fun link: ESGF replication LLNL->{ANL,LLNL}: https://dashboard.globus.org/esgf/
Space to watch: Globus-JupyterLab interface in development