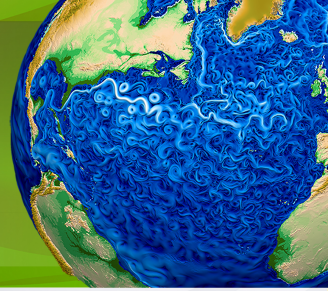


R: Publication Service for ACME

Rachana Ananthkrishnan, Lukasz Lacinski
Argonne National Lab



Objective

For users:

- Make data publication accessible to scientists via an intuitive easy interface
- Provide a fire and forget flow that allows scientists to submit data for publication, check status and be informed of progress and errors

For administrators:

- Facilitate publication without providing users escalated privilege on machines
- Reduce overhead of all publication tasks via an administrator of the site

For integration:

- Provide programmatic access to the capability for use cases where automation and integration with other tools such as workflow is needed

Approach

- A service and web application that wraps the various steps involved in data publication and makes it remotely accessible
- Users can use browse to publish data using the service- A publication workflow that walks the user through the various steps:
 - Providing any metadata as user input
 - Determining search facets and values for the published data
 - transfer of data to the ESGF data node
- On submission, automated process to extract metadata, generate THREDDs catalogs and publish to the ESGF search service
 - User can get status information and view progress any time during the process
 - User is send an email when errors are encountered and when publication is completed
- REST service that provides an API for integrating with publication service
- Python client to the publication service that can be used for scripting.

Impact

- Scientists have successfully published over 19TB of data (AMIP ne30 and ne120) from runs at NERSC and Argonne to Oak Ridge ESGF node without any administrative overhead at the sites.
 - They did not have to learn file system specific organization, and just specified facets on the web interface
 - Globus Transfer integrated with publication provided a simple mechanism to transfer files over to the ESGF data node
- REST API enables integration of publication as a step in the model run workflow via Pegasus
- REST API with status enables integration with ACME webapp/GUI for status information

Publication Service is available at <https://acme.globuscs.info/>

1. Choose collection and ESGF node to publish data to.

2. Choose facets needed for publication from drop down options.

3. Choose folder/files to publish and start publication.

4. Check status by logging in. User will be notified via email when publication is completed.