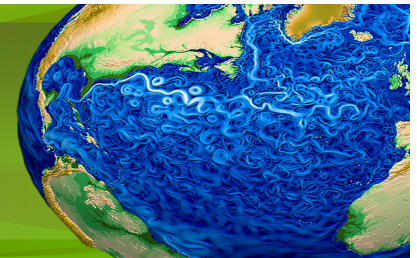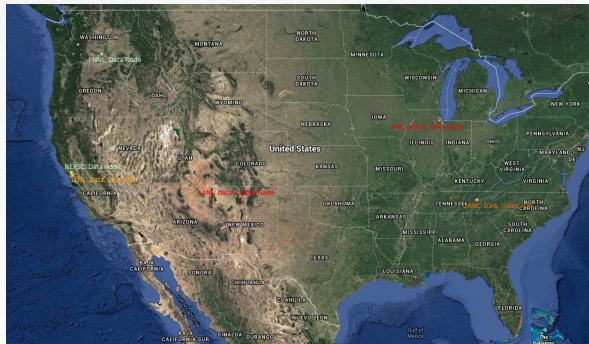# Earth System Grid Federation

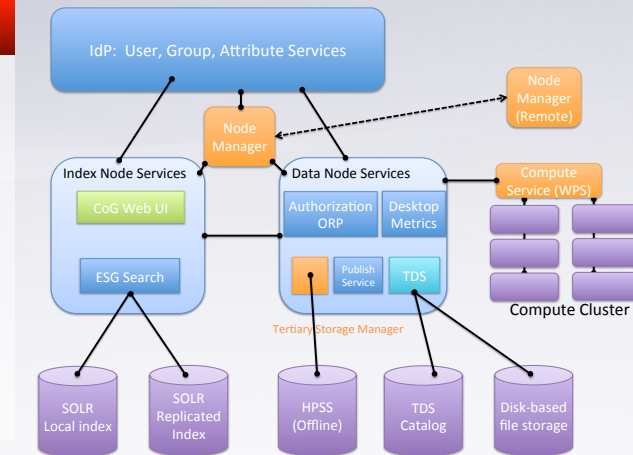Sasha Ames, Dean Williams (LLNL), and the international ESGF team

Abstract

Problem: ACME needs a means to manage and disseminate data.

Solution: The Earth Systems Grid Federation (ESGF) is a data management platform designed for the climate science community. The Federation is comprised of a collection of nodes running the ESGF software stack at various sites. Each node supports categorical publishing of data sets, attribute-based search for stored data, and access control. Additionally, the software supports inter-node coordination. ACME has been using ESGF to make its simulation data available. The ACME site is transitioning from the previous "front-end" user interface to CoG: the latest Web-based content delivery system adopted by ESGF. Some other developments to ESGF that will directly benefit ACME include: HPSS archive access; server-side and cluster analysis capabilities; an improved "node-manager" for ACME node coordination.
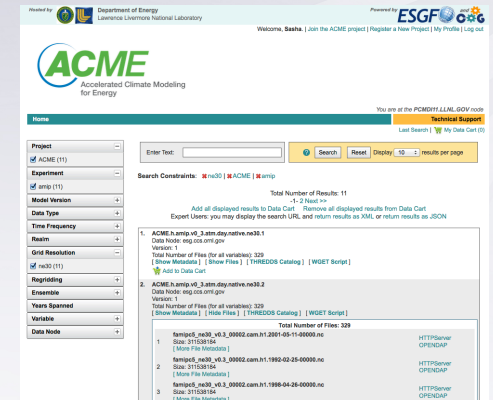
ESGF Nodes will be hosted at six of the labs participating in ACME. ORNL's node serves ACME data today. LLNL, as a major ESGF site, serves as an IDP for other ACME nodes. PNNL, NERSC and ANL have had previous ESGF installations. We assign different node types to balance project needs and federation scalability.
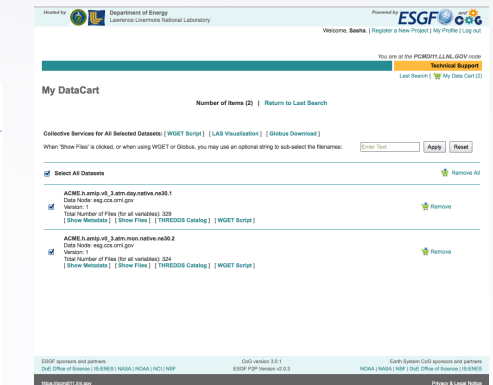
IdP: User, Group, Attribute Services

Node Manager

Node Manager (Remote)

Index Node Services
- CoG Web UI
- ESG Search

Data Node Services
- Authorization ORP
- Desktop Metrics
- Publish Service
- TDS

Compute Service (WPS)

Compute Cluster

Tertiary Storage Manager

- SOLR Local index
- SOLR Replicated Index
- HPSS (Offline)
- TDS Catalog
- Disk-based file storage

- Existing Services
- Updated Services
- New Service
- Future Services
- Third-party, store, or external component

The CoG web user interface enables ESGF users to search for datasets based on facets (left), browse results down to individual files, and select methods for downloads

Data Access Login

The following URL requires authentication:
http://esg.ccs.ornl.gov/thredds/fileServer/ACME/h/amip/v0_3/atm/mon/native/ne30/1/famipc5_ne30_v0.3_00001.cam.h0.1979-03.nc

Please enter your OpenID and you will be redirected to the login page at that site

Status: not logged-in

OpenID https://esg.ccs.ornl.gov/esgf-idp/openid/sashakames   GO

☑ Remember my OpenID on this computer

After logging in, you will be redirected to:
http://esg.ccs.ornl.gov/thredds/fileServer/acme_dataroot/ACME/h/amip/v0_3/atm/mon/native/ne30/1/famipc5_ne30_v0.3_00001.cam.h0.1979-03.nc

ESGF access controls require user logins and group membership in order to access project data sets, per node configuration.

ESGF / CoG contains a "data-cart" feature, which allows for batch downloads of multiple data sets, with automatic resume after failure.