GHRC Cloud Migration

Will Ellett

IT/Operations Manager

Cumulus Prototype

Participated in Cumulus Prototype in October 2016

- Cumulus is a lightweight, cloud optimized framework for data ingest, archive, distribution and management
- Exploratory process to help ESDIS better understand DAAC requirements
- Used GHRC field campaign and mission datasets to prototype and test ingest, processing and archive workflows
- GHRC provided findings & recommendations to ESDIS at the end of the prototype evaluation period

Cloud Migration

GHRC Cloud Migration Effort began in November 2017

- Attended quarterly planning at GSFC
- Developers participated in hands-on cloud training
- Focused on learning Cumulus in AWS environment
- Developed sample workflows with GHRC datasets in NGAP sandbox environment
- Collaborated with the Cumulus core development team

Cloud Migration

Identified as Pathfinder for operating DAAC in cloud

- Created drafts of documents for publishing using Cumulus
- Onboarded all public GHRC datasets in Operational Earthdata Cloud
- Created Glacier Granule Restore Tool to streamline onboarding from AWS Glacier using Cumulus
- Performed Egress testing
- Updated security and developer documentation

Cloud Operations

GHRC began parallel operations on-prem and cloud on October 1st, 2019

- Publish datasets on-premise and in cloud
- Provide cloud training for Operators and Data team
- Perform cloud dataset validation
- Implement cloud disaster recovery solution
- Prepare Operational readiness review for cloud-only operations
- Prepare for GHRC's first cloud-only dataset: IMPACTS

Lessons Learned

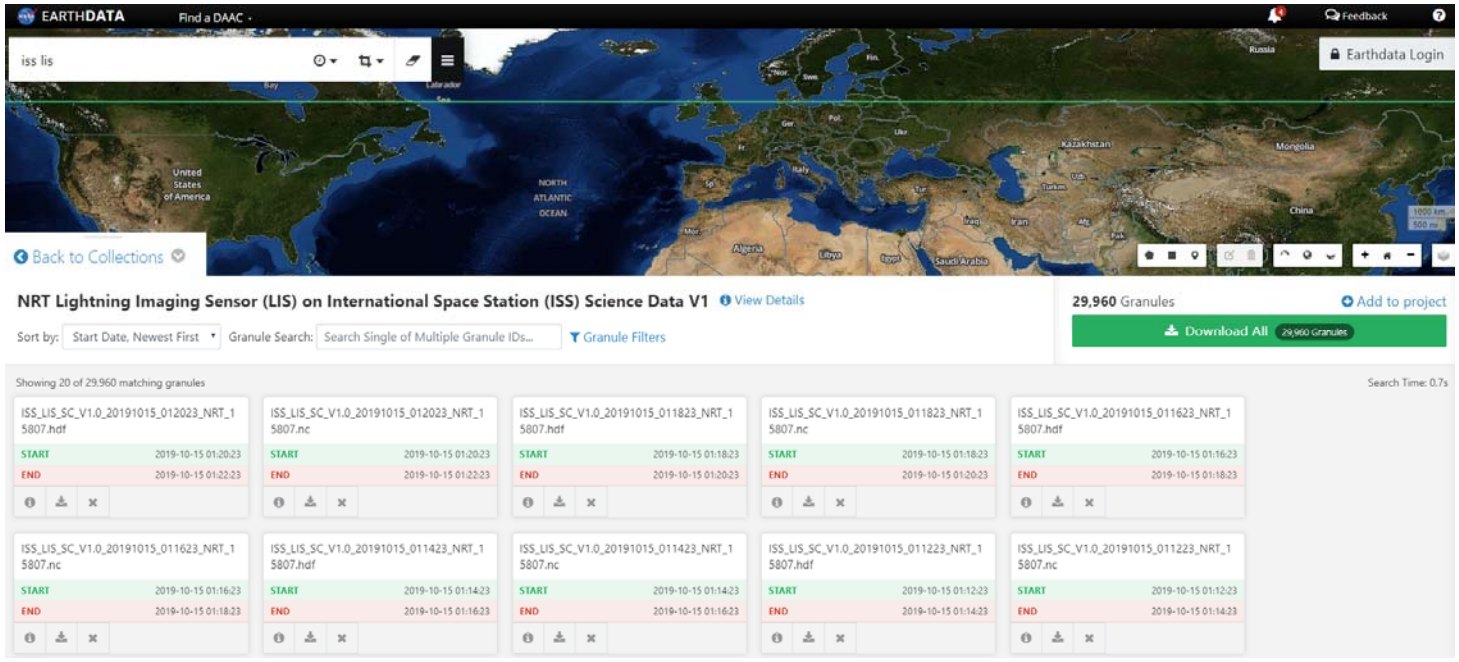
- Quarterly planning established by ESDIS provided insights that helped improve GHRC processes
- Large learning curve for Cumulus onboarding
- Developers & Operators need to understand AWS environment
- Cumulus Cloud Operator experience needs improvement
- Need improved data publication tool
- Cumulus supports code reuse
- Pros and Cons of updating to latest Cumulus
- Cloud transport layers need improvement

Earthdata Pub

- Build a reusable framework that provides a consistent cross-DAAC experience for data providers
- Support for cloud data products
- Track publication process and data provider interactions
- Build a reusable framework that enables DAACs to customize the publication process
- Create a consistent experience for data providers across DAACs by using common vocabulary and visual appearance
- Earthdata Pub Working Groups: (Led by Justin Rice)
  - Technical: GHRC, ORNL
  - Info: ASDC, GHRC, GES DISC, ORNL, NSIDC, PO.DAAC

Demos

GHRC Cloud Data in Earthdata Search



THANK YOU!

QUESTIONS?