NASA’S EARTH OBSERVATION DATA

FOSS4G 2018, Dar es Salaam
Introduction

Aimee Barciauskas
Data Engineer at Development Seed
@_aimeeb
aimee@developmentseed.org
Team of engineers and designers HQ in DC, distributed across the globe
Building tools to ingest, process and store earth observation data
Using earth observation data, machine learning and human mappers to solve development challenges
SatSummit 2018
Satellite data for global development

ABOUT
SatSummit convenes leaders in the satellite industry and experts in global
Discount code "ExploreSatSummit" for 20% off
Some earth observations

● Dar has beautiful beaches…
● But bad traffic
Some earth observations

● Dar has beautiful beaches…
● But bad traffic
● People see the world differently
FOSS4G Takeaway

● **Objective + subjective** earth observations = Earth insights

● **Remote + human** earth observations = Earth insights

● **Satellites + field data** earth observations = Earth insights
EOSDIS BIG DATA EVOLUTION

- NASA is gathering **new data from upcoming NISAR and other satellite missions** to better understand climate change.
- These new satellites will **collect data using Cumulus**.
- Cumulus is software for building and configuring **data ingest, process, and archive workflows** in the AWS cloud.
- NASA has a number of ways for you to get involved.
What is NISAR?
What is NISAR?

- NISAR = NASA-ISRO Synthetic Aperture Radar
- A satellite being launched in 2022
- A joint mission between NASA and the Indian Space Research Organisation
NISAR’s Key Science Objectives

- Interaction between the ice sheets, sea ice and climate change
- Determine likelihood of volcanic eruptions, earthquakes and landslides
- Aid future resource and hazard management
NASA Earth Science Missions: Present through 2023

- MAIA (~2021)
- TROPICS (~2021)
- EVM-2 (~2021)
- Landsat 8 (USGS) (>2022)
- Suomi NPP (NOAA) (>2022)
- CloudSat (~2018)
- Aqua (>2021)
- Terra (>2021)
- Landsat 7 (USGS) (~2022)
- GRACE-FO (2) (2018)
- ICESat-2 (2018)
- CYGNSS (2016)
- ISS
- SORCE, (2017)
- TCTE (NOAA)
- NISTAR, EPIC (NOAA's DISCOVR)
- QuikSCAT (2017)
- JPS S-2 (NOAA)
- RBI, OMPS-Limb (2018)
- PACE (2022)
- SWOT (2021)
- TEMPO (2018)
- LWPSC-B (2021)
- Landsat 9 (2021)
- InVEST – In-Space Validation CubeSats:
- RAVAN (2016)
- HARP (2016)
- IceCube (2016)
- MiRaTA (2017)
- RainCube (2017)
- TEMPEST-D (2018)
- CIRES (2018)
- CubeRRT (2018)
- LMPC (TBD)

Earth Science Instruments on ISS:
- RapidScat, (2017)
- CATS, (2020)
- LIS, (2016)
- SAGE III, (2016)
- ECOSTRESS, (2019)
- GEDI, (2018)
- OCO-3, (2018)
- CLARREO-PF, (2020)
- TSIS-2 (2020)
EOSDIS Big Data Evolution
2017 → EOSDIS ingests 3.9 PB / year
2022 → NISAR produces 7.2 PB / year
2022 → EOSDIS ingests 47.7 PB / year
>10x increase
NASA’s Cloud Initiative
“To prepare for this tremendous growth and efficiently provide access to these data, EOSDIS ... is testing how EOSDIS data collections can be archived collectively and disseminated in the cloud.

As befitting the cloud environment, this prototype is called Cumulus.”

https://earthdata.nasa.gov/eosdis-cumulus-project
What is Cumulus?
Cumulus is a collection of resources for deploying and configuring a data pipeline in the cloud.
Cumulus Major System Components

A lightweight framework consisting of:

**Tasks** a discrete action in a workflow, invoked as a Lambda function or EC2 service, common protocol supports chaining

**Orchestration engine** (AWS Step Functions) that controls invocation of tasks in a workflow

**Database** store status, logs, and other system state information

**Workflows(s)** file(s) that define the ingest, processing, publication, and archive operations (json)

**Dashboard** create and execute workflows, monitor system
Who uses Cumulus and how?
DAACs

DAACs are NASA’s Distributed Active Archive Centers. They ingest and disseminate NASA’s satellite data.

DAACs is now using Cumulus to discover, ingest, process and manage their data products.
Discipline-oriented Distributed Active Archive Centers (DAACs)

DAACs ingest, archive, process and distribute data to users

www.nasa.gov
Everyone!

- Cumulus is open source
- Cumulus Core Repository → https://github.com/nasa/cumulus
- Cumulus documentation → https://nasa.github.io/cumulus
  - How to deploy
  - Example workflows
How can you get involved?
High Temporal Resolution Air Quality Observations from Space

Tuesdays, September 4-25
12:00-13:00 EDT (UTC-4)

Register Now
Welcome to NASA Solve!

Interested in helping NASA solve tough problems? You are in the right place! This one-stop-shop website is where you’ll find opportunities to participate in challenges, prize competitions, and citizen science activities that develop solutions for problems related to NASA’s mission. So, jump in and become a NASA SOLVE-r!
Welcome to NASA DEVELOP

Integrating NASA's Earth observations to help meet the challenges of environmental change and improve life on our planet.

Fresh updates and events

VIEW ALL POSTS

AUG 01, 2018
DEVELOPers of the Term
Congratulations to Harrison Kepp from NASA Goddard Space Flight Center for being selected as a DEVELOP Thanksgiving Scholarship recipient.

AUG 01, 2018
SSAI Scholarship Award Recipients

AUG 01, 2018
VIRTUAL POSTER SESSION

DEVELOP VPS Summer Competition!
How can you get involved?

● NASA ARSET
● NASA SOLVE
● NASA DEVELOP
● AWS Earth on AWS
Thanks!

Aimee Barciauskas
Development Seed
@_aimeeb
aimee@developmentseed.org